Labor Market Matching, Wages, and Amenities

Thibaut Lamadon Jeremy Lise Costas Meghir
Jean-Marc Robin*

December 30, 2025

Abstract

This paper develops an equilibrium labor market model that jointly incorporates worker heterogeneity, firm heterogeneity, compensating differentials, search frictions, and Becker–type complementarities. We show that the primitives of the model are nonparametrically identified using matched employer–employee data. Unobserved heterogeneity can be recovered in the first stage in a fully model consistent way, making estimation very tractable. We apply the framework to Swedish administrative data and find substantial heterogeneity in worker preferences over firms, large non-pecuniary contributions to wage dispersion, and meaningful but imperfect sorting. Our estimates provide a structural interpretation of observed wage premia, mobility patterns, and sorting patterns in the data.

*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. The authors thank IFAU (www.ifau.se) for access and help with 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. We gratefully acknowledge the support of the European Research Council (grant reference ERC-2020-ADG-101018130), the Institution for Social Policy Studies, the Cowles Foundation at Yale University, and the NSF Grant SES-1658920.

Introduction

Explaining wage determination and labor market flows is crucial for understanding both the nature of dispersion and the functioning of the labor market. This understanding can, in turn, help to formulate policy. The sorting of workers to firms and the consequent implications for wages has been understood even before Becker's seminal paper (Becker, 1973).[1] However, empirical work on sorting and its effects took off with the increasing availability of matched employer-employee data, which spawned a large literature. The importance of such work lies, in part, in the need to understand wage setting, where seemingly identical individuals receive different pay rates depending on where they work. In his influential lecture series, Mortensen (2003) identifies four key economic forces that contribute to the noticeably large cross-sectional wage dispersion 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 among employers. The third is the wage dispersion channel caused by market frictions, which grants some degree of local monopsony power to firms. The fourth is the endogenous matching of workers with firms in the presence of match complementarities, as in Becker (1973) and Sattinger (1993).

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

In this paper, we make three key contributions. First, we develop a theoretical equilibrium model of the labor market that encompasses all four channels. Second, we prove that the model's primitives are nonparametrically identified using matched employer-employee data. Specifically, while we assume discrete types, we do not impose any shape restrictions on match production or the disutility of labor. Our identification approach demonstrates that we can use the nonlinear estimator of Bonhomme, Lamadon, and Manresa (2019, hereafter BLM), which provides a purely statistical description of wages and transitions, to reveal the unobserved heterogeneity of workers and firms in the first stage in a fully model-consistent manner. We then illustrate

[1] Early papers addressing the assignment problem include Koopmans and Beckmann (1957) and Shapley and Shubik (1971).

how to non-parametrically recover the primitives of our structural model from the type-specific wages and mobility patterns estimated in the first stage. Third, the identification proof suggests an estimation strategy that we apply to analyze the sources of wage variation in Swedish administrative records.

Our first main empirical finding is that workers do not agree on a common ranking of firms. This indicates that both worker heterogeneity and differences in worker composition between firms are important factors in explaining mobility and wage dispersion between and within employers. In particular, while all workers agree on which firms pay the most, they disagree on which firms they prefer to work for. They make systematically different job-to-job transitions based on worker type. Our second finding is that non-pecuniary aspects of jobs significantly contribute to wage dispersion, accounting for half of the within-worker wage dispersion.

Our theoretical model integrates the theory of complementarities of Becker (1973), the presence of compensating differentials of Rosen (1986), and wage setting in a frictional labor market that includes on-the-job search of Postel-Vinay and Robin (2002, PVR). The economy consists 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 experiences a disutility of working (or, equivalently, enjoys an amenity). Both production and disutility are fully flexible functions of the worker and firm types, allowing for arbitrary complementarity or substitutability between chosen partners. Firms are risk-neutral, workers are risk-averse, and both parties are forward-looking. Meetings between firms and workers are constrained by search frictions. Both employed and unemployed workers search for jobs, and when they encounter a vacancy, they also draw a mobility preference shock.

Solving the equilibrium is more challenging than what is presented 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. However, by defining the match surplus as the highest present value that the firm can offer to the worker, we can characterize the equilibrium concisely. We first demonstrate that the optimal contract is a straightforward extension of matching outside offers of Postel-Vinay and Robin (2002), even when firm types are private information, while worker types and mobility costs are common knowledge. Second, we demonstrate that our surplus satisfies a surprisingly simple and intuitive equation that extends Postel-Vinay and Robin (2002). 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 we apply 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 remains relevant in our context. Importantly, different workers have different rankings 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, Holzheu, Lamadon, Manresa, Mogstad, and Setzler (2023), indicate that differences in firm premia account for between 5% and 15% of the cross-sectional variance in earnings, while the sorting of high-paid workers into 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.

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. A similar point has been made by Eeckhout and Kircher (2011) and Hagedorn, Law, and Manovskii (2017). Here, we consider the log-additive decomposition as an informative moment of the data; however, we utilize a different approach to identify and estimate worker and firm-types.

Our identification employs a series of steps and results. First, we prove that our framework is compatible with the assumptions of the nonlinear estimator of BLM. In contrast to the estimation of AKM, the types from the estimation of BLM correspond to the model types. Conditional on the estimated types, the estimated wages, mobility, and allocations of 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, 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 for the identification of surpluses in a broader sense. 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 an arbitrarily flexible mobility shock distribution, there is a very transparent interpretation in the case of the logistic distribution. In this case, we can use mobility alone from the BLM estimation stage to directly recover the match surplus (up to scale)[2]. Therefore, wages can be used to separate production from disutility (amenities) given mobility patterns. This is the strategy that we employ in estimation.

We use our model to investigate the underlying forces that lead to the observed sorting, including the potential roles of complementarities in production and amenities. We find that heterogeneity across workers accounts for the majority of the wage dispersion observed in the data; there is substantial positive sorting of workers across jobs; and just under half of the 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) markedly differs by worker type. It accounts for 12% of the variation in wages for workers of the highest type, but more than 65% for the lowest type. We estimate that the process of sorting workers into firms accounts for 27% of the wage dispersion between workers and for 20% of the overall wage dispersion. Finally, we use the model to interpret several well-documented empirical regularities reported in Card et al. (2013) and Di Addario et al. (2023).

Related literature. Our paper is related to several recent contributions. The first is Hagedorn, Law, and Manovskii (2017), which demonstrates the identification of

[2] This uses a revealed preference argument similar to Sorkin (2018), but we relax the assumption that all workers have common preferences. Arcidiacono et al. (2023) presents a search model with amenities and preference shocks that exploits insights from dynamic discrete choice estimation. Both of these models are partial equilibrium, and the shocks are not priced into wages.

the equilibrium frictional sorting model of Shimer and Smith (2000), wherein workers search only from unemployment. In this model, a worker's wage rank within a firm is equivalent to a worker's productivity rank within that firm. 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 the amenity values of 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). Interestingly, the identifying assumptions in Hagedorn et al. (2017) rule out the possibility of compensating differentials arising from amenities, while the assumptions in Sorkin (2018) rule out the possibility of worker-firm sorting. 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 does not necessarily order worker types within a firm due to compensating differentials from various 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 model's structure without the need to repeatedly solve and simulate, thereby offering additional transparency.

Lamadon, Mogstad, and Setzler (2022) proposes a static and frictionless model with several features that we study, including interactions in production and amenities. They demonstrate the fundamental role of interactions in amenities for understanding the distribution of firm size, wage premia, and sorting, while clarifying that the presence of firm effects could be driven by compensating differentials rather than market power. In that paper, there is a one-to-one mapping between amenities and firm size. It also remains silent on the differential mobility patterns of low- and high wage workers. In a dynamic context with search frictions, we address the key identification 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. Additionally, in our dynamic context, we can derive the elasticities of hires and separations with respect to firm wages, objects that are prevalent in the empirical literature but are not defined in a static model.

Unlike the papers discussed above, we derive the optimal contract offered by risk-neutral firms to risk-averse workers. The approach shares features with Balke and Lamadon (2022) and provides a comprehensive microfoundation for the types of contracts assumed in Postel-Vinay and Robin (2002). It extends to a much richer environment that includes non-wage attributes of jobs and mobility cost shocks.

In contemporaneous work, Borovickova and Shimer (2024) proposes an elegant model that extends Shimer and Smith (2000) by adding a match-specific shock to production. The selection induced by the match shock allows the model to successfully reproduce the wage and sorting patterns estimated by BLM. While their model does successfully replicate these cross-sectional observations, it is unable to generate job-to-job mobility accompanied by a wage cut, which is also a feature of the BLM estimates and a well established empirical characteristic of worker transitions (Jolivet et al., 2006; Sorkin, 2018; Morchio and Moser, 2023). This feature of the data plays an important role in identifying the value that workers place on non-wage attributes of jobs.

Finally, Lentz, Piyapromdee, and Robin (2023) provides a statistical model based on BLM. They impose additional parametric restrictions on mobility to facilitate interpretation and propose a slightly different iterative estimation procedure than BLM. Their reduced-form analysis emphasizes the apparent disconnect between the wage and mobility processes. Our theoretical model provides a coherent interpretation of this data that is not possible with any existing equilibrium search model.

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

1 The labor market model

We consider a steady-state economy populated by heterogeneous workers and firms. Time is discrete. Workers and firms are forward-looking, discounting the future at a rate r, and have an infinite planning horizon.

1.1 Agents and states

There are X worker types indexed by $x = 1, \dots, X$. We denote by ℓ_x the measure of type x workers in the population, with the total measure normalized to one. There are Y firm types indexed by $y = 1, \dots, Y$. Let n_y denote the measure of jobs of type y, which may be vacant or matched to a worker, with total measure N. The measure of unemployed workers of type x is ℓ^0_x, and the measure of matches of type (x, y) is $\ell^1_{x,y}$, with $\ell^0_x + \sum_{y=1}^{Y} \ell^1_{x,y} = \ell_x$.

Firms hire workers and advertise vacancies for unfilled positions. Each firm may 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 across jobs. We denote by v_y the measure of job openings of type y and by $V = \sum_{y=1}^{Y} v_y$ the total measure of vacancies, with $n_y = v_y + \sum_{x=1}^{X} \ell^1_{x,y}$.

We assume that ℓ_x and n_y are exogenous, while v_y, ℓ^0_x, and $\ell^1_{x,y}$ are determined in equilibrium.

1.2 Timing of events

At the beginning of each period, a worker may either be unemployed or employed. The timing of events during the period 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 a probability of v_y/V. Upon meeting, the workers draw an instantaneous, one-time utility shock ξ from the distribution G^0 (which has negative support and density g^0). They will experience this shock only if they change state, specifically if they accept the job. We assume that firms make take-it-or-leave-it offers to unemployed workers in a manner that we will describe below.

Employed workers. A match between a worker of type x and a firm of type y produces $f_{x,y}$ during the period. Workers receive a wage w that they value at a utility flow of $u(w)$. They incur a deterministic flow cost of $c_{x,y}$ from providing labor in this job, net of any amenity value. We will use the terms disutility of labor and amenity interchangeably. Their net flow utility from a job is $u(w) - c_{x,y}$. Job costs or amenities are predetermined and set at the foundation of the firm, either as fixed characteristics or due to 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 in that state until the following period. The match continues with probability $\overline{\delta}_{x,y} \coloneqq 1 - \delta_{x,y}$, and the worker contacts an alternative vacancy with probability λ^1. This vacancy is also drawn from the cross section of vacancies, where a vacancy of type y' is drawn with probability $v_{y'}/V$. At the time of meeting, the worker also draws an instantaneous, one-off mobility shock ξ from the distribution G^1 (with full support and density g^1). The worker experiences the mobility shock if they accept the alternative job or choose to become unemployed but not if they remain in their current job. Unlike unemployed workers, who always experience cost shocks, mobility shocks for the employed may be negative or positive. 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 $L^0 = \sum_{x=1}^{X} \ell_x^0$, $L^1 = \sum_{x=1}^{X} \sum_{y=1}^{Y} \overline{\delta}_{x,y} \ell_{x,y}^1$, and κ represents the relative search efficiency of the employed workers. We define the equilibrium meeting rates for unemployed and employed workers as follows:

$$\lambda^0 = \frac{M}{L}, \quad \lambda^1 = \kappa \frac{M}{L}, \quad \text{and} \quad \overline{\lambda}^j = (1 - \lambda^j). \tag{1}$$

Similarly, a vacancy meets an unemployed worker with probability $\lambda^0 L^0 / V$, drawing a type x from the distribution ℓ_x^0 / L^0. It meets an employed worker with probability $\lambda^1 L^1 / V$, drawing from the cross-sectional distribution of matches $\ell_{x,y}^1 / L^1$.

Further discussion of types and shocks. Worker and firm types are assumed to be discrete, with production, disutility, and the exogenous separation rate differing arbitrarily across (x, y) matches. This representation offers substantial flexibility. It encompasses scenarios where worker types are characterized by multiple traits, with production, disutility, and separations varying according to these traits. For example, consider a type x defined by a vector of K characteristics (e.g., analytical skill, verbal skill, strength, dexterity, empathy, charisma, self-esteem, health status, etc.). Moreover, subsets of these characteristics (which may overlap) influence production, disutility, and the probability of separation in different ways across firm types. In this paper, we focus the analysis at the type level, which is sufficient for modeling mobility and wage determination. We do not explicitly define types in terms of characteristics, partly because most matched employer-employee datasets lack detailed measurements of the various skills of workers and the task requirements of jobs.[3]

[3] For the purpose of plotting the figures in Section 4, we need to choose a label for the types. We follow BLM and label worker types according to the average wage received across all jobs, while firm types are labeled by the average wage they pay to all workers. In this discussion, we tend to refer to high wage workers and high wage firms as high types. Note that this ordering will not necessarily correspond to the ordering of any single characteristic that defines a type. Indeed, as pointed out by Lindenlaub and Postel-Vinay (2023), it is generally not possible to create a label that is monotonic with respect to all the characteristics that define a type. That said, the equilibrium allocations and wages can be summarized by types and do not require a detailed specification of the characteristics. For this reason, the average wage is a natural label to adopt when considering the allocations of worker types to firm types. There are certainly cases in which researchers would be interested in how the different characteristics of a type affect mobility and wages. For example, Lise and Postel-Vinay (2020) estimate an equilibrium labor search model using a combination of NLSY and O*Net data, explicitly representing worker (firm) types by a vector of cognitive, manual, and interpersonal skills

Match production $f_{x,y}$ is a well-defined concept in principle; however, it is generally not directly measurable with data, as we only observe the sum over match outputs within a firm. Match specific disutility $c_{x,y}$ encompasses various difficult-to-measure factors, such as unpleasantness, difficulty, pain, stress, and the physical demands of a job, along with location-specific aspects like weather and city size. Moreover, there is no reason to believe that different worker types would share similar perceptions of disutility across firm types, partly because they are likely to perform very different tasks. For instance, an accountant would engage in roughly the same tasks at a large coffee chain as they would at a coal mine. Conversely, an unskilled laborer might be sweeping up at the coffee shop while shoveling coal down the shaft at the mine.

Finally, the mobility shock ξ encompasses all the idiosyncratic aspects of the job-changing process that are not directly modeled. This includes the monetary costs of physically relocating and the capital gain or loss in the housing market associated with changing locations. It also captures the desire to relocate in order to cohabit with a spouse who has already accepted a job elsewhere, as well as the momentary satisfaction of informing a supervisor about the resignation. Heuristically, mobility shocks ξ will lead to idiosyncratic job-to-job transitions associated with a wage cut, while match-specific disutility/amenities $c_{x,y}$ will generate systematic differences across worker types regarding the likelihood of job-to-job transitions associated with a wage cut, both conditional on match production $f_{x,y}$.

1.3 Competition, hiring, and wage determination

Define W_x^0 as the value of an unemployed worker and Π_y^0 as the value of a vacancy. An employment contract promises a value W to the worker, where W must be strictly greater than W_x^0 for employment to be preferred over unemployment. We refer to the difference $R = W - W_x^0 \geq 0$ as the surplus for the worker or the value of the contract.

Define $\Pi_{x,y}^1(R)$ as the value of profit that a firm can achieve when employing a 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 $R = 0$. Whenever $\Pi_{x,y}^1(0)$ 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 we 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

(task requirements) and parametrically specifying the interactions in production.

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, meets another firm of type y' and draws a mobility shock ξ. The worker now has the possibility of making a move and incurring the shock 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.[4]

Definition 1. *The poaching mechanism.* The incumbent and the poaching firms report their reservation surpluses, namely B and B', which are not necessarily equal to $S_{x,y}$ and $S_{x,y'}$. There are three possible mobility outcomes: either the worker moves to y', stays with the incumbent employer, or becomes unemployed. The mechanism prescribes the following policy:

1. If $B' + \xi > \max\{B, \xi\}$, the poacher wins and must deliver the surplus of $\max\{B, \xi\} - \xi$ to the worker, who receives a total of $\max\{B, \xi\}$ following 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. This means that, even in cases where the incumbent is certain to lose, there must exist a feasible strategy to deliver B to the worker, i.e., $B \leq S_{x,y}$.

1.4 Value functions

The firm's problem The employer faces two distinct decision problems. First, it must choose a bid $B(\xi)$ in the event that the worker encounters a vacancy and draws a mobility shock ξ. Second, for a specified promised value R, it must determine how to implement it through wages and separations.

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 B', which will be characterized later.

Consider a firm type y employing a worker type x with a promised surplus R. In calculating its value $\Pi^1_{x,y}(R)$, the firm anticipates all future events. In the event of poaching at the end of the current period, the firm assumes that the poacher will adhere

[4] See Online Appendix B.1 for an alternating offer interpretation of this mechanism.

to the mechanism. Therefore, if $B(\xi) < B'^{+} + \xi$ the poacher wins, the incumbent firm receives Π^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. This payment is a minimum because it may be below the promised surplus R_1. For any promised worker surplus $R \leq S$, the firm's problem is:

$$\begin{aligned}\Pi^1_{x,y}(R) = \max_{\{w, R_0, R_1(B',\xi), B(\xi)\}} \Big\{ & f_{x,y} - w + \frac{\delta_{x,y}}{1+r}\Pi^0_y + \frac{\overline{\delta}_{x,y}\overline{\lambda}^1}{1+r}\Pi^1_{x,y}(R_0) \\ & + \frac{\overline{\delta}_{x,y}\lambda^1}{1+r} \iint \Big[\mathbf{1}\{B(\xi) \geq B'^{+} + \xi\}\Pi^1_{x,y}(R_1(B',\xi)) \\ & + \mathbf{1}\{B(\xi) < B'^{+} + \xi\}\Pi^0_y \Big] \, \mathrm{d}F^0(B'|x,\xi)\, \mathrm{d}G^1(\xi) \Big\}, \end{aligned} \tag{2}$$

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

$$\begin{aligned} R = u(w) - c_{x,y} + \frac{\overline{\delta}_{x,y}\overline{\lambda}^1}{1+r}R_0 + \frac{\overline{\delta}_{x,y}\lambda^1}{1+r} \iint & \Big[\mathbf{1}\{B(\xi) \geq B'^{+} + \xi\} R_1(B',\xi) \\ & + \mathbf{1}\{B(\xi) < B'^{+} + \xi\} \max\{B(\xi), \xi\} \Big] \, \mathrm{d}F^0(B'|x,\xi)\, \mathrm{d}G^1(\xi), \end{aligned} \tag{PK}$$

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

$$R_0 \geq 0, \text{ and } \Pi^0_y \leq \Pi^1(x,y,R'), \text{ for } R' \in \{R_0, R_1, B\}, \tag{C}$$

where $\frac{1}{1+r}$ is the discount factor corresponding to the discount rate r, and $\mathbf{1}\{\cdot\}$ is an indicator function equal to 1 if the statement is true and zero otherwise.

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 this case, the firm recovers a vacancy in the next period, which has a value of Π^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, as well as the worker surplus $R_1(B',\xi)$ to deliver when it retains the worker. Finally, it chooses R_0 for the scenario in which no external contact is made. Each of these decisions is conditional on the state (x, y, R).

When making these choices, the firm is subject to the promise-keeping constraint (PK). This states that the contract $\{w, R_0, R_1(B',\xi), B(\xi)\}$ needs to provide the worker with the promised surplus R. The right-hand side of (PK) represents the surplus that the worker gains from this contract. The worker receives the flow utility $u(w) - c_{x,y}$. Then, with probability $\delta_{x,y}$, the worker is laid off; in this 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 a 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 the amount required by the mechanism. Finally, (C) represents the worker's participation constraint, the firm's limited commitment, and the credibility of the bid.

The value of unemployment. We assume that employers have full monopsony power when they encounter an unemployed worker, which, in our auction mechanism, is equivalent to being with a firm that has zero surplus. When an unemployed worker of type x meets a vacancy of type y, it draws a negative mobility cost shock ξ from G^0. Employers must compensate the worker for this cost in order to hire them. There are two possible scenarios. 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 mobility cost. The worker receives a net surplus of zero. If $S_{x,y} + \xi \leq 0$, no offer is made, and the worker remains unemployed, also receiving zero surplus. The annuity value of unemployment for a type-x worker is then simply

$$rW_x^0 = (1 + r)b_x.$$

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

$$r\Pi_y^0 = \max_{\{B_x^0(\xi), B_x^1(\xi)\}} \frac{\lambda^0 L^0}{V} \sum_{x=1}^{X} \frac{\ell_x^0}{L^0} \int \mathbf{1}\{B_x^0(\xi) + \xi > 0\} \left[\Pi_{x,y}^1(-\xi) - \Pi_y^0\right] \mathrm{d}G^0(\xi)$$
$$+ \frac{\lambda^1 L^1}{V} \sum_{x=1}^{X} \iint \mathbf{1}\{B_x^1(\xi) + \xi > B'\} \left[\Pi_{x,y}^1(B' - \xi) - \Pi_y^0\right] \mathrm{d}F^1(B', x|\xi) \, \mathrm{d}G^1(\xi). \quad (3)$$

$F^1(B', x|\xi)$ represents the distribution of workers and retention bids conditioned on the mobility shock. When the worker is unemployed, the implicit retention bid B' will be 0. When the worker is employed, the firm faces a distribution of retention bids derived from the cross-sectional distribution of employed workers, along with the reporting strategy of each of these firms. The vacant firm selects the values to report to the auction, denoted $B_x^0(\xi)$ and $B_x^1(\xi)$. We adopt the minimum value constraint from the auction as the promised value at the beginning of the match, as it satisfies the worker's participation constraint. Unlike the incumbent firm, there is no preexisting promised value. Whenever the firm wins an auction, it realizes a capital gain of $\Pi_{x,y}^1(R) - \Pi_y^0$.

Theorem 1. *Assume that the utility function u is bounded, twice continuously dif-*

ferentiable, 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 it is differentiable almost everywhere.*
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* $B'^{+} + \xi \le S_{x,y}$ *.*

Firms have a reservation value $S_{x,y}$ that is equal to the contract surplus that generates minimal profit Π^0_y. The incumbent firm is truthful and bids exactly its reservation value. It will provide a constant value to the worker over time until a meeting with an outside firm necessitates increasing this 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 poaching firms.

1.5 Equilibrium

In a stationary truth-telling equilibrium, the inflows and outflows of $\ell^1_{x,y}$ are equal:

$$\ell^1_{x,y}\left(\delta_{x,y} + \overline{\delta}_{x,y}\lambda^1 \sum_{y'=1}^{Y} \overline{G}^1\left(S_{x,y} - S_{x,y'}\right)v_{y'}\right)$$
$$= \lambda^0 \frac{v_y}{V}\overline{G}^0(-S_{x,y})\ell^0_x + \lambda^1 \frac{v_y}{V}\sum_{y'=1}^{Y} \overline{G}^1\left(S_{x,y'} - S_{x,y}\right)\overline{\delta}_{x,y'}\ell^1_{x,y'}. \quad (4)$$

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

$$F^0(B'|x,\xi) = \sum_{y'=1}^{Y} \mathbf{1}\{S_{x,y'} \le B'\}\frac{v_{y'}}{V}, \quad F^1(B',x|\xi) = \sum_{y'=1}^{Y} \mathbf{1}\{S_{x,y'} \le B'\}\overline{\delta}_{x,y'}\frac{\ell^1_{x,y'}}{L^1}. \quad (5)$$

This states that employed workers receive outside bids from the cross-section of vacancies, vacancies draw from the cross-section of employed workers who have not been laid off, and all firms submit truthful bids.

Definition 2. A *stationary search equilibrium with sequential auctions* is defined by the following components: meeting probabilities λ^0 and λ^1, employment measure $\ell^1_{x,y}$, bid distributions $F^0(B'|\xi, x)$ and $F^1(B'|\xi, x)$, and firm value functions Π^0_y and $\Pi^1_{x,y}(R)$, along with their corresponding policies, such that:

1. The meeting probabilities λ^0, λ^1 are consistent with the meeting technology (1).
2. Taking F^0 and F^1 as given, $\Pi^1_{x,y}(R)$ and Π^0_y solve equations (2) and (3), which take into account the mobility decisions of the workers.
3. The policies of $\Pi^1_{x,y}(R)$ and Π^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 (4) and $F^0(B'|x, \xi), F^1(B', x|\xi)$ are generated by (5).

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. At equilibrium, the value function for an employed worker can be inverted to obtain a wage equation in terms of the current surplus R and the maximum surplus $S_{x,y}$:

$$\begin{aligned} u(w_{x,y}(R)) = c_{x,y} + \frac{r + \delta_{x,y}}{1+r} R + \frac{r}{1+r} W^0_x \\ - \frac{\lambda^1 \overline{\delta}_{x,y}}{1+r} \sum_{y'=1}^{Y} \Bigg[\int_{R - S^+_{x,y'}}^{S_{x,y} - S^+_{x,y'}} \left(S^+_{x,y'} + \xi - R \right) g^1(\xi)\, \mathrm{d}\xi \\ + \int_{S_{x,y} - S^+_{x,y'}}^{\infty} \left(\max\{\xi, S_{x,y}\} - R \right) g^1(\xi)\, \mathrm{d}\xi \Bigg] \frac{v_{y'}}{V}. \end{aligned} \tag{6}$$

The wage is increasing in the promised surplus R. For a given R, the wage increases with the disutility of labor, net of amenities $c_{x,y}$, and decreases with 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 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 a match (x, y) is

$$\overline{w}_{x,y} = w_{x,y}(S_{x,y}) = f_{x,y} - \frac{r}{1+r} \Pi^0_y. \tag{7}$$

We also know that, at this point, the worker is receiving maximum surplus. We can substitute equation (7) and $R = S_{x,y}$ into the wage equation (6) to obtain the following equation for maximum worker surplus $S_{x,y}$:

$$(r + \delta_{x,y})\, S_{x,y} = (1+r) \left[u(\overline{w}_{x,y}) - c_{x,y} \right] - r W^0_x + \lambda^1 \overline{\delta}_{x,y} \int_{S_{x,y}}^{\infty} \overline{G}^1(\xi)\, \mathrm{d}\xi. \tag{8}$$

The flow value of the maximum worker surplus comprises the flow utility $(1 + r)\left[u(\overline{w}_{x,y}) - c_{x,y}\right]$ minus the unemployment annuity rW_x^0, plus the expected value of $\max\{\xi - S_{x,y}, 0\}$ (the last term). If the mobility shock ξ is greater than the current surplus $S_{x,y}$, the value of the worker's outside option when moving is ξ, not the incumbent's reservation value $S_{x,y}$.

We will also refer to $S_{x,y}$ as 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 definition of the surplus of a match that is independent of how 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, and their job valuations are expressed in units of utility.

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

$$\Pi_{x,y}^1(R) = \Pi_y^0 + \int_R^{S_{x,y}} \frac{\mathrm{d}R'}{u'(w_{x,y}(R'))}. \tag{9}$$

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

$$r\Pi_y^0 = \frac{\lambda^0 L^0}{V} \sum_{x=1}^{X} \mathbf{1}\{S_{x,y} > 0\} \frac{\ell_x^0}{L^0} \int_{-S_{x,y}}^{0} \mathbf{1}\{\xi \in Supp(G^0)\} \frac{\overline{G}^0(\xi)}{u'(w_{x,y}(-\xi))}\, \mathrm{d}\xi$$
$$+ \frac{\lambda^1 L^1}{V} \sum_{x=1}^{X} \mathbf{1}\{S_{x,y} > 0\} \sum_{y'=1}^{Y} \frac{\overline{\delta}_{x,y'} \ell_{x,y'}^1}{L^1} \int_{S_{x,y'} - S_{x,y}}^{S_{x,y'}} \mathbf{1}\{\xi \in Supp(G^1)\} \frac{\overline{G}^1(\xi)}{u'(w_{x,y}(S_{x,y'} - \xi))}\, \mathrm{d}\xi. \tag{10}$$

In the next section, we will utilize these equilibrium properties 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. It includes amenities and mobility shocks that can, in principle, enable a complex structure of transitions and wage growth. However, the value of such a rich specification depends on our ability to identify the key components without relying on assumed parametric forms. This will define the empirical content of the model. We will 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 demonstrate 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 worker types ℓ_x and job types n_y, the distribution of the mobility shocks for the unemployed G^0 and the employed G^1, and parameters λ^0, λ^1.[5]

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 a latent type x_i, and each employer j is associated with a latent type y_j, but neither is observed by the econometrician. For each period t and individual i, we observe the employment status $e_{it} \in \{\text{U}, \text{E}\}$, 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 $m_{it} = \text{EE}$ if the worker remains employed with the same firm between t and $t+1$; $m_{it} = \text{EU}$ if the worker separates from a job and becomes unemployed; and $m_{it} = \text{JJ}$ if the worker changes employers. We denote $m_{it} = \text{UU}$ if the worker is unemployed during both t and $t+1$; and $m_{it} = \text{UE}$ if the worker transitions 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) regarding 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 presented in Section 1 generates a Markovian law of motion for wages and employment, conditional on worker type, as follows:*

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

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 encompasses the structural model. This will be one of the rare cases in which indirect inference can be proven to be consistent. Usually, the identification of a

[5] 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.

structural model is based on intuitive considerations regarding how a moment should help identify a parameter. However, the injectivity of the binding function linking 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 the second step, we will formally prove the identification of structural parameters from the auxiliary model. 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 estimates of unobserved heterogeneity, as it is difficult to separately identify different latent groups using aggregate moments. Hagedorn et al. (2017) was the first to propose an identification and estimation procedure that incorporates some of these ideas for the model in Shimer and Smith (2000).

Assuming a finite number of worker types K, and 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 $\mathbb{P}[m_t = m \,|\, x, y_t = y, e_t = \mathrm{E}] \coloneqq p_m(x, y)$ for $m = \mathrm{EU}, \mathrm{EE}, \mathrm{JJ}$, the transition probability $\mathbb{P}[y_{t+1}, m_t = \mathrm{JJ} \,|\, x, y_t] \coloneqq p_{\mathrm{JJ}}(y_{t+1}|x, y_t)$, the law of motion for within-job wages $\mathbb{P}[w_{t+1} \leq w \,|\, w_t, x, y_t, m_t = \mathrm{EE}] \coloneqq F_{\mathrm{EE}}(w|x, y_t, w_t)$, and the distribution of wages after a move $\mathbb{P}[w_{t+1} \leq w \mid x, y_{t+1}, y_t, m_t = \mathrm{JJ}] \coloneqq F_{\mathrm{JJ}}(w|x, y_t, y_{t+1})$ are nonparametrically identified from data on movers with two periods before and after the move.

We include unemployment as an additional 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 $\mathbb{P}[m_t = \mathrm{UE} \,|\, x, e_t = \mathrm{U}] \coloneqq p_{\mathrm{UE}}(x)$, 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 $\mathbb{P}[w_{t+1} \leq w|x, y_{t+1}, m_t = \mathrm{UE}] \coloneqq F_{\mathrm{UE}}(w|x, y_{t+1})$ (see the online appendix B.4 for complete details).

From this point onward, we can consider each of these distributions to be 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 $\phi_{x,y} = \mathbf{1}\{S_{x,y} > 0\}$. 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 increases,*

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

- *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 based on 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 (6) that differentiating the worker surplus $R_{x,y}(w)$ with respect to the wage w yields:

$$\frac{\partial R_{x,y}(w)}{\partial w} = \frac{(1+r)u'(w)}{r + \delta_{x,y} + \lambda^1 \overline{\delta}_{x,y} \sum_{y'=1}^{Y} \overline{G}^1 \left[R_{x,y}(w) - S^+_{x,y'}\right] \frac{v_{y'}}{V}}, \tag{11}$$

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

$$\begin{aligned}\delta_{x,y} + \lambda^1 \overline{\delta}_{x,y} \sum_{y'=1}^{Y} \overline{G}^1 \left[R_{x,y}(w) - S^+_{x,y'}\right] \frac{v_{y'}}{V} \\ = \mathbb{P}[m_t = \mathrm{EU} | x, y_t = y, e_t = \mathrm{E}] + \mathbb{P}[m_t = \mathrm{JJ} | x, y_t = y, e_t = \mathrm{E}] \\ + \mathbb{P}[m_t = \mathrm{EE}, w_{t+1} > w | x, y_t = y, w_t = w, e_t = \mathrm{E}].\end{aligned}$$

The probability of any of these changes, conditional on a match (x,y) , is known from Step 1 and is therefore 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 (11) for wages w in the support $[\underline{w}_{x,y}, \overline{w}_{x,y}]$:

$$R_{x,y}(w) = \int_{\underline{w}_{x,y}}^{w} \frac{\partial R_{x,y}(w')}{\partial w} \, \mathrm{d}w'.$$

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 (11) 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, along 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 of a change in state occurring.

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) reiterates the support restriction for $G^0(\xi)$, which rules out positive preference shocks when transitioning out of unemployment. The assumptions 1(b) and 1(d) impose restrictions on the families of distributions that we can consider. This is an extrapolation assumption. It encompasses a large set of functions; for example, the family could include both normal and logistic distributions, or it could incorporate any polynomial transformation of ξ within 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* $F_{\text{JJ}}(w|x,y,y')$*, along with knowledge of* $R_{x,y}(w)$ *and* $S_{x,y}$.
2. *Additionally, 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 drawing y', the probability of moving is equal to the probability of drawing ξ such that $S_{x,y'} + \xi > S_{x,y}$. Conditional on moving, the wage $w_{x,y'}(R')$ is determined by $R_{x,y'}(w') + \xi = S_{x,y}$. Thus the distribution of wages at a job change, conditional on x, y, y' is given by

$$F_{\text{JJ}}(w|x,y,y') = \frac{\overline{G}^1\left[S_{x,y} - R_{x,y'}(w)\right]}{\overline{G}^1\left[S_{x,y} - S_{x,y'}\right]}.$$

Lemma 4. *The probability of separation* $\delta_{x,y}$*, the rate of job meetings* λ^1*, and the probability mass of vacancies* $v_{y'}/V$ *are identified from* $p_{\text{JJ}}(y'|x,y)$*,* $p_{\text{EU}}(x,y)$*, and* $G^1(\xi)$.

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

$$p_{\text{JJ}}(y'|x,y) = \overline{\delta}_{x,y}\lambda^1 \frac{v_{y'}}{V}\overline{G}_1(S_{x,y} - S_{x,y'}). \tag{12}$$

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 the relative flows $p_{\text{JJ}}(y'_1|x, y)/p_{\text{JJ}}(y'_2|x, y)$ to recover $v_{y'}/V$.

We note here that equation (12) 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 the relaxation of the use of Assumption 1(b). Combining this with $p_{\text{EU}}(x, y)$, the layoff rate $\delta_{x,y}$ is identified by the conditional probability of separation implied by the model.

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

1. *$\lambda^0\overline{G}_0(\xi)$ is non parametrically identified on $\xi \in [-\bar{S}, 0]$.*
2. *Additionally, with Assumption 1(d), $\lambda^0\overline{G}_0(\xi)$ is identified everywhere.*

This result arises from the expression of the flow of unemployment exits alongside the wage conditional upon exiting unemployment. Importantly, we can learn $G^0(\xi)$ very flexibly. 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 with knowledge of λ^1, $G^1(\xi)$, $\lambda^0 G^0(\xi)$, $S_{x,y}$, and v_y/V, we identify $f_{x,y}$, $\tilde{c}_{x,y} = c_{x,y} + b_x$, V, and n_y based on the 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 Π^0_y:

$$f_{x,y} = w_{x,y}(S_{x,y}) + \frac{r}{1+r}\Pi^0_y, \tag{13}$$

where the only unknown part is Π^0_y. Using equation (10), we see that Π^0_y is only unknown up to the total number of vacancies V.[6] Hence, the match output is known up to one coefficient. Integrating the output of the match against the already identified distribution of matches $\ell^1_{x,y}/L^1$ expresses V as a function of the labor share and known quantities:

$$\text{labor share} = \frac{\mathbb{E}\left[w_{it}\right]}{\mathbb{E}\left[w_{x_i,y_{j(i,t)}}(S_{x_i,y_{j(i,t)}}) + \frac{1}{V}\frac{rV}{1+r}\Pi^0_y\right]}. \tag{14}$$

This is a substantive result since we did not assume transferable utility, where knowledge of $S_{x,y}$ would have directly provided us with Π^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 (10).

[6] Equation (10) equates the value of a vacancy with the expected gain from matching. If an additional period of vacancy incurs a type specific cost k_y, this vacancy cost would not be identifiable.

The final part of Lemma 6 demonstrates 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, as well as the forgone leisure and home production. We can identify this from the surplus equation (8) 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.

Taking stock. In this section, we have developed identification of the model structure in a transparent manner. The surplus of any (x, y) match is identified by aggregating the marginal utility of the wages observed in the match, weighted by 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, 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 rankings 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.

3.5 Identification Under Parametric G^0 and G^1

Assumption 2. *Assume the following about the shock distributions:*

(a) $G^0(\xi) = 2G(\rho_0\xi)$ *and* $\xi \leq 0$.

(b) $G^1(\xi) = G(\rho_1\xi)$.

(c) G *is known, is invertible, is not linear, and has a known* $G(0)$.

In Web Appendix B.5, we show that, under Assumption 2, we can first identify v_y/V, λ^0, λ^1, ρ_0, and all the match surpluses $S_{x,y}$ up to the scalar ρ_1 using only data on worker transitions. Then, in a second stage, we can use the wage data to separately identify $f_{x,y}$ and $\tilde{c}_{x,y}$. In this case, identification is very transparent: conditional on worker type, relative mobility between job types y_1 and y_2 informs us about the relative surplus workers enjoy at these jobs. Wages, conditional on the known surplus, then allow us to construct these surpluses in terms of the pecuniary and non-pecuniary aspects of the jobs. In Web Appendix B.5, we also provide a proof of the consistency of the estimator. For estimation in the following section, we adopt this more transparent version of identification, assuming that G is represented by the logistic distribution.

4 Data, Estimation and Results

Before we describe the data, we provide an overview of our estimation procedure. While we build on the identification argument, estimation requires us to make some functional-form assumptions. In particular, we adopt Assumption 2 with a logistic distribution for G. 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, given our knowledge of the parameters of $G^0(\xi)$ and $G^1(\xi)$ from the mover probabilities using Equation (12). 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, $f_{x,y}$, $\tilde{c}_{x,y}$ and $\delta_{x,y}$ remain completely unrestricted.

Data We use matched employer-employee data from Sweden. The data comprises annual tax records for all 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 the 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 include all workers under the age of 50. See the online appendix B.8 as well as Balke and Lamadon (2022) and Friedrich et al. (2022) for additional details.

Estimation of the reduced-form model We start by classifying the firms according to BLM. We group firms based on the empirical cumulative distribution function (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 discrete worker types and y to denote discrete firm types

We continue to denote ℓ_x^0 as the probability that a worker is unemployed and of type x, and $\ell_{x,y}^1$ as the probability that a worker is employed, of type x, and matched with a firm of type y. We leave these probabilities completely unrestricted. We also leave the probability of exiting unemployment unrestricted, as well as the probability for a worker of type x to join a firm of type y, which we denote $p_{\text{UE}}(x, y)$. Finally, we specify a statistical model for the hiring wage for a worker of type x who exits unemployment and joins firm y as log-normal with unrestricted mean $\mu_{\text{UE}}(x, y)$ and variance $\sigma_{\text{UE}}(x, y)$. Hence, the probability of observing a wage w for a worker joining firm y in the data is given by:

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

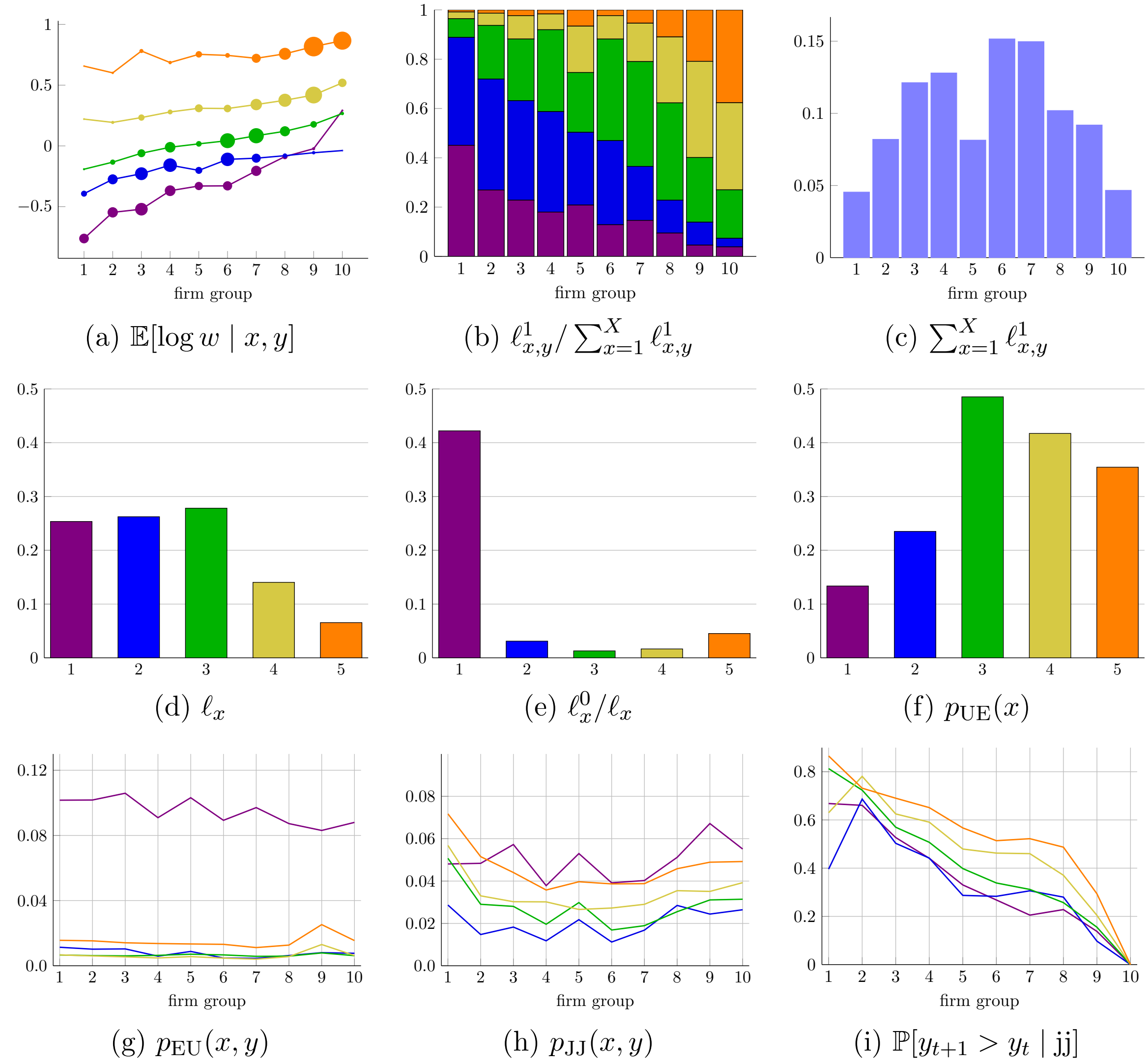
where $\mathcal{N}$ is the density of the log-normal distribution. The evolution of wages on the job is specified as an auto-regressive process with (x, y) specific intercept and a common auto-regressive coefficient:

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

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'. It is denoted as $p_{\text{JJ}}(x, y, y')$. Importantly, following the model presented in the previous section, the wage conditional on a move follows a distribution that depends only on the worker and firm types. However, conditional on these types, it 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 ℓ_x^0 and $\ell_{x,y}^1$, the mobility parameters $p_{\text{UE}}(x, y)$, $p_{\text{JJ}}(x, y, y')$, and $\delta_{x,y}$, and the parameters of the statstical 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 X worker types and Y firm types, this amounts to a number of parameters of the order of $3X \cdot Y^2$. In estimation, we focus on $X = 5$ and $Y = 10$.

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 the mean log wage for each worker type x 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 ranking of firms, where all workers are generally paid more as we move to higher firm types. The figure also reveals a relatively stable ordering of workers, except for the lowest worker type concerning the top two firm



(a) $\mathbb{E}[\log w \mid x, y]$ (b) $\ell^1_{x,y} / \sum_{x=1}^{X} \ell^1_{x,y}$ (c) $\sum_{x=1}^{X} \ell^1_{x,y}$

(d) ℓ_x (e) ℓ^0_x / ℓ_x (f) $p_{\text{UE}}(x)$

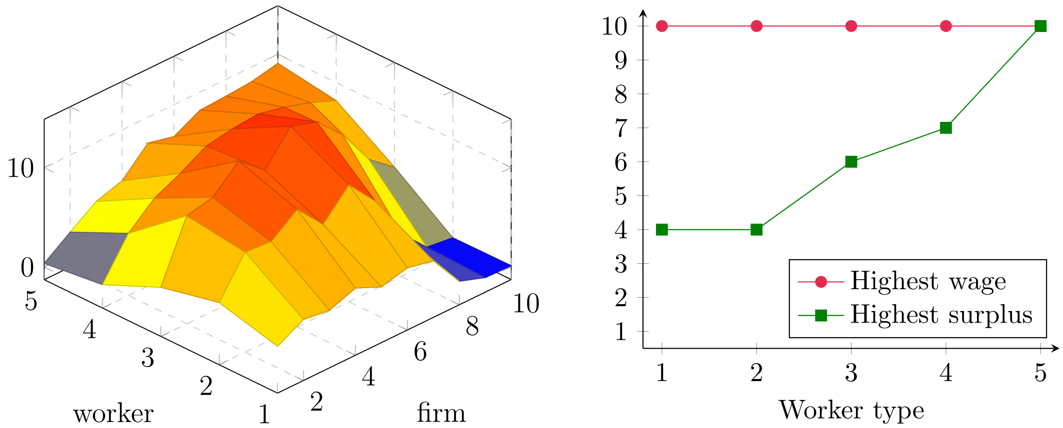
(g) $p_{\text{EU}}(x, y)$ (h) $p_{\text{JJ}}(x, y)$ (i) $\mathbb{P}[y_{t+1} > y_t \mid \text{jj}]$

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

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

types, where the data are sparse.

In panel (b), we plot the composition of worker types employed in each type of firm. The cross-sectional distribution exhibits a clear and strong pattern of separating high-paid workers into high-paying firms. In panel (c), we plot the share of total employment for each type of firm in the cross section. 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 match specific probabilities of transitioning from employment to unemployment, changing jobs, and moving to a higher firm type following a job change.



(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.

The estimates in Step 1 provide a clear picture of what distinguishes the different worker types from one another. 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, but they also have the highest unemployment rate (over 40%), which results from the lowest probability of moving from unemployment to employment, combined with a probability of transitioning into unemployment from any type of firm that is about 10 times higher than the average for the other types of workers. These workers also have high job-to-job transition probabilities; however, they have among the lowest probabilities of moving to a higher paying firm when changing jobs. On the other hand, workers in the highest category, comprising about 7% of the population, have the highest wages when employed, unemployment rates below 5%, a high probability of transitioning from unemployment to employment, and a low probability of separating from employment to unemployment. Like the lowest type, 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.[7]

[7] In Online Appendix Figure 7 , we plot the stationary distribution of worker types across firm 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.

Estimation of surpluses, vacancy rates, and meeting rates. Under Assumption 2 and a logistic G we have:

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

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.9 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:

$$\begin{aligned} m_1 =& \mathbb{E}\left[\log w_t - \log w_{t-4} \middle| s_t{=}s_{t-4}{=}1,\ m_{t-4}{=}\text{EE}\right], \\ m_2 =& \mathbb{E}\left[\mathbb{E}\left[\log w_t|x, y, s_t{=}1, m_{t-4} = \text{JJ}\right] - \mathbb{E}\left[\log w_t|x, y, s_t{=}1, m_{t-4} = \text{UE}\right]\right], \end{aligned}$$

where t represents a quarter and $s_t{=}1[m_{t-1}{=}m_{t-2}{=}m_{t-3}{=}\text{EE}]$ 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 estimates of $\hat{\rho}_1 = 0.42$ and $\hat{\rho}_0 = 0.07$. This suggests that there is a greater variance in costs when leaving unemployment than when transitioning 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 support this ranking. 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 regarding which firms offer high-pay, there is complete disagreement among worker types about which firms provide high surplus. These results align well with the reduced form empirical findings of Lentz et al. (2023), demonstrating that inferences about sorting based solely on wage information differ significantly from those that incorporate 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 (6) 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 Π^0_y using equation (10) 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 (14). We then use equation (13) to recover the production

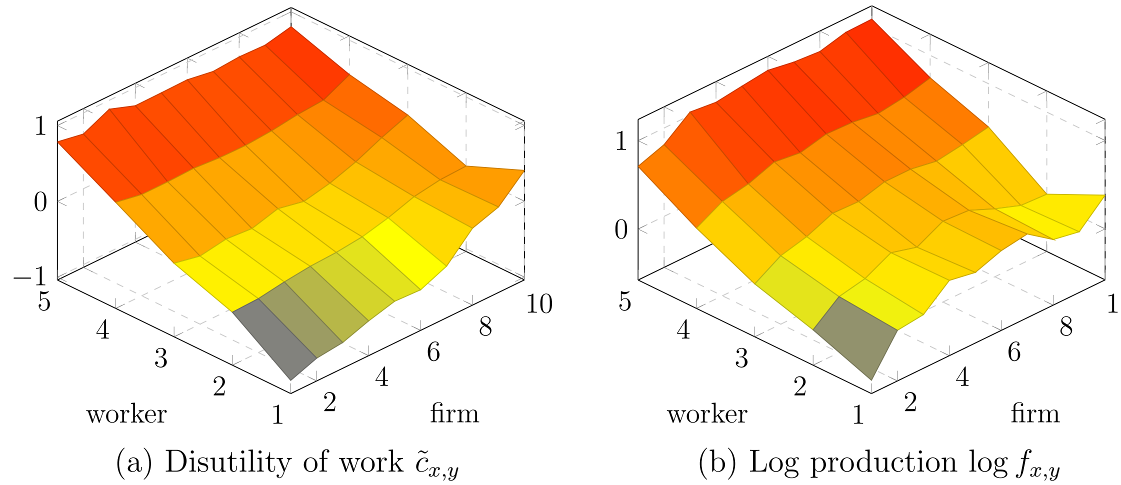


Figure 3: Disutility net of amenities and production function (Steps 4 and 5)

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 noteworthy patterns here. First, the average disutility of work increases monotonically with worker type; higher worker types face a greater opportunity cost of work. For the lowest worker types, the disutility of labor increases with firm type, indicating that while these firms offer good pay, they are unattractive to low-type workers in non-wage dimensions. In contrast, for the highest worker type, the disutility of labor is high but effectively independent of firm type. Finally, the estimates reveal striking monotonicity in both dimensions x and y (except for the lowest worker type paired with the highest firm, where there is limited data). Turning to the estimates of the production function, we observe striking monotonicity in both x and y, again with the exception of the pairing 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, interpret well-documented empirical regularities through the structure of the model, and derive and present model consistent separation and hiring elasticities.

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

$$\underbrace{Var(\log w)}_{\text{total}} = \underbrace{\mathbb{E}[Var(\log w|x)]}_{\text{within worker: 28\%}} + \underbrace{Var(\mathbb{E}[\log w|x])}_{\text{between worker: 72\%}}. \tag{15}$$

Turning first to the within worker (within x-type) term, we decompose this into interpretable sources in equation (16). 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, as this allows a firm to offer the same worker a different value over time as their outside option evolves. This variation within a worker-firm match is induced by the sequential auction mechanism (Postel-Vinay and Robin, 2002, PVR), which prices in both the effect of the outside firm and the mobility cost.[8] 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.[9]

For a fixed worker type x, we first decompose wages into variation within and between a fixed surplus value R. Second, we decompose the variation between R into the variation between firms and within firms. To simplify notation, we first define $\mu_R = \mathbb{E}[\log w|R, x]$ and write:[10]

$$\underbrace{Var(\log w|x)}_{\text{within worker}} = \underbrace{\mathbb{E}[Var(\log w|R, x)|x]}_{\text{(a) Rosen}} + \underbrace{Var(\mathbb{E}(\mu_R|y, x)|x)}_{\text{(b) Mortensen}} + \underbrace{\mathbb{E}[Var\left[\mu_R|y, x\right]|x]}_{\text{(c) PVR}} \tag{16}$$

We present this decomposition in Figure 4 for each worker type separately, as well as for the average across 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 50%. This is an important finding, as it implies that for the typical worker, half of the variation in observed wages can be attributed to compensating differentials that do not

[8] While there are a finite number of combinations of x, y, and y', there is a continuum of mobility shocks ξ. Since ξ is priced in, there is a continuum of worker values in each (x, y) match $R \in [0, S_{x,y}]$. This implies a continuum of wages, even conditional on x, y, and y'.

[9] 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.

[10] See Online Appendix B.6 for complete derivation.

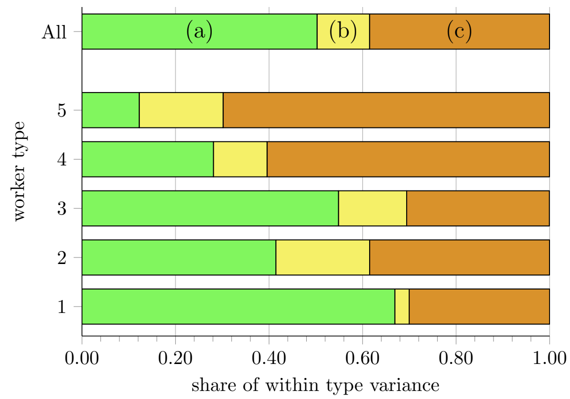


Figure 4: Within worker wage variance contribution

Notes: We plot the three terms of the equation (16) 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 28% for the total variance.

reflect utility differences. Second, there are substantial differences between workers. At the bottom of the distribution, there is considerably less variation in value compared to the variation in wages. About 67% of the wage variation for the lowest group is the compensating differential and occurs at a constant utility value. At the top, only 12% of the wage variation is due to compensating differentials. High-paid workers face important utility differences between employers and notable wage dynamics within jobs, resulting from the offer-counteroffer mechanism.

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

$$\underbrace{Var(\mathbb{E}[\log w|x])}_{\text{between worker: 72\%}} = \underbrace{\mathbb{E}[Var(\mu_x|y)]}_{\text{within firm, between worker: 52\%}} + \underbrace{Var(\mathbb{E}[\mu_x|y])}_{\text{between firm, sorting: 20\%}} . \tag{17}$$

The first term of equation (17), $\mathbb{E}[Var(\mu_x|y)]$, captures the average within-firm variance of average worker values. This reflects the fact that even when all firms are identical and pay identical wages, they will hire a distribution of workers, each of whom has a different average market wage. The second term, $Var(\mathbb{E}[\mu_x|y])$, represents the between firm variance in wages, net of firm-specific pay policies. This term is the contribution attributable to the fact that different firms hire workers with different average market values. This highlights that firms do differ from eachother and have differing worker compositions, directly reflecting the sorting of workers across firms.

We summarize the decomposition in Table 1. Most of the wage variance arises from differences among workers within firms, accounting for 52% of the overall wage variation. The sorting of workers to firms accounts for another 20% of the overall

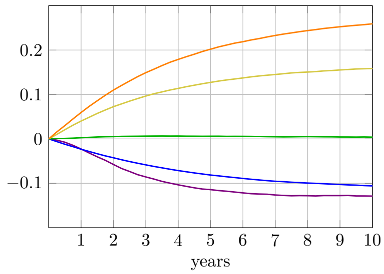
Table 1: Structural variance decomposition of log wages

Total variance of log wages		0.1423
Between worker		
├ Within firm	Worker effects	51.85%
└ Between firm	Sorting effect	20.47%
Within worker		
├ Within R	Rosen compensating differential	13.83%
└ Between R		
├ Between firm	Mortensen search friction	3.03%
└ Within firm	PVR search friction	10.82%

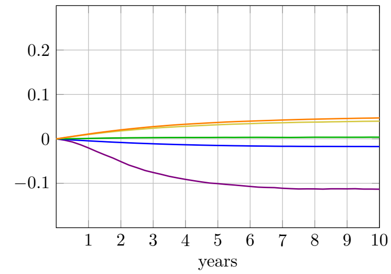
Note: While the model wage equation is not linear in worker and firm types, the variance decomposition we report is exact. An alternative to this decomposition would be to report the measures of worker-, firm heterogeneity and sorting that arise from a two way best linear predictor decomposition in the spirit of AKM. Applying AKM to model-simulated data, we find that the shares of wage dispersion are as follows: worker fixed effect 50%, firm fixed effect 10%, covariance 20%, nonlinearities 1.4%, and a residual of 19%. While not directly comparable to our decomposition, the shares attributable to worker heterogeneity and sorting (covariance) are strikingly similar. See Online Appendix Table 4.

variance. The remaining variance occurs within the worker type and can be split as follows: 14% is due to compensating differentials at a fixed utility level, 3% results from different firms offering different values to the same workers, and the remaining 11% stems from sequential auctions. The last two represent the combined effect of market frictions, indicating that search frictions account for 14% of wage variation for a typical worker; however, this share varies substantially between worker types.

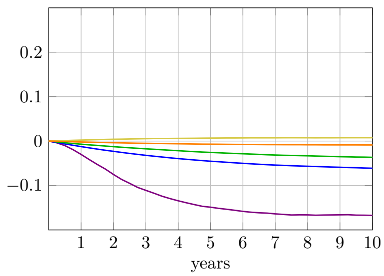
Contributions to the dynamics of sorting The sorting pattern observed in Figure 1.b is generated by a combination of the match specific production function and the match specific disutility of labor. To illustrate the effect of each of these forces, we conduct the following model experiment. For each worker type x, we start with workers uniformly allocated across firm types. Next, we simulate forward and follow these workers as they converge to the steady state. Define $\log w_y$ as the mean wage in a type y firm, and $\log w_{x,y}$ as the mean wage in an x, y match, both calculated in the steady state. In Figure 5 we plot $\mathbb{E}_t[\log w_y|x]$, $\mathbb{E}_t[\log w_{x,y}|x]$, $\mathbb{E}_t[c_{x,y}|x]$, and $\mathbb{E}_t[\log f_{x,y}|x]$ for the first 10 years of the simulation. At each period of the simulation, expectations



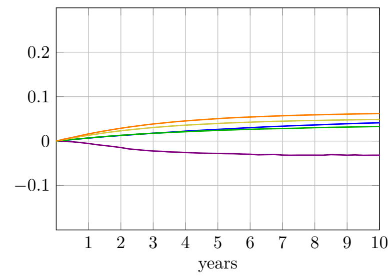
(a) Mean firm wage, $\mathbb{E}_t[\log w_y|x]$



(b) Mean match wage, $\mathbb{E}_t[w_{x,y}|x]$



(c) Mean disutility of labor, $\mathbb{E}_t[c_{xy}|x]$

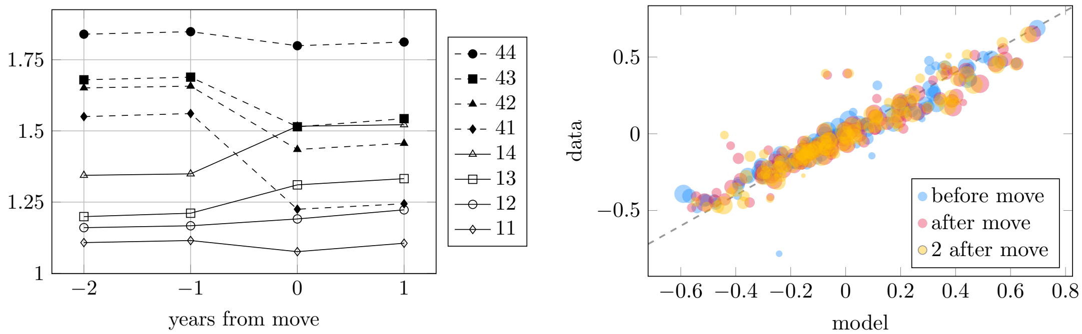


(d) Mean match production, $\mathbb{E}_t[\log f_{xy}|x]$

Figure 5: Reallocation dynamics for workers starting at randomly drawn firms.
Note: The lines correspond to worker types —1, —2, —3 , —4, and —5.

are formed based on the current distribution of workers across firms, beginning with a uniform distribution in period 0 and approaching the steady state distribution $\ell^1_{x,y}$ after 10 years. All series are normalized against period 0.

The dynamic sorting is clearly illustrated in panel (a), where we observe that high type workers systematically reallocate to firms with higher mean wages, while low type workers systematically reallocate to firms paying the lowest mean wages. In panel (b), we observe that the same pattern holds for the match-specific mean wages; however, these changes are muted because the firm component of wages is dominated by the worker specific and sorting components. In panel (c), we see that while all worker types reallocate toward firms with a lower type-specific disutility of labor, this is particularly important for the lowest worker type. Finally, in panel (d), we observe that, with the exception of the lowest type, all other worker types reallocate toward higher productivity matches. The desire to reduce match disutility is particularly strong for the lowest worker type, even leading them to reallocate to firms with lower match productivity.



(a) Movers event study in model

(b) Fit around the move

Figure 6: Event Study plot.

Notes: Panel (a) reproduces the event study plot from Card et al. (2013) using model simulated data. It plots yearly earnings before and after a move for different firms, grouped by quartiles of earnings. Lines are labeled by origin and destination quartile of the firm. For example, "−⧫ − 41" plots the mean log wage of workers who move from a firm in quartile 4 to a firm in quartile 1 at period 0. Panel (b) shows the fit of the model for annual earnings before and after a move for all pairs of model firm types.

The quantitative role of the preference shock An important part of our framework arises 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 remain 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, current firm type, and poaching firm type, rather than by a combination of types and the preference shock. In this counterfactual, focusing on employed workers who have received at least one offer, the wage variance decreases by 1% relative to the baseline log wage variance.[11] In summary, the model attributes 19% of job-to-job moves and 1% of log wage variation to the preference shock.

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 6, we reproduce this study based on our estimated model. We simulate data

[11] While the total variance decreases by only 1% there is a substantial reallocation within and between workers.

from the model and then aggregate it to annual earnings, exactly as described in Card et al. (2013).

We notice several important features. First, as typically documented, we also find that the movers from the top quartile to the bottom quartile experience an earnings loss that is not alleviated in the period following the move. Sequential contracting provides a rationale for why workers might accept a wage cut when moving, expecting that they will recover later 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 wages within worker types. Similarly, the movers from the bottom quartile to the top experience earnings gains that remain remarkably consistent in the following period. In panel (b), we plot the model fit to the data for annual earnings before and after a move between all pairs of firm types.

In the canonical AKM specification, log wages are assumed to be additively separable into a worker effect and a firm effect. BLM provides a specification that allows for nonlinearities and finds the log-additive specification to be a good approximation, as confirmed here in Figure 1.a. Additionally we find that there is significant sorting of high type workers into high type jobs (Figure 1.b). W also find that when making job-to-job transitions, high type workers tend to move to better paying firms, while low type workers tend to move to lower paying firms (Figure 1.i). Without compensating differentials, reconciling these three patterns would be difficult. However, the surplus represents a combination of productivity and amenity. This means that the model can rationalize both the wages and the surplus by incorporating both $f_{x,y}$ and $\tilde{c}_{x,y}$. The combination of log-additive wages and strong sorting results in estimated dis-amenities of work that increase across firm types for low type workers, generating positive sorting.

Recently, Di Addario, Kline, Saggio, and Sølvsten (2023) augmented the AKM specification to allow for an effect of the previous job on the current wage. They find that the firm where a worker was employed prior to a move does not significantly affect the wage after the move. At face value, this suggests that sequential contracting may not play a significant role in wage determination. An important aspect of a model like ours, in which sorting is modeled structurally, is that the shape of the surplus leads to heterogeneous effects of previous firms on different workers. This will tend to muddle the average effect of the previous firm. Indeed, when we estimate the linear additive formulation on data simulated from our model, aggregated as in Di Addario et al. (2023), we also find little effect of the previous employer on the wage immediately following a move. This is despite the fact that, in the model, contracts are indeed determined with reference to the previous employer. In Online Appendix Table 5, we present this decomposition, showing that the previous firm type explains less than half a percent of the variance in log wages following a move.

Separation and hiring elasticities In a dynamic model with labor flows, it is natural to compute separation and hiring elasticities. In the canonical wage posting model of Burdett and Mortensen (1998), firms choose a fixed wage policy that directly affects hiring and separation rates. In the current model, firms make state contingent offers and counteroffers and do not have a fixed wage. We can construct wage elasticities by making small adjustments to the firm's recruiting and retention policies, assuming that an individual firm bids $B = (1 - \Delta)S_{x,y}$ when competing for workers. All other firms continue to bid $S_{x,y'}$, allowing us to directly express how this deviation affects both hiring and separation probabilities for each pair x, y:

$$h_{x,y}(B) = \frac{\lambda^0 L_0}{V} \frac{\ell^0_x}{L_0} G^0(B) + \sum_{y'} \frac{\lambda^1 L_1}{V} \frac{\ell^1_{x,y}}{L_1} G^1(B - S_{xy'}),$$

$$q_{x,y}(B) = \delta_{x,y} + (1 - \delta_{x,y})\lambda^1 \sum_{y'} \frac{v_{y'}}{V} G^1(S_{xy'} - B).$$

From here, we want to compute an elasticity with respect to wages. We use the average wage paid by firm y to worker x in the cross-section under policy $B = (1 - \Delta)S_{x,y}$. The change in the surplus offer generates changes in average wages, hiring probabilities, and separation probabilities compared to bidding $S_{x,y}$. From these induced changes, we construct elasticities with respect to the wage. The resulting elasticities for each firm type, as well as the average, are reported in Table 2 using $\Delta = 0.05$. Hiring elasticities range from 2.8 to 5.1, with an average of 3.6, while separation elasticities range from -5.8 to -1.7, with an average of -3.8.

In the second to last row, we compute the size elasticity for a job, which is the filling rate. It is interesting to note that, in this environment, the relationship between the separation, hiring, and size elasticities is not the same as in Burdett and Mortensen (1998). This is due to the assumption that firms have a fixed capacity, where the job is either vacant or filled. We obtain a different expression from the typical $\varepsilon_{n,w} = \varepsilon_{h,w} - \varepsilon_{q,w}$:

$$\varepsilon_{n,w} = \frac{q}{h + q}\big(\varepsilon_{h,w} - \varepsilon_{q,w}\big),$$

Table 2: Separation, hire and employment elasticities

Firms	1	2	3	4	5	6	7	8	9	10	all
Hiring probability	0.05	0.08	0.08	0.11	0.09	0.11	0.10	0.05	0.03	0.02	0.07
Separation probability	0.10	0.06	0.06	0.04	0.06	0.03	0.04	0.05	0.06	0.05	0.05
Employment probability	0.31	0.58	0.59	0.75	0.61	0.78	0.71	0.50	0.30	0.28	0.54
Hire elasticity	3.9	3.6	3.6	2.9	2.8	2.8	3.1	5.1	4.2	3.8	3.6
Separation elasticity	-1.7	-3.2	-3.3	-4.5	-3.1	-5.8	-4.6	-4.4	-2.8	-4.1	-3.8
Employment elasticity	3.6	3.0	3.0	2.0	2.3	2.0	2.5	4.8	5.1	5.5	3.5
Output elasticity	3.7	3.1	3.0	2.0	2.5	2.1	2.6	5.2	5.5	5.9	3.7

Notes: Elasticities are calculated using the wage and transition rate changes generated by simulating firms who bid $B = 0.95 \times S_{x,y}$ rather than the equilibrium $B = S_{x,y}$ when competing for workers. For a model with fixed capacity we have $n = h/(h+q)$, where n is share of filled jobs, h is the hiring rate and q is the separation rate. The employment elasticity is $\varepsilon_{n,w} = \frac{q}{h+q}(\varepsilon_{h,w} - \varepsilon_{q,w})$. Employment and output elasticity are very close since the experiment involves little change in composition.

where q and h represent the separation rate and the hiring rate. See Appendix B.7 for details.

We want to put this in the context of the literature on separation elasticities. We think the most comparable estimates are from Bassier et al. (2022). Among the several specifications, we can focus on two estimates that we can relate to our model. First, they report an elasticity of separations with respect to the average firm wage of -0.282 in the first column of Table 1. The same regression in the cross-section generated by our model produces a very similar separation elasticity of -0.20. When using the firm premium, their estimate increases to around -1.3. In the working paper version, they also report an estimate that classifies firms first, where the estimate approaches -2. The same cross-firm regression of log separation on log-wage premium in our case gives an elasticity of -3.4, which is higher. Finally, controlling for compensating differentials across firms, our structural estimate of the average elasticity of separation comes to -3.8. This magnitude is at the upper end of the reported results.

The size elasticities, on the other hand, are on the lower end but well within the usual range of reported values at 3.5. It is key to apply the formula derived for the capacity constrained model, as simply doubling the separation elasticity would produce a very large size elasticity.

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 nonparametrically identified from a short panel of matched employer-employee data. We find that for the lowest type workers, compensating differentials account for 67% of the within-worker wage variation, but only 12% 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 can be used to conduct counterfactual policy experiments and assess the efficiency and distributional effects of, for example, changing the progressivity of taxes, earned income tax credits, or minimum wages. Policy changes that affect the relative value of wages and amenities will directly influence how workers value different jobs and 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.

Bassier, I., A. Dube, and S. Naidu (2022): "Monopsony in movers: The elasticity of labor supply to firm wage policies," *Journal of Human Resources*, 57, S50–s86.

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.

Borovickova, K. and R. Shimer (2024): "Assortative Matching and Wages: The Role of Selection," NBER wp33184.

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, C. Meghir, and L. Pistaferri (2022): "Earnings dynamics and firm-level shocks," Tech. rep., NBER wp25786.

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

Jolivet, G., F. Postel-Vinay, and J.-M. Robin (2006): "The Empirical Content of the Job Search Model: Labor Mobility and Wage Dispersion in Europe and the U.S." *European Economic Review*, 50, 877–907.

Koopmans, T. C. and M. Beckmann (1957): "Assignment Problems and the Location of Economic Activities," *Econometrica*, 25, 53–76.

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.

Lindenlaub, I. and F. Postel-Vinay (2023): "Multi-dimensional Sorting Under Random Search," *Journal of Political Economy*, 131, 3497—3539.

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

Lise, J. and F. Postel-Vinay (2020): "Multidimensional skills, sorting, and human capital accumulation," *American Economic Review*, 110, 2328–76.

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," in *Handbook of Labor Economics*, 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.

Shapley, L. S. and M. Shubik (1971): "The assignment game I: The core," *International Journal of Game Theory*, 1, 111–130.

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. All dependence on (x,y) is suppressed for notational clarity. 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:

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

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

$$
\begin{aligned}
-\Delta_R \mu = \Delta_R g(v) &\leq \gamma \Delta_R v \\
-\gamma \Delta_R v &\leq \Delta_R \mu.
\end{aligned}
$$

We then turn to the PK constraint, we have

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

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:

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

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 $\Pi^1 \le \Pi^1_*$ 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 $\Pi(R) \ge \Pi^0$. 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 $\alpha \ge 0$ 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 $\tilde{\delta} \ge \delta$, this gives us that:

$$T[\Pi^1 + \alpha](R) \le T[\Pi^1](R) + \frac{1-\delta}{1+r}\alpha.$$

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:

$$\mathcal{H}(B', \xi, R_1, \mu, \bar{\mu}, \underline{\mu}) = \mathbf{1}\{S \geq B'^{+} + \xi\}\Pi^1(R_1) + \mathbf{1}\{S < B'^{+} + \xi\}\Pi^0 - \mu\mathbf{1}\{S \geq B'^{+} + \xi\ \}R_1$$
$$-\mu\mathbf{1}\{S < B'^{+} + \xi\}\max\{S, \xi\} + \mathbf{1}\{S \geq B'^{+} + \xi\}\left(\underline{\mu}\cdot(R_1 - B'^{+} - \xi) + \bar{\mu}\cdot(R_1 - S)\right)$$

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

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

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:

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

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

$$\Pi^{1\prime}(R_1(B',\xi)) = \mu = \Pi^{1\prime}(R_0),$$

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 $\Pi^{1\prime}(R) = \mu = \Pi'(R_0)$ 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 $\Pi^{1\prime}(R) = -\frac{1}{u'(w(R))}$. □

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:*

$$F_{\text{JJ}}(w|x,y,y') = \frac{\overline{G}^1(S_{x,y} - R_{x,y'}(w))}{\overline{G}^1(S_{x,y} - S_{x,y'})} \tag{20}$$

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 $w \in]w_{x,y'}(0), w_{x,y'}(S_{x,y'})[$, the CDF following a job-to-job transition is given by

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

Therefore, we have for all $w \in [w_{x,y'}(0), w_{x,y'}(S_{x,y'})[$ that $F_{\text{JJ}}(w|x,y,y') = \overline{G}^1(S_{x,y} - R_{x,y'}(w))/\overline{G}^1(S_{x,y} - S_{x,y'})$. □

Lemma A.9. *We show that $F_{\text{UE}}(w|x,y') = \overline{G}^0(-R_{x,y'}(w))/\overline{G}_0(-S_{x,y'})$.*

$$\begin{aligned} F_{\text{UE}}(w|x,y') &= \mathbb{P}[w_{t+1} \le w | x, y_{t+1} = y', m_t = \text{UE}] \\ &= \mathbb{P}[w_{x,y'}(-\xi) \le w | \xi > -S_{x,y'}] \\ &= \overline{G}^0(-R_{x,y'}(w))/\overline{G}_0(-S_{x,y'}). \end{aligned}$$

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 $\sum_{y=1}^{Y} v_y = V$, 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:

$$\mathbb{P}[m_t = \text{EU}|x, y_t = y] = \delta_{x,y} + \overline{\delta}_{x,y}\lambda^1 \left(\sum_{y'=1}^{Y}(1 - \phi_{x,y'})\frac{v_{y'}}{V}\right)\overline{G}^1(S_{x,y}),$$

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:

$$\mathbb{P}[y_{t+1} = y, m_t = \text{UE}|x] = \lambda^0 \frac{v_y}{V}\overline{G}_0(-S_{x,y})$$

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:

$$F_{\text{UE}}(w|x,y) = \overline{G}^0(-R_{x,y}(w))/\overline{G}_0(-S_{x,y}).$$

We introduce $\xi = -R_{x,y}(w)$ and use $R_{x,y}(w_{x,y}(-\xi)) = \xi$ to write

$$\forall \xi \in [-S_{x,y}, 0] \qquad \lambda^0\overline{G}_0(\xi) = \frac{V}{v_y}\mathbb{P}[y_{t+1}{=}y, m_t{=}\text{UE}|x]F_{\text{UE}}(w_{x,y}(-\xi)|x,y),$$

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 Π^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.

$$r V \Pi^0_y = \lambda^0 \sum_{x=1}^{X} \int_{-S_{x,y}}^{0} \frac{\overline{G}^0(\xi)}{u'(w_{x,y}(-\xi))} \ell^0_x \,\mathrm{d}\xi + \lambda^1 \sum_{x=1}^{X}\sum_{y'=1}^{Y} \int_{-S_{x,y}}^{0} \frac{\overline{G}^1(\xi + S_{x,y'})}{u'(w_{x,y}(-\xi))}\overline{\delta}_{x,y}\ell^1_{x,y'}\,\mathrm{d}\xi. \tag{22}$$

Hence, we know Π^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, Π^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 $n_y = \sum_{x=1}^{X} \ell^1_{x,y} + v_y$, 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 $S := S_{x,y}$ and $S' := S_{x,y'}$. 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 is convenient 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:

$$
\begin{aligned}
rW^1_{x,y}(w) = {} & u(w) - c_{x,y} + \frac{\delta_{x,y}}{1+r}\left(W^0_x - W^1_{x,y}(w)\right) \\
& + \frac{\lambda^1\overline{\delta}_{x,y}}{1+r}\sum_{y'=1}^{Y}\Bigg[\int_{W^1_{x,y}(w)-W^0_x-S_{x,y'}}^{S_{x,y}-S_{x,y'}} \left[W^0_x + S_{x,y'} + \xi - W^1_{x,y}(w)\right] g(\xi)\,\mathrm{d}\xi \\
& \qquad + \int_{S_{x,y}-S_{x,y'}}^{S_{x,y}} \left[S_{x,y} + W^0_x - W^1_{x,y}(w)\right] g(\xi)\,\mathrm{d}\xi \\
& \qquad\qquad + \int_{S_{x,y}}^{\infty} \left[W^0_x + \xi - W^1_{x,y}(w)\right] g(\xi)\,\mathrm{d}\xi\Bigg]\frac{v_{y'}}{V}.
\end{aligned}
$$

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 (6).

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

$$
\begin{aligned}
u(w_{x,y}(R)) = {} & \frac{r+\delta_{x,y}}{1+r}R + \frac{r}{1+r}W^0_x + c_{x,y} \\
& - \frac{\overline{\delta}_{x,y}\lambda^1}{1+r}\int_{S_{x,y}}^{\infty}\overline{G}^1(\xi)\,\mathrm{d}\xi - \frac{\overline{\delta}_{x,y}\lambda^1}{1+r}\sum_{y'=1}^{Y}\left[\int_{R}^{S_{x,y}}\overline{G}^1(R'-S^+_{x,y'})\,\mathrm{d}R'\right]\frac{v_{y'}}{V}. \quad (24)
\end{aligned}
$$

The minimum wage is $\underline{w}_{x,y} = w_{x,y}(0)$ that produces zero surplus. The maximum wage is $\overline{w}_{x,y} = w_{x,y}(S_{x,y})$ that produces the maximum 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:

$$
(1+r)u'(w_{x,y}(R))\frac{\partial w_{x,y}(R)}{\partial R} = r + \delta_{x,y} + \overline{\delta}_{x,y}\lambda^1\sum_{y'=1}^{Y}\overline{G}^1\left[R - S^+_{x,y'}\right]\frac{v_{y'}}{V} > 0. \quad (25)
$$

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

$$
(1+r)u'(w_{x,y}(R))\frac{\partial w_{x,y}(R)}{\partial S_{x,y}} = -\overline{\delta}_{x,y}\lambda^1\sum_{y'=1}^{Y}\left[G^1(S_{x,y}) - G^1(S_{x,y} - S^+_{x,y'})\right]\frac{v_{y'}}{V} < 0. \quad (26)
$$

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 (7) for the maximum wage along with $R = S_{x,y}$ into the wage equation (6) we have

$$
\begin{aligned}
u(f_{x,y} - r(1+r)^{-1}\Pi^0_y) = c_{x,y} &+ \frac{r+\delta_{x,y}}{1+r} S_{x,y} + \frac{r}{1+r} W^0_x \\
&- \frac{\lambda^1 \overline{\delta}_{x,y}}{1+r} \sum_{y'=1}^{Y} \phi_{x,y'} \Bigg[\int_{S_{x,y}-S_{x,y'}}^{S_{x,y}-S_{x,y'}} \left[S_{x,y'} + \xi - S_{x,y} \right] g(\xi)\, \mathrm{d}\xi \\
&\qquad + \int_{S_{x,y}-S_{x,y'}}^{\infty} \left[\max\left\{ S_{x,y} - \xi, 0 \right\} + \xi - S_{x,y} \right] g(\xi)\, \mathrm{d}\xi \Bigg] \frac{v_{y'}}{V}.
\end{aligned}
$$

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 (8).

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

$$
\begin{aligned}
\left[r + \delta_{x,y} \right] \Pi^1_{x,y}(R) &= (1+r) \left[f_{x,y} - w_{x,y}(R) \right] + \delta_{x,y} \Pi^0_y \\
&+ \overline{\delta}_{x,y} \lambda^1 \sum_{y'=1}^{Y} \Big[\int_{R-S^+_{x,y'}}^{S_{x,y}-S^+_{x,y'}} \left(\Pi^1_{x,y}(S^+_{x,y'} + \xi) - \Pi^1_{x,y}(R) \right) \mathrm{d}G^1(\xi) \\
&\qquad + \int_{S_{x,y}-S^+_{x,y'}}^{S_{x,y}} \left(\Pi^0_y - \Pi^1_{x,y}(R) \right) \Big] \mathrm{d}G^1(\xi) \frac{v_{y'}}{V}. \qquad (27)
\end{aligned}
$$

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

$$\Pi^1_{x,y}(S_{x,y}) = \Pi^0_y. \tag{28}$$

We can therefore also deduce firm profits from worker wages as

$$\Pi^1_{x,y}(R) = \Pi^0_y + \int_R^{S_{x,y}} \frac{\mathrm{d}R'}{u'(w_{x,y}(R'))}. \tag{29}$$

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:

$$r\Pi^0_y = \lambda^0 \frac{L^0}{V} \sum_{x=1}^{X} \int_{-S_{x,y}}^{\infty} J_{x,y}(\max\{-\xi, 0\})\,\mathrm{d}G^0(\xi) \frac{\ell^0_x}{L^0}$$
$$+ \lambda^1 \frac{L^1}{V} \sum_{x=1}^{X} \sum_{y'=1}^{Y} \int_{S_{x,y'}-S_{x,y}}^{\infty} J_{x,y}(\max\{S_{x,y'} - \xi, 0\})\,\mathrm{d}G^1(\xi)(1 - \delta_{x,y'}) \frac{\ell^1_{x,y'}}{L^1}.$$

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

$$r\Pi^0_y = \lambda^0 \frac{L^0}{V} \sum_{x=1}^{X} \phi_{x,y} \left[\overline{G}^0(0) J_{x,y}(0) + \int_{-S_{x,y}}^{0} J_{x,y}(-\xi)\,\mathrm{d}G^0(\xi) \right] \frac{\ell^0_x}{L^0}$$
$$+ \lambda^1 \frac{L^1}{V} \sum_{x=1}^{X} \sum_{y'=1}^{Y} \phi_{x,y} \Big[\overline{G}^1(S_{x,y'}) J_{x,y}(0)$$
$$+ \int_{S_{x,y'}-S_{x,y}}^{S_{x,y'}} J_{x,y}(S_{x,y'} - \xi)\,\mathrm{d}G^1(\xi) \Big] (1 - \delta_{x,y'}) \frac{\ell^1_{x,y'}}{L^1}. \tag{30}$$

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

$$\int_{-S_{x,y}}^{0} J_{x,y}(-\xi)\,\mathrm{d}G^0(\xi) = -\left[J_{x,y}(-\xi)\overline{G}^0(\xi) \right]_{-S_{x,y}}^{0} - \int_{-S_{x,y}}^{0} \frac{\partial J_{x,y}}{\partial R}(-\xi)\overline{G}^0(\xi)\,\mathrm{d}\xi$$
$$= -J_{x,y}(0)\overline{G}^0(0) + \int_{-S_{x,y}}^{0} \frac{\overline{G}^0(\xi)}{u'(w_{x,y}(-\xi))}\,\mathrm{d}\xi,$$

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 $\frac{\partial J_{x,y}}{\partial R}(R) =$

$-\frac{1}{u'(w_{x,y}(R))}$. Similarly, in the second case we have

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

Substituting these two expressions into equation (30) we obtain (10), 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):

$$\mathbb{P}\left[m_{t+1} = \mathrm{UE}|x, y_{t+1} = 0, \Omega_t\right] = \lambda^0 \sum_{y=1}^{Y} G^0(-S_{x,y})\phi_{x,y}\frac{v_y}{V},$$

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

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

$$\mathbb{P}\left[m_{t+1} = \text{EU}|x, y_{t+1} = y, w_{t+1}, \Omega_t\right] = \delta_{x,y} + \overline{\delta}_{x,y}\lambda^1 \sum_{y'=1}^{Y} \left[1 - \phi_{x,y'}\right] \frac{v_{y'}}{V} \times \overline{G}^1(S),$$

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

$$\mathbb{P}\left[m_{t+1} = \text{JJ}|x, y_{t+1} = y, w_{t+1}, \Omega_t\right] = \overline{\delta}_{x,y}\lambda^1 \sum_{y'=1}^{Y} \overline{G}^1(S - S')\phi_{x,y'} \frac{v_{y'}}{V},$$

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

$$\mathbb{P}\left[m_{t+1} = \text{EE}|x, y_{t+1} = y, w_{t+1}, \Omega_t\right] = \overline{\delta}_{x,y}\overline{\lambda^1} + \overline{\delta}_{x,y}\lambda^1 \sum_{y'=1}^{Y} G^1(S - S'^{+}) \frac{v_{y'}}{V}.$$

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 $m_t \in \{\text{UU}, \text{EU}\}$ 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 $m_t = \text{UE}$, 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:

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

Similarly when $m_t = \text{JJ}$, 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 $S_{x,y} - S_{x,y'}$:

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

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

We then consider our final case of $m_t = \text{EE}$. 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 $m_t = \text{EE}$ 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 $S^+_{x,y'} + \xi \leq R_{x,y}(w')$:

$$\mathbb{P}\left[m_t = \text{EE}, w_{t+1} \leq w' | x, w_t = w, y_t = y, \Omega_{t-1}\right]$$
$$= \overline{\delta}_{x,y}(1-\lambda^1) + \overline{\delta}_{x,y}\lambda^1 \sum_{y'=1}^{Y} G^1\left[R_{x,y}(w') - S^+_{x,y'}\right]\frac{v_{y'}}{V}.$$

Hence

$$\mathbb{P}\left[w_{t+1} \leq w' | x, w_t, y_t, y_{t+1}, m_t = \text{EE}, \Omega_{t-1}\right]$$
$$= \mathbf{1}\left\{w' \geq w\right\} \frac{\mathbb{P}\left[m_t = \text{EE}, w_{t+1} \leq w' | x, w_t = w, y_t = y, \Omega_{t-1}\right]}{\mathbb{P}\left[m_t = \text{EE} | x, y_t\right]}$$
$$= \mathbf{1}\left\{w' \geq w\right\} \frac{1 - \lambda^1 \sum_{y'=1}^{Y} \overline{G}^1\left[R_{x,y}(w') - S^+_{x,y'}\right]\frac{v_{y'}}{V}}{1 - \lambda^1 \sum_{y'=1}^{Y} \overline{G}^1\left[S_{x,y} - S^+_{x,y'}\right]\frac{v_{y'}}{V}} := F_{\text{EE}}(w'|x,y,w).$$

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

Firm y_{t+1} Next we turn to y_{t+1}. When $m_t = \text{EU}$, $y_{t+1} = 0$. When $m_t = \text{EE}$, $y_{t+1} = y_t$. When $m_t = \text{UE}$, $\mathbb{P}\left[y_{t+1} = y' | x, m_t = \text{UE}, \Omega_{t-1}\right] \propto \frac{v_{y'}}{V}\overline{G}^0(-S_{x,y'})$. 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

$$M_1(y_1) = \Big[\mathbb{P}[w_1 \leq w_p|x = q, y_1, m_1 = \text{JJ}, m_2 = \text{EE}]\Big]_{p,q},$$

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

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

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

This implies that

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

We then construct

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

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. □

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:

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

where $\mathbb{P}[w_2|x, m_1 = \text{EE}, y_1, w_1]$ is known and the column rank assumption of Lemma B.1 gives that $\mathbb{P}[x|m_1 = \text{EE}, y_1, w_1]$ 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 using transition probabilities from Corollary 1.*

B.5 Parametric G

Assumption 3. $S_{x,y} > 0$*, for all* x, y*.*

Let

$$p_{JJ}(y'|y, x) = \overline{\delta}_{x,y}\lambda^1 \frac{v_{y'}}{V} \overline{G}^1 \left(S_{x,y} - S_{x,y'}\right),$$
$$p_{UE}(y|x) = \lambda^0 \frac{v_y}{V} \overline{G}^0 \left(-S_{x,y}\right),$$

and with $S_{x,y} > 0$,

$$p_{EU}(x, y) = \delta_{x,y}.$$

1) Within-group JJ transitions,

$$p_{JJ}(y'|y',x) = \overline{\delta}_{x,y'}\lambda^1 \frac{v_{y'}}{V}\overline{G}(0),$$

identify

$$\lambda^1 \frac{v_{y'}}{V} = \frac{p_{JJ}(y'|y',x)}{\overline{G}(0)p_{EU}(x,y')},$$

and λ^1 follows by integration.

2) Then,

$$\rho_1\left(S_{x,y} - S_{x,y'}\right) = \overline{G}^{-1}(\theta_1(x,y,y')).$$

for

$$\theta_1(x,y,y') := \frac{\overline{G}(0)p_{JJ}(y'|y,x)}{p_{JJ}(y'|y',x)}.$$

Moreover,

$$p_{UE}(y|x) = \frac{\lambda^0}{\lambda^1}\lambda^1 \frac{v_y}{V}\left[1 - 2G\left(-\rho_0 S_{x,y}\right)\right]$$

yields

$$\rho_0 S_{x,y} = (1-2G)^{-1}\left(\frac{\lambda^1}{\lambda^0}\theta_0(x,y)\right)$$

for

$$\theta_0(x,y) := \frac{p_{UE}(y|x)}{\lambda^1 \frac{v_y}{V}} = \overline{G}(0)\frac{p_{EU}(x,y)p_{UJ}(y|x)}{p_{JJ}(y|y,x)}.$$

3) We can write, for any triple (y_1, y_2, y_3),

$$\frac{(1-2G)^{-1}\left(\frac{\lambda^1}{\lambda^0}\theta_0(x,y_3)\right) - (1-2G)^{-1}\left(\frac{\lambda^1}{\lambda^0}\theta_0(x,y_1)\right)}{(1-2G)^{-1}\left(\frac{\lambda^1}{\lambda^0}\theta_0(x,y_2)\right) - (1-2G)^{-1}\left(\frac{\lambda^1}{\lambda^0}\theta_0(x,y_1)\right)} = \frac{\overline{G}^{-1}\left(\theta_1(x,y_1,y_3)\right)}{\overline{G}^{-1}\left(\theta_1(x,y_1,y_2)\right)},$$

which identifies $\frac{\lambda^1}{\lambda^0}$ if G is not linear. Finally,

$$\begin{aligned}(1-2G)^{-1}\left(\frac{\lambda^1}{\lambda^0}\theta_0(x,y)\right) - (1-2G)^{-1}\left(\frac{\lambda^1}{\lambda^0}\theta_0(x,y')\right) &= \rho_0\left(S_{x,y} - S_{x,y'}\right) \\ &= \frac{\rho_0}{\rho_1}\overline{G}^{-1}\left(\theta_1(x,y,y')\right)\end{aligned}$$

identifies $\frac{\rho_0}{\rho_1}$. Note that this equation caries a lot of information on the shape of G itself. So, we learn a lot from the assumption that G^0 and G^1 have the same shape up to different scale parameters.

At this stage, we have identified v_y/V, λ^1, λ^0, ρ_0 and all $S_{x,y}$'s up to the multiplicative scale ρ_1.

B.5.1 Wage equation

We first show by integration by part in equation (7) that we can write

$$u(w_{x,y}(R)) = c_{x,y} + b_x + \frac{r+\delta_{x,y}}{1+r}R - \frac{\lambda^1\overline{\delta}_{x,y}}{1+r}\sum_{y'=1}^{Y}\left[\int_{R-S_{x,y'}}^{S_{x,y}-S_{x,y'}}\overline{G}^1(\xi)\,\mathrm{d}\xi + \int_{S_{x,y}}^{\infty}\overline{G}^1(\xi)\,\mathrm{d}\xi\right]v_{y'},$$

since

$$\int_{R-S_{x,y'}}^{S_{x,y}-S_{x,y'}} (S_{x,y'} + \xi - R)\, g^1(\xi)\, \mathrm{d}\xi = -\left(S_{x,y} - R\right) \overline{G}^1\left(S_{x,y} - S_{x,y'}\right) + \int_{R-S_{x,y'}}^{S_{x,y}-S_{x,y'}} \overline{G}^1(\xi)\, \mathrm{d}\xi$$

and

$$\begin{aligned}\int_{S_{x,y}-S_{x,y'}}^{\infty} &(\max\{\xi, S_{x,y}\} - R)\, g^1(\xi)\, \mathrm{d}\xi \\ &= \int_{S_{x,y}-S_{x,y'}}^{\infty} (\max\{\xi - S_{x,y}, 0\} + S_{x,y} - R)\, g^1(\xi)\, \mathrm{d}\xi \\ &= \int_{S_{x,y}}^{\infty} \overline{G}^1(\xi)\, \mathrm{d}\xi + (S_{x,y} - R)\, \overline{G}^1\left(S_{x,y} - S_{x,y'}\right).\end{aligned}$$

We finally obtain

$$u(w_{x,y}(R)) = c_{x,y} + b_x + \frac{r + \delta_{x,y}}{1+r} R - \frac{\lambda^1 \overline{\delta}_{x,y} V}{1+r} \left(\sum_{y'=1}^{Y} \left[\int_{R-S_{x,y'}}^{S_{x,y}-S_{x,y'}} \overline{G}(\rho_1 \xi)\, \mathrm{d}\xi \right] \frac{v_{y'}}{V} + \int_{S_{x,y}}^{\infty} \overline{G}(\rho_1 \xi)\, \mathrm{d}\xi \right).$$

Expressions for transition wages follow. Wages out of unemployment occur after drawing ξ from G^0 such that $-S_{x,y} < \xi < 0$, and have $R = 0$ ($\underline{w}_{x,y} = w_{x,y}(0)$):

$$u(\underline{w}_{x,y}) = c_{x,y} + b_x - \frac{\lambda^1 \overline{\delta}_{x,y} V}{1+r} \left(\sum_{y'=1}^{Y} \left[\int_{-S_{x,y'}}^{S_{x,y}-S_{x,y'}} \overline{G}(\rho_1 \xi)\, \mathrm{d}\xi \right] \frac{v_{y'}}{V} + \int_{S_{x,y}}^{\infty} \overline{G}(\rho_1 \xi)\, \mathrm{d}\xi \right).$$

Next, a JJ transition from y to y' occurs after drawing ξ from G^1 such that $\xi > S_{x,y} - S_{x,y'}$, and the hiring wage is $w_{x,y'}(\max\{S_{x,y}, \xi\})$.

$$u(w_{x,y'}(\max\{S_{x,y}, \xi\})) = c_{x,y'} + b_x + \frac{r + \delta_{x,y'}}{1+r} \max\{S_{x,y}, \xi\} - \frac{\lambda^1 \overline{\delta}_{x,y'} V}{1+r} \left(\sum_{y''=1}^{Y} \left[\int_{\max\{S_{x,y},\xi\}-S_{x,y''}}^{S_{x,y'}-S_{x,y''}} \overline{G}(\rho_1 \xi')\, \mathrm{d}\xi' \right] \frac{v_{y''}}{V} + \int_{S_{x,y'}}^{\infty} \overline{G}(\rho_1 \xi')\, \mathrm{d}\xi' \right).$$

In these wage equations, we do not know $c_{x,y} + b_x$, ρ_1, or V. A JJ transition within the same group of firms yields

$$u(w_{x,y}(\max\{S_{x,y}, \xi\})) - u(\underline{w}_{x,y}) = \frac{r + \delta_{x,y}}{1+r} \max\{S_{x,y}, \xi\} + \frac{\lambda^1 \overline{\delta}_{x,y} V}{1+r} \sum_{y'=1}^{Y} \left[\int_{-S_{x,y'}}^{\max\{S_{x,y},\xi\}-S_{x,y'}} \overline{G}(\rho_1 \xi')\, \mathrm{d}\xi' \right] \frac{v_{y'}}{V}.$$

Again, the nonlinearity of the integral allows to identify ρ_1 separately from V. We obtain ρ_1 after eliminating V using two different firm types. Then, $c_{x,y} + b_x$ follows from level wages.

B.5.2 ML estimation of structural parameters

Let $f_{UE}(w|x,y)$ and $f_{JJ}(w|x,y,y')$ denote the PDF of the distributions of transitions wages estimated in the first step. We also allow for Gaussian measurement error with variance ω^2.

We estimate $\theta = \left(\frac{v_y}{V}, \delta_{x,y}, \lambda^1, \lambda^0, \rho_1, \rho_0, S_{x,y}, c_{x,y} + b_x, V\right)$ by maximizing the pseudo likelihood:

$$L = \sum_{x=1}^{X} \ell_x^0 L_x^0 + \sum_{x=1}^{X} \sum_{y=1}^{Y} \ell_{x,y}^1 L_x^1$$

where

$$\begin{aligned} L_x^0 = \sum_y p_{UE}(y|x) \ln\left(\lambda^0 \frac{v_y}{V}(1-2G)\left(-\rho_0 S_{x,y}\right)\right) \\ + p_{UU}(x) \ln\left(1 - \sum_y \lambda^0 \frac{v_y}{V}(1-2G)\left(-\rho_0 S_{x,y}\right)\right) \\ + \sum_y p_{UE}(y|x) \int f_{UE}(w|x,y) \ln \frac{1}{\omega}\phi\left(\frac{u(w) - u(\underline{w}_{x,y})}{\omega}\right) \mathrm{d}u(w) \end{aligned}$$

and

$$\begin{aligned} L^1(x,y) = \sum_{y'} p_{JJ}(y'|x,y) \ln\left(\overline{\delta}_{x,y} \lambda^1 \frac{v_{y'}}{V} \overline{G}\left(\rho_1\left(S_{x,y} - S_{x,y'}\right)\right)\right) \\ + p_{EU}(x,y) \ln \delta_{x,y} \\ + p_{EE}(x,y) \ln\left(1 - \delta_{x,y} - \overline{\delta}_{x,y} \sum_{y'} \lambda^1 \frac{v_{y'}}{V} \overline{G}\left(\rho_1\left(S_{x,y} - S_{x,y'}\right)\right)\right) \\ + \sum_{y'} p_{JJ}(y'|x,y) \int f_{JJ}(w|x,y,y') \times \\ \ln\left(\int_{S_{x,y} - S_{x,y'}}^{\infty} \frac{1}{\omega}\phi\left[\frac{u(w) - u(w_{x,y'}(\max\{S_{x,y}, \xi\}))}{\omega}\right] \frac{\rho_1 g\left(\rho_1 \xi\right)}{\overline{G}\left(\rho_1\left(S_{x,y} - S_{x,y'}\right)\right)} \mathrm{d}\xi\right) \\ \mathrm{d}u(w). \end{aligned}$$

Note that we make no use of within-job wage changes. There are lots of wage cuts on the job that the model does not generate. There is thus little hope of fitting within-spell wage dynamics well. One can however check ex post if the model delivers average wage changes well.

The reason why this works is that if first-stage estimators are consistent, then

$$\sum_y p_{UE}(y|x)\frac{\partial}{\partial\theta}\ln\left(\lambda^0\frac{v_y}{V}(1-2G)\left(-\rho_0 S_{x,y}\right)\right)$$

$$+\,p_{UU}(x)\frac{\partial}{\partial\theta}\ln\left(1-\sum_y\lambda^0\frac{v_y}{V}(1-2G)\left(-\rho_0 S_{x,y}\right)\right)$$

$$=\sum_y\frac{p_{UE}(y|x)}{\lambda^0\frac{v_y}{V}(1-2G)\left(-\rho_0 S_{x,y}\right)}\frac{\partial}{\partial\theta}\left(\lambda^0\frac{v_y}{V}(1-2G)\left(-\rho_0 S_{x,y}\right)\right)$$

$$+\frac{p_{UU}(x)}{1-\sum_y\lambda^0\frac{v_y}{V}\overline{G}\left(-\rho_0 S_{x,y}\right)}\frac{\partial}{\partial\theta}\left(1-\sum_y\lambda^0\frac{v_y}{V}\overline{G}\left(-\rho_0 S_{x,y}\right)\right)$$

$$\simeq\frac{\partial}{\partial\theta}\sum_y\left(\lambda^0\frac{v_y}{V}(1-2G)\left(-\rho_0 S_{x,y}\right)\right)$$

$$+\frac{\partial}{\partial\theta}\left(1-\sum_y\lambda^0\frac{v_y}{V}(1-2G)\left(-\rho_0 S_{x,y}\right)\right)=\frac{\partial 1}{\partial\theta}=0.$$

The fractions in red are asymptotically equation 1 if first stage estimators are consistent. Hence, the true parameter value maximizes

$$\sum_y p_{UE}(y|x)\ln\left(\lambda^0\frac{v_y}{V}(1-2G)\left(-\rho_0 S_{x,y}\right)\right)$$

$$+\,p_{UU}(x)\ln\left(1-\sum_y\lambda^0\frac{v_y}{V}(1-2G)\left(-\rho_0 S_{x,y}\right)\right).$$

Similarly,

$$\sum_y p_{UE}(y|x)\int f_{UE}(w|x,y)\frac{\partial}{\partial\theta}\ln\frac{1}{\omega}\phi\left(\frac{u(w)-u(\underline{w}_{x,y})}{\omega}\right)\mathrm{d}u(w)$$

$$=\sum_y p_{UE}(y|x)\int\frac{f_{UE}(w|x,y)}{\frac{1}{\omega}\phi\left(\frac{u(w)-u(\underline{w}_{x,y})}{\omega}\right)}\frac{\partial}{\partial\theta}\frac{1}{\omega}\phi\left(\frac{u(w)-u(\underline{w}_{x,y})}{\omega}\right)\mathrm{d}u(w)$$

$$\simeq\sum_y p_{UE}(y|x)\frac{\partial}{\partial\theta}\int\frac{1}{\omega}\phi\left(\frac{u(w)-u(\underline{w}_{x,y})}{\omega}\right)\mathrm{d}u(w)=0.$$

Note that we do not need to weigh all y's by $p_{UE}(y|x)$ since it does not depend on θ.

The argument also applies to $L^1(x,y)$.

B.6 Variance Decomposition

We apply the law of total variance, given by:

$$Var(A|C)=\mathbb{E}[Var(A|B,C)|C]+Var(\mathbb{E}[A|B,C]|C).$$

First, by using $A = \log w$, $C = x$, and $B = R$, and defining $\mu_R = \mathbb{E}[\log w|R, x]$, we obtain:

$$Var(\log w|x) = \mathbb{E}[Var(\log w|R, x)|x] + Var(\mu_R|x)$$

Second, we decompose the last term using $A = \mu_R$, $C = x$, and $B = y$ to obtain:

$$Var(\log w|x) = \mathbb{E}[Var(\log w|R, x)|x] + Var(\mathbb{E}(\mu_R|y, x)|x) + \mathbb{E}[Var\left[\mu_R|y, x\right]|x]$$

Finally, one might be concerned that the within R terms incorrectly attribute the within firm variation in delivering R to worker x within firm y to a Rosen compensating differential, which should strictly apply between firms. However, in our model, this within R, x, y variance is 0 theoretical, so we focus directly on the within R term.

Formally, we can apply the decomposition to $Var(\log w|R, x)$ using $A = \log w$, $B = y$, and $C = R, x$ to obtain

$$Var(\log w|R, x) = \mathbb{E}[Var(\log w|R, x, y)|R, x] + Var(\mathbb{E}(\log w|y, R, x)|R, x)$$

In our model, the wage is a deterministic function of R, x, y so the conditional variance is 0, $Var(\log w|R, x, y) = 0$, and hence, in our case, there is only variation between y.

B.7 Elasticities in fixed capacity firm

Consider a firm with n jobs. When jobs are vacant, they meet workers with a hiring probability of $h(w)$. When a job is filled, the match separates with a probability of $q(w)$. We consider the stationary relationship where separations equal hires:

$$h(w)(C - n(w)) = n(w)q(w).$$

This implies that

$$n(w) = C\frac{h(w)}{h(w) + q(w)}.$$

From there, we can derive the elasticity of $n(w)$ from the elasticities of $h(w)$ and $q(w)$:

$$\begin{aligned}\frac{\partial \log n(w)}{\partial \log w} &= \frac{\partial \log h(w)}{\partial \log w} - \frac{\partial \log h(w) + q(w)}{\partial \log w} \\ &= w\frac{h'(w)}{h(w)} - w\frac{h'(w) + q'(w)}{h(w) + q(w)} \\ &= \frac{q(w)}{h(w) + q(w)}\left(\frac{\partial \log h(w)}{\partial \log w} - \frac{\partial \log q(w)}{\partial \log w}\right)\end{aligned}$$

In the Manning type setting, we start from

$$\tilde{n}(w) = \frac{h(w)}{q(w)}$$

which gives immediately that

$$\frac{\partial \log \tilde{n}(w)}{\partial \log w} = \frac{\partial \log h(w)}{\partial \log w} - \frac{\partial \log q(w)}{\partial \log w}$$

B.8 Data, sample, and variable construction

We used the raw matched employer-employee data set constructed in Friedrich et al. (2022). This links information from three data sources made available by The Institute for Evaluation of Labour Market and Education Policy (IFAU).

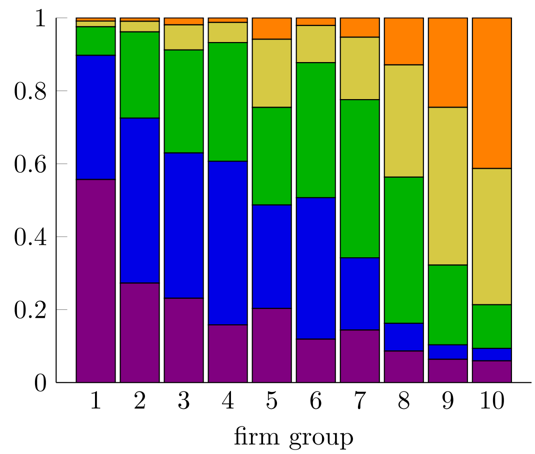
The primary data sources for this study are three datasets. The first is the Longitudinal Database on Education, Income, and Employment (LOUISE), which provides comprehensive information on demographic and socioeconomic variables for the entire working-age population in Sweden, spanning from 1990 to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

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

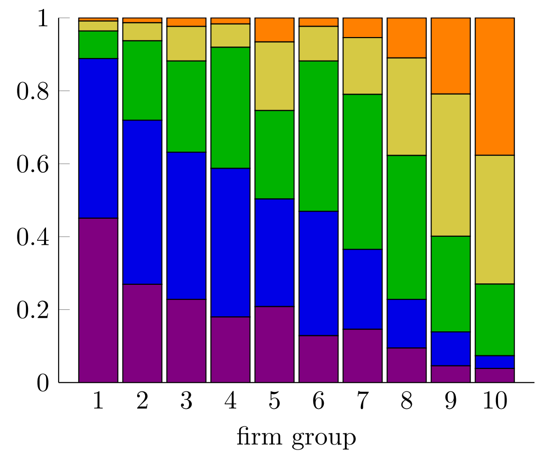
On the firm-related side, RAMS also records information about the industries and types of legal entities for all firms that employ workers. Finally, we draw from the third data source, the Structural Business Statistics (SBS), which encompasses accounting and balance sheet information for all nonfinancial 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 annual levels. 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 Swedish kronor in 2008 are excluded from the sample, and 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 yearly means from the data and limit it to workers under the age of 50.



(a) Model Stationary



(b) Data (Step 1)

Figure 7: Model stationary distribution and distribution in first step.
Notes: comparing distribution implied by the surplus to distribution estimated with empirical model in Step 1.

B.9 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 moment constraints in the text:

$$\begin{aligned}\max_\theta \sum_x \ell^0_x \Big[& \sum_{y'} \mathbb{P}[y_{t+1}{=}y', m_t{=}\text{UE}|x] \times \log \lambda^0 \tilde{v}_{y'} G^0(\rho_0 S_{x,y'}) \\ & + \mathbb{P}[m_t{=}\text{UU}|x] \times \log \Big(1 - \sum_{y'} \lambda^0 \tilde{v}_{y'} G^0(\rho_0 S_{x,y'})\Big)\Big] \\ + \sum_{x,y} \ell^1_{x,y} \bar{\delta}_{x,y} \Big[& \sum_{y'} \mathbb{P}[y_{t+1}{=}y', m_t{=}\text{JJ}|y_t{=}y, x] \times \log \lambda^1 \tilde{v}_{y'} G^1(\rho_1 S_{x,y'} - \rho_1 S_{x,y}) \\ & + \mathbb{P}[m_t{=}\text{EE}|y_t{=}y, x] \times \log \Big(1 - \sum_{y'} \lambda^1 \tilde{v}_{y'} G^1(\rho_1 S_{x,y'} - \rho_1 S_{x,y})\Big)\Big]\end{aligned}$$

$$s.t. \quad m_1(\theta) = m_1, m_2(\theta) = m_2$$

where $\mathbb{P}[y_{t+1}{=}y', m_t{=}\text{UE}|x]$, $\mathbb{P}[m_t{=}\text{UU}|x]$, $\mathbb{P}[y_{t+1}{=}y', m_t{=}\text{JJ}|y_t{=}y, x]$, $\mathbb{P}[m_t{=}\text{EE}|y_t{=}y, x]$ and $\ell^1_{x,y}$ and ℓ^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.

B.10 Estimated parameters

References Online Appendix

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

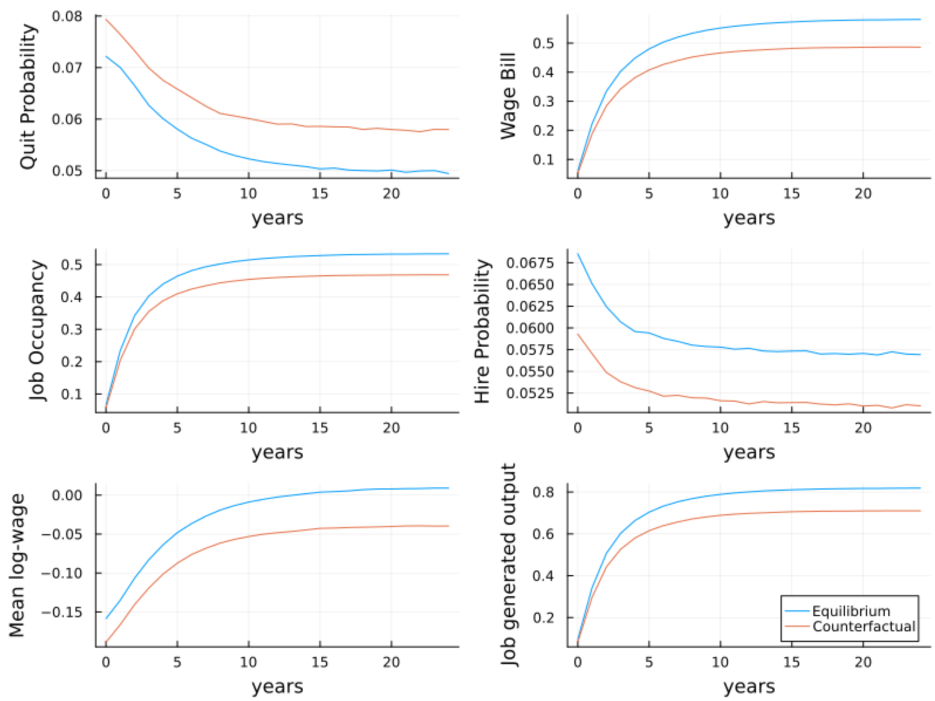


Figure 8: Dynamics within job: Equilibrium and Bidding Deviation

Notes: We conduct the following model experiment. First we start by sampling a large number of vacant jobs drawn from v_y. We then simulate forward as these vacancies become filled and eventually converge to the stationary equilibrium. The blue lines in each panel are the average outcome across these jobs for (a) the sparation probability, (b) total wage bill, (c) job occupancy rate, (d) hiring probability, (e) mean log wage, and (f) match output. In each case the outcome has converged after 20 years. Next, we simulate the same path but assume that each firm uses a strategy $B = (1 - \Delta)S$ when competing for workers with other firms bidding the equilibrium $B' = S_{x,y'}$. We then use the differences, for $\Delta = 0.05$, in the generated average wage, hiring probability, and separation probability to calculate the elasticities presented in Table 2.

Table 3: Model Parameters

Firm group:	1	2	3	4	5	6	7	8	9	10
	Production $f_{x,y}$									
1	0.66	0.99	1.03	1.32	1.24	1.39	1.45	1.18	1.12	1.46
2	0.84	1.06	1.12	1.31	1.20	1.40	1.37	1.21	1.09	1.07
3	1.03	1.22	1.31	1.48	1.46	1.59	1.58	1.51	1.58	1.66
4	1.39	1.64	1.73	1.92	1.85	1.97	1.96	1.87	1.96	2.11
5	2.03	2.23	2.79	2.74	2.82	2.98	2.80	2.73	2.95	2.97
	Amenity $\tilde{c}_{x,y}$									
1	-0.85	-0.74	-0.70	-0.57	-0.49	-0.49	-0.36	-0.02	0.09	0.41
2	-0.34	-0.33	-0.27	-0.23	-0.22	-0.18	-0.09	0.02	0.08	0.09
3	-0.06	-0.13	-0.08	-0.11	-0.01	-0.08	-0.01	0.10	0.26	0.39
4	0.36	0.30	0.30	0.31	0.29	0.28	0.31	0.35	0.38	0.55
5	0.79	0.71	0.87	0.77	0.76	0.78	0.71	0.73	0.73	0.80
	Separation $\delta_{x,y}$									
1	0.10	0.10	0.11	0.09	0.10	0.09	0.10	0.09	0.08	0.09
2	0.01	0.01	0.01	0.01	0.01	0.00	0.00	0.01	0.01	0.01
3	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
4	0.01	0.01	0.01	0.00	0.01	0.00	0.00	0.01	0.01	0.01
5	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.02
	Common parameters									
r	0.0125									
ρ_0	0.108									
ρ_1	0.460									
κ	0.086									
Firm Mass	1.88									

Note: All the parameters needed to solve and simulate the model. See graphical versions in 3.a,b, and 1.g.

Table 4: Two-way linear decomposition in the model

Total variance of log wages	0.1408
Percent of total variance explained by	
Worker effects	49.52%
Firm effects	9.65%
Covariance of worker and firm	20.18%
Non-linear component	1.36%
Residual	19.29%

Notes: log wage variance decomposition in the cross-section generated by the model. We project the model-simulated wages on a dummy for each x, a dummy for each y and joint dummy x, y.

Table 5: Statistical variance decomposition of wages just after a move.

Variance of log poaching wages	0.1432
Percent of total variance explained by	
Worker effects	49.75%
Destination firm effects	8.51%
Origin effects	0.47%
Covariance of worker, destination	20.04%
Covariance of worker, origin	-2.12%
Covariance of destination, origin	-0.50%
Residual	23.84%

Note: This table is based on model simulated data and mirrors Table 5 of Di Addario et al. (2023). The numbers are quantitatively similar with meaningful shares attributed to worker effects, destination effect and sorting on destination, with very little effect from firm of origin, despite wages being generated from sequential auctions.

Friedrich, B., L. Laun, C. Meghir, and L. Pistaferri (2022): "Earnings dynamics and firm-level shocks," Tech. rep., NBER wp25786.

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," in *Handbook of Labor Economics*, Elsevier, vol. 1 of *Handbook of Labor Economics*, 641–692.