# Assortative Matching and Wages: The Role of Selection*

Katarína Borovičková
FRB Richmond

Robert Shimer
University of Chicago

August 31, 2024

**Abstract**

We develop a random search model of the labor market with two-sided heterogeneity and match-specific productivity shocks. Our model has two robust predictions: i) the average log wage that a worker receives is increasing in the worker's and employer's productivity and ii) there is positively assortative matching between high wage workers and high wage firms. Wages and sorting are driven by the same selection force. All workers are equally likely to meet all firms, but low (high) productivity workers have a higher average surplus from meeting low (high) productivity firms. The high surplus meetings result in matches more frequently, generating positive assortative matching. Since only meetings that result in matches are observed in administrative wage data, such data contain only a selected subset of meetings, driving the result that average log wages are increasing and submodular. We show that our findings are consistent with recent results in the empirical wage literature.

*We are grateful for comments from Fernando Alvarez, Stéphane Bonhomme, Jaroslav Borovička, Kyle Herkenhoff, Philipp Kircher, Thibaut Lamadon, Guido Menzio, Simon Mongey, Christian Moser, Benjamin Schoefer, and Isaac Sorkin, as well as workshop participants at Duke University, the Minnesota Macro Workshop, New York University, the NBER Summer Institute (Micro and Macro Perspectives of the Aggregate Labor Market), the SITE Summer Conference (Micro and Macro of Labor Markets), Universidad Andres Bello, and the University of Chicago. Any remaining errors are our own.

# 1 Introduction

Why do high-wage workers tend to work at high-wage firms, even though low-wage workers seem to gain at least as much from such matches? Bonhomme, Lamadon and Manresa (2019) write that these patterns are "difficult to reconcile with models where sorting is driven by complementaries in production, as in Becker (1973) [p. 701]." We address this puzzle by developing a theory of frictional labor markets that highlights the crucial role of selective job acceptance decisions for observed wages.

We develop a random search framework where workers and firms are heterogeneous in their productivity. Output in a match depends on the worker's productivity, the firm's productivity, and a match-specific shock. The decision to form a match depends on the realization of this shock relative to the outside options of both the worker and the firm, with wages determined by Nash bargaining. This setup allows us to analyze how selection in the matching process affects observed wage patterns and sorting outcomes.

We prove that our model can generate outcomes that align with the empirical literature. We find conditions which ensure that (i) average log wages increase with both worker and firm productivity, while possibly being submodular—meaning low-productivity workers' wages are more sensitive to firm type than high-productivity workers' wages; (ii) there is positively assortative matching between worker and firm productivities; and (iii) high-productivity workers are high-wage workers and high-productivity firms are high-wage firms.

Our theoretical results arise from a novel mechanism, the selective nature of realized matches. While all workers are equally likely to meet all firms, the likelihood of a match forming depends on the interplay between worker and firm types. Low-productivity workers rarely receive acceptable offers from high-productivity firms, but when they do, these offers are exceptionally good. Similarly, high-productivity workers seldom accept offers from low-productivity firms unless the match-specific productivity is unusually high. This selection process creates a subset of realized matches and wages that differs systematically from the set of all meetings, driving both the observed sorting patterns and the behavior of average log wages.

Our theoretical findings are consistent with a growing body of empirical research on log wages and worker-firm sorting. Studies by Card, Heining and Kline (2013); Card, Cardoso and Kline (2016); Card, Cardoso, Heining and Kline (2018); Song, Price, Guvenen, Bloom and Von Wachter (2019); Bonhomme, Lamadon and Manresa (2019); Kline, Saggio and Sølvsten (2020); Bonhomme, Holzheu, Lamadon, Manresa, Mogstad and Setzler (2023), among others, have found that average log wages increase with both worker and firm type and are approximately additively separable. These papers have also found that there is

strong correlation between matched worker and firm types. We show that our sparsely-parameterized model is able to match the behavior of both wages and sorting. In particular, our model can match the variance decomposition of wages into the part attributable to worker heterogeneity, to firm heterogeneity, to worker-firm sorting, and to a residual. We demonstrate this in two versions of the model, one where only unemployed workers search for jobs and one with on-the-job search.

Our final contribution is to distinguish between the surplus derived from meetings and the surplus derived from matches. The surplus from matches can readily be measured using administrative wage data. We prove that the surplus from meetings is identified using administrative wage data and the assumption of random search. We also derive an expression for the surplus from meetings. This analysis reveals that the surplus from meetings is supermodular, with low-type workers actually preferring to meet intermediate-type firms rather than high-type firms. Importantly, we show that this identification result and estimator hinge on the random search assumption. This finding underscores the importance of carefully considering search frictions when interpreting wage data and highlights a novel aspect of our model's implications for labor market dynamics.

Our paper contributes to several strands of literature. First, we build on empirical work using matched employer-employee data to study wage determination and sorting (Abowd, Kramarz and Margolis, 1999; Andrews, Gill, Schank and Upward, 2008; Card, Heining and Kline, 2013; Card, Cardoso, Heining and Kline, 2018; Song, Price, Guvenen, Bloom and Von Wachter, 2019; Bonhomme, Lamadon and Manresa, 2019; Bonhomme, Holzheu, Lamadon, Manresa, Mogstad and Setzler, 2023; Kline, Saggio and Sølvsten, 2020). While these papers document important empirical regularities, we provide a theoretical framework that can explain these patterns through a novel mechanism.

Our theory cautions against the prevailing interpretation of these empirical regularities. For example, Card, Heining and Kline (2013) regress wages on worker and establishment fixed effects, with a focus on measuring "establishment-specific wage premiums" $\psi_j$. They "interpret the establishment effect $\psi_j$ as a proportional pay premium (or discount) that is paid by establishment $j$ to all employees (i.e., all those with $\mathcal{J}(i,t) = j$). Such a premium could represent rent-sharing, an efficiency wage premium, or strategic wage posting behavior [p. 987]." We argue that selection may shape the establishment-specific wage premium.

Second, we contribute to the theoretical literature on conditions for assortative matching in labor markets with random search (Shimer and Smith, 2000; Hagedorn, Law and Manovskii, 2017; Lopes de Melo, 2018; Bagger and Lentz, 2019). Our model extends this literature by incorporating match-specific shocks, as in Goussé, Jacquemet and Robin (2017)'s model of the marriage market, and highlighting the role of selection in job acceptance deci-

sions. Our finding that there is positively assortative matching if the production function is strictly increasing and weakly log-supermodular recalls similar results in competitive search models (Shi, 2001, 2005; Shimer, 2005; Eeckhout and Kircher, 2010) and models with non-transferable utility (Smith, 2006; Bonneton and Sandmann, 2023). The details of the mechanisms are quite different, and in particular selection does not play a role in those papers. Those earlier papers also do not try to fit the empirical findings on how wages depend on worker and firm types. Our model bridges this gap by providing a framework that not only explains sorting patterns but also shows how the same selection force that generates sorting drives the observed wage patterns in administrative data sets.

Third, the existing literature recognizes that endogenous mobility can create a selection bias in estimates of the AKM wage equation. Following Card, Heining and Kline (2013), much of the literature has focused on selection working through job-to-job selection transitions, as in Burdett and Mortensen (1998) and Lopes de Melo (2018). We show why existing tests for selection cannot detect the type we highlight here. Moreover, building on an insight from Flinn and Heckman (1982), we explain why it is impossible to construct such a test without auxiliary assumptions, such as random search.

By highlighting the role of selection in job acceptance, our model offers a new interpretation of empirical wage patterns without requiring assumptions about unobserved job amenities or worker preferences for similar peers. This perspective has important implications for understanding labor market dynamics and interpreting wage data. The rest of the paper is structured as follows. Section 2 presents our model. In Section 3, we derive conditions for the average log wage to be an increasing and submodular function of worker and firm type. Section 4 derives conditions for assortative matching both of high productivity workers to high productivity jobs and of high wage workers to high wage jobs. Section 5 shows that our model can quantitatively match the wage decomposition in Bonhomme, Lamadon and Manresa (2019). Section 6 extends our model to allow for on-the-job search, whereby employed workers can systematically climb a job ladder. Section 7 discusses existing tests for exogenous mobility and shows how our model fares in those tests. Section 8 compares the surplus from meetings with the surplus from matches. Section 9 concludes with a discussion of how our results affect our understanding of the empirical wage regression literature.

# 2 Model

We formulate a search model with two-sided heterogeneity (Shimer and Smith, 2000) and match-specific heterogeneity (Goussé, Jacquemet and Robin, 2017). The model is formulated in continuous time and we focus on steady states and so drop time arguments in what follows.

## 2.1 Assumptions

There is measure $M$ of risk-neutral workers and measure $N$ of risk-neutral firms. Everyone discounts the future at rate $r > 0$. There are $X$ worker types indexed by $x = 1, \ldots, X$. The population measure of type-$x$ workers is $m_x > 0$, with $\sum_{x=1}^{X} m_x = M$. There are $Y$ firm types indexed by $y = 1, \ldots, Y$, with population measure $n_y > 0$ and $\sum_{y=1}^{Y} n_y = N$. Workers can be either unemployed or matched to one firm; likewise, firms can be either vacant or matched to one worker. Thus, in this model, a firm and a job are identical.

Search is random and only unmatched firms and workers can search.[1] Let $u_x$ be the population measure of unemployed type-$x$ workers, so that $\frac{u_x}{m_x}$ is the unemployment rate for type $x$. Similarly, let $v_y$ be the population measure of type-$y$ vacancies, with $\frac{v_y}{n_y}$ is the vacancy rate of $y$. All unemployed workers contact vacant type-$y$ firms according to a Poisson process with arrival rate $\rho v_y$ for $y = 1, \ldots, Y$, where $\rho > 0$. Likewise, all vacant firms contact unemployed type-$x$ workers at rate $\rho u_x$ for $x = 1, \ldots, X$.[2]

When a worker and firm meet, they draw a match-specific productivity shock $z$ from a distribution function with density $s(z)$ and survival function $S(z)$. Draws are independent across matches for every worker and firm. By definition, the survival function $S : \mathbb{R}_+ \to \mathbb{R}_+$ is non-increasing. For expositional simplicity, we also assume that $S$ is continuous and strictly positive for all $z > 0$. Finally, we assume the mean productivity draw, $z^* \equiv \int_0^\infty z s(z) dz$, is finite, which is necessary for existence of an equilibrium.

After an unemployed type-$x$ worker meets a vacant type-$y$ firm and draws a match-specific productivity shock $z$, they decide whether to match. If they match, they stop searching and produce flow output $z f_{x,y}$. A matched worker and firm split the surplus according to Nash bargaining, with worker's bargaining power equal to $\gamma \in (0, 1)$. We assume $f_{x,y} > 0$ for all $x$ and $y$. Matches end at rate $\delta > 0$, leaving the worker unemployed and the job vacant.

## 2.2 Value Functions

We start by formulating the value functions of workers and firms. For a type-$x$ unemployed worker, let the value be $V_x^u$:

$$r V_x^u = \rho \sum_{y=1}^{Y} v_y \int_0^\infty \max\left\{ V_{x,y}^e(z, W_{x,y}(z)) - V_x^u, 0 \right\} s(z) dz. \tag{1}$$

[1] In Section 6, we extend the model to allow for search by employed workers and show numerically that our main conclusions carry over to this more realistic environment.

[2] This is a quadratic matching technology (Diamond and Maskin, 1979), so the total number of matches is a homogeneous of degree two in unemployment $\{u_x\}_{x=1}^X$ and vacancies $\{v_y\}_{y=1}^Y$. It is straightforward to prove Propositions 2–5 with a linear matching technology (homogeneous of degree one), at the cost only of additional notational complexity. Similarly, our numerical results carry over exactly to an environment with

At rate $\rho v_y$, the worker meets a vacant type $y$ firm. They then draw match-specific productivity $z$ from a distribution with density $s(z)$. After that, they decide whether to match. If they do, the worker's value jumps to $V^e_{x,y}(z, W_{x,y}(z))$, the value of a type $x$ worker matched to a type $y$ firm in a match with productivity $z$ and earning the equilibrium wage $W_{x,y}(z)$. Nash bargaining implies that both parties agree on whether to match, so the worker matches whenever the value of being in the match, $V^e_{x,y}(z, W_{x,y}(z))$, exceeds the value of being unmatched, $V^u_x$.

Once the worker is in the match, we have the corresponding Bellman equation

$$rV^e_{x,y}(z, W) = W + \delta(V^u_x - V^e_{x,y}(z, W)). \tag{2}$$

This equation describes a type $x$ worker at a type $y$ firm with match-specific shock $z$ earning an arbitrary wage $W$. The worker earns the wage until the match ends exogenously. This implies that the worker will accept the match ($V^e_{x,y}(z, W) \geq V^u_x$) if and only if $W \geq rV^u_x \equiv \bar{w}_x$, the worker's reservation wage.

The Bellman equations for firms are symmetric:

$$rV^v_y = \rho \sum_{x=1}^{X} \int_0^\infty u_x \max\left\{V^f_{y,x}(z, W_{x,y}(z)) - V^v_y, 0\right\} s(z)dz \tag{3}$$

$$rV^f_{y,x}(z, W) = zf_{x,y} - W + \delta(V^v_y - V^f_{y,x}(z, W)). \tag{4}$$

Notably, a type $y$ firm earns flow profit $zf_{x,y} - W$ when employing a type $x$ worker in a match with productivity $z$ and paying a wage $W$. The firm will accept the match ($V^f_{y,x}(z, W) \geq V^v_y$) if and only if $zf_{x,y} - W \geq rV^v_y \equiv \bar{\pi}_y$, the firm's reservation profit.

We define the match surplus as

$$V^s_{x,y}(z) \equiv V^e_{x,y}(z, W) + V^f_{x,y}(z, W) - V^u_x - V^v_y = \frac{\max\{zf_{x,y} - \bar{w}_x - \bar{\pi}_y, 0\}}{r + \delta}, \tag{5}$$

where the second equation follows from equations (2) and (4) and the definitions of the reservation wage $\bar{w}_x$ and reservation profit $\bar{\pi}_y$. Equivalently, define the reservation productivity level $\bar{z}_{x,y}$:

$$\bar{z}_{x,y} \equiv \frac{\bar{w}_x + \bar{\pi}_y}{f_{x,y}}. \tag{6}$$

Then

$$V^s_{x,y}(z) = \frac{f_{x,y} \max\{z - \bar{z}_{x,y}, 0\}}{r + \delta}. \tag{7}$$

a linear matching technology. This is a consequence of some normalizations that we discuss in Appendix D.2 in the context of our more general model with on-the-job search.

The match surplus is positive, so there exists a $W$ which is acceptable both to the worker ($V^e_{x,y}(z,W) \geq V^u_x$) and firm ($V^f_{y,x}(z,W) \geq V^v_y$), if and only if productivity exceeds the reservation level $\bar{z}_{x,y}$.

Finally, when the match surplus is positive, we use Nash bargaining to pin down the equilibrium wage $W_{x,y}(z)$:

$$W_{x,y}(z) = \arg\max_W (V^e_{x,y}(z,W) - V^u_x)^\gamma (V^f_{y,x}(z,W) - V^v_y)^{1-\gamma}. \tag{8}$$

Using equations (2) and (4), as well as the definition of the reservation productivity level in equation (6), it is straightforward to show that this implies

$$W_{x,y}(z) = \bar{w}_x + \gamma f_{x,y}(z - \bar{z}_{x,y}). \tag{9}$$

Putting this together with the fact that a worker and firm match whenever $z \geq \bar{z}_{x,y}$, we get that type $x$ workers match with type $y$ firms at rate $\rho u_x v_y S(\bar{z}_{x,y})$, with matches formed whenever $W_{x,y}(z) \geq \bar{w}_x$, a type-$x$ worker's reservation wage. Symmetrically, a type-$y$ firm agrees to match with a type-$x$ worker whenever $zf_{x,y} - W_{x,y}(z) \geq \bar{\pi}_y$, a type-$y$ firm's reservation profit.

We can now combine these equations to get a simpler expression for a worker's reservation wage and a firm's reservation profit. Eliminate the wage from equation (2) using equation (9) to get $V^e_{x,y}(z, W_{x,y}(z)) - V^u_x = \frac{\gamma f_{x,y}(z - \bar{z}_{x,y})}{r+\delta}$. Substitute that into equation (1) to get

$$\bar{w}_x = \frac{\gamma\rho}{r+\delta} \sum_{y=1}^{Y} v_y f_{x,y} \int_{\bar{z}_{x,y}}^{\infty} (z - \bar{z}_{x,y}) s(z) dz. \tag{10}$$

Analogous steps lead to the equation for a firm's reservation profit:

$$\bar{\pi}_y = \frac{(1-\gamma)\rho}{r+\delta} \sum_{x=1}^{X} u_x f_{x,y} \int_{\bar{z}_{x,y}}^{\infty} (z - \bar{z}_{x,y}) s(z) dz. \tag{11}$$

## 2.3 State Variables and Equilibrium

To close the model, we need to find the steady state values of $u_x$ and $v_y$. To do this, we first define the steady state measure of $(x, y)$ matches, $\phi_{x,y}$. This satisfies

$$\delta\phi_{x,y} = \rho u_x v_y S(\bar{z}_{x,y}). \tag{12}$$

The left hand side is the rate that these matches end, while the right hand side is the rate that unmatched type $x$ agents (measure $u_x$) meet unmatched type $y$ agents ($\rho v_y$) in a match with an acceptable $z$ (share $S(\bar{z}_{x,y})$).

By adding partner types, we can then recover the unemployment and vacancy measures:

$$u_x = m_x - \sum_{y=1}^{Y} \phi_{x,y} \tag{13}$$

$$v_y = n_y - \sum_{x=1}^{X} \phi_{x,y}. \tag{14}$$

A steady state equilibrium is given by $(\bar{w}, \bar{\pi}, \bar{z}, \phi, u, v)$ satisfying equations (6), (10), (11), (12), (13), and (14). We can prove

**Proposition 1** *An equilibrium exists. In any equilibrium, the reservation wage $\bar{w}_x$ and reservation profit $\bar{\pi}_y$ are strictly positive for all $x$ and $y$.*

All the proofs are in Appendix A.

We note that a model like this may have multiple equilibria (Burdett and Coles, 1997). All of our claims apply to any steady state equilibrium.

## 2.4 Monotonicity

We first prove a useful preliminary result, that the reservation wage and reservation profit are increasing if the production function is increasing:

**Lemma 1** *Assume $f_{x,y}$ is strictly increasing in $x$ and $y$. Then the reservation wage $\bar{w}_x$ and reservation profit $\bar{\pi}_y$ are strictly increasing.*

The proof is in Appendix A.

We note that since $\bar{w}_x$ and $\bar{\pi}_y$ are strictly positive, equation (6) implies $\bar{z}_{x,y}$ is strictly positive as well. And since there are a finite number of types of workers and firms, $\underline{z} \equiv \min_{x,y} \bar{z}_{x,y}$ is strictly positive as well. It follows from the definition of equilibrium that in any steady state equilibrium, the behavior of $S(z)$ at $z < \underline{z}$ does not affect the equilibrium $(\bar{w}, \bar{\pi}, \bar{z}, u, v)$. We build on this observation in Section 3.1 below.

# 3 Average Log Wage

Define the *average log wage* in an $(x, y)$ match:

$$w^*_{x,y} \equiv \frac{\int_{\bar{z}_{x,y}}^{\infty} \log(W_{x,y}(z))s(z)dz}{S(\bar{z}_{x,y})}. \tag{15}$$

We think of the average log wage as a reduced-form wage equation. This is in contrast to the structural wage equation (9), which defines the wage as a function of the worker type $x$, the firm type $y$, and the match-specific shock $z$. The $z$-contingent decision to form an $(x, y)$ match, described by the threshold $\bar{z}_{x,y}$ in equation (6), influences the behavior of $w^*_{x,y}$. We illustrate this influence in this section by combining these two equations with some distributional assumptions on the match-specific shock $z$.

The average log wage is of considerable empirical interest. Bonhomme, Lamadon and Manresa (2019) prove that under reasonable conditions, the average log wage is identified in administrative wage data. They also estimate it using Swedish data and find that the average log wage is increasing in worker and firm type and perhaps slightly submodular, meaning that the lowest wage workers get a slightly larger increase in their wage by moving to more productive jobs.

The average log wage is related to the more familiar log-linear wage equation proposed by Abowd, Kramarz and Margolis (1999), as we discuss in Appendix B. In short, that paper proposed that the log wage $w_{i,t}$ of worker $i$ at firm $j$ in period $t$ can be expressed as

$$w_{i,t} = \alpha_i + \psi_j + \varepsilon_{i,t}, \tag{16}$$

where $\varepsilon_{i,t}$ is a random variable with mean zero for all $i$, $j$, and $t$. Under these conditions, $w^*_{x_i,y_j} = \alpha_i + \psi_j$, where $x_i$ is the type of worker $i$ and $y_j$ is the type of firm $j$. Our formulation allows for non-separabilities in $w^*$.

## 3.1 Pareto-Distributed Match Quality

We first characterize wage behavior under the assumption that $z$ has a Pareto distribution with minimum value $z_0 > 0$ and tail parameter $\theta > 1$, so $S(z) = (z/z_0)^{-\theta}$ for $z \geq z_0$ and $S(z) = 1$ otherwise. We also parameterize the contact rate as $\rho = \bar{\rho} z_0^{-\theta}$, so the rate that an unemployed worker contacts a vacant type-$y$ firm and has productivity at least $z$ is $\bar{\rho} z^{-\theta} v_y$ for $z > z_0$. Notably, this is independent of $z_0$. We focus throughout on cases where there is an interior solution for the threshold $\bar{z}_{x,y}$ for all pairs $(x, y)$, $z_0 < \underline{z} = \min_{x,y} \bar{z}_{x,y}$, as will be

the case if $z_0$ is sufficiently small.[3] Under these conditions, we have the following result:

**Proposition 2** *Assume $S(z) = (z/z_0)^{-\theta}$ with $z_0 > 0$ and $\theta > 1$. Then*

$$w^*_{x,y} \equiv \int_0^\infty \log\left(\bar{w}_x + \gamma(\bar{w}_x + \bar{\pi}_y)q\right)\theta(1+q)^{-\theta-1}dq. \tag{17}$$

*Moreover,*

1. *for all $x_1$ and $x_2$ with $\bar{w}_{x_1} < \bar{w}_{x_2}$, $w^*_{x_1,y} < w^*_{x_2,y}$ for all $y$;*
2. *for all $y_1$ and $y_2$ with $\bar{\pi}_{y_1} < \bar{\pi}_{y_2}$, $w^*_{x,y_1} < w^*_{x,y_2}$ for all $x$;*
3. *for all $x_1$, $x_2$, $y_1$, and $y_2$ with $\bar{w}_{x_1} < \bar{w}_{x_2}$ and $\bar{\pi}_{y_1} < \bar{\pi}_{y_2}$, $w^*_{x_1,y_2} + w^*_{x_2,y_1} > w^*_{x_1,y_1} + w^*_{x_2,y_2}$.*

*In particular, if $f_{x,y}$ is strictly increasing, the average log wage is strictly increasing and strictly submodular.*

Our proof in Appendix A allows for any transformation of the wage, not just the log. We focus in the text on the average log wage because this is the focus of the empirical literature.

There are many pieces to unpack from Proposition 2. First, the average log wage depends only on four numbers: the worker's reservation wage $\bar{w}_x$, the firm's reservation profits $\bar{\pi}_y$, the worker's bargaining power $\gamma$, and the Pareto tail parameter $\theta$. The production technology $f_{x,y}$ does not explicitly enter this expression. We view this as both good news and bad news for empirical research. The bad news is that average log wages are not useful for learning about the production technology $f$. The good news is that the model makes strong and testable predictions for how the average wages behaves across different types of matches. We discuss these predictions in Section 8.

Second, workers who have a higher reservation wage earn more at any type of employer, and similarly firms that have a higher reservation profit level pay more to any type of worker. To understand the power of this result, it is useful to consider a non-monotonic production function $f$. That is, suppose that there are worker types $x_1$ and $x_2$ and firm types $y_1$ and $y_2$ such that $f_{x_1,y_1} > f_{x_1,y_2}$ and $f_{x_2,y_1} < f_{x_2,y_2}$. One might conjecture that with Nash bargaining, this non-monotonicity would imply a similar ordering of average log wages, $w^*_{x_1,y_1} > w^*_{x_1,y_2}$ and $w^*_{x_2,y_1} < w^*_{x_2,y_2}$. Proposition 2 establishes that this cannot happen: $\bar{\pi}_{y_2} \gtreqless \bar{\pi}_{y_1} \Rightarrow w^*_{x,y_2} \gtreqless w^*_{x,y_1}$ for all $x$.

Proposition 2 also implies that all workers have the same ranking of average log wages

[3]For given $\bar{\rho}$, a change in $z_0$ does not affect the equilibrium allocation as long as $z_0 \leq \underline{z}$. Thus focusing on small values of $z_0$ does not change the arrival rate of "good" matches. We could sidestep any discussion of $z_0$ by assuming that meetings with match quality at least equal to $z$ occur at rate $\bar{\rho}z^{-\theta}$ times the relevant vacancy or unemployment rate for any $z$, as in Oberfield (2018) and Buera and Oberfield (2020).

across firm types $y$, determined by the reservation profit $\bar{\pi}_y$. This result stands in contrast to one in Shimer and Smith (2000), a similar model except that $z$ has a degenerate distribution. In that case, an unproductive worker's wage may be maximized at low-type firm, while a more productive worker's wage is maximized at a high-type firm. See Lopes de Melo (2018) for an elaboration of this observation.

Finally, we prove that the average log wage is submodular in the worker's reservation wage $\bar{w}_x$ and the firm's reservation profits $\bar{\pi}_y$. That means that the average increase in the log wage of a given worker who moves from a low $\bar{\pi}_y$ firm to a high $\bar{\pi}_y$ firm is decreasing in the worker's reservation wage $\bar{w}_x$. Again, this places strong restrictions for how wages vary across different types of matches.

## 3.2 Exponentially-Distributed Match Quality

Next we characterize the average log wage under the assumption that $z$ has an exponential distribution with parameter $\theta$, so $S(z) = e^{-\theta z}$ for all $z \geq 0$.

**Proposition 3** *Assume $S(z) = e^{-\theta z}$ with $\theta > 0$. Then*

$$w^*_{x,y} = \theta \int_0^\infty \log(\bar{w}_x + \gamma q f_{x,y}) e^{-\theta q} dq. \tag{18}$$

*Assume that for all $x_1 < x_2$ and $y$, $f_{x_1,y} < f_{x_2,y}$. Then $w^*_{x_1,y} < w^*_{x_2,y}$. Additionally, take any $x$, $y_1$, and $y_2$. Then $w^*_{x,y_1} \gtreqless w^*_{x,y_2}$ if and only if $f_{x,y_1} \gtreqless f_{x,y_2}$.*

*If $f_{x,y} = f^1_x f^2_y$ for strictly increasing functions $f^1$ and $f^2$, then $w^*_{x,y}$ is strictly increasing and strictly submodular.*

Our proof of the first part of the Proposition in Appendix A allows for any transformation of the wage, not just the log. The result for multiplicatively-separable production function relies on the logarithmic assumption.

Some of these results parallel the earlier results for a Pareto distribution. In particular, if $f$ is strictly increasing in both $x$ and $y$, then so is the average log wage. Thus the model predicts that more productive workers are paid higher wages at any firm, and that more productive firms pay higher wages to any worker.

On the other hand, when $S$ is exponential, the average log wage depends on the entire production function $f$, not just on the reservation wage and profit, as was the case with a Pareto. This allows for some interesting possibilities. For example, a type-$x$ worker's average log wage at a type-$y$ firm is increasing in $f_{x,y}$, and so different types of workers may have a different ordering of average log wages across firm types.

In one important case, we can prove that the average log wage is both increasing and submodular with an exponential distribution, exactly as we established for the Pareto distribution. Assume that $f_{x,y}$ is strictly increasing and multiplicatively separable. Proposition 3 establishes that the average log wage is increasing and submodular. Thus once again we see that complementarity in the production function does not necessarily carry over to the average log wage.

## 3.3 Other Distributions

A natural question is whether monotonicity of the production function $f$ guarantees monotonicity of the average log wage for other distributions of the idiosyncratic shock $z$. In general, the answer is no. First, if the distribution of idiosyncratic productivity shocks is degenerate, wages are generally a hump-shaped function of a firm's type for a given worker (Shimer and Smith, 2000; Lopes de Melo, 2018). Second, we find that if $z$ has a normal distribution, possibly truncated at zero, then the average log wage for some workers may be decreasing in firm type even if the production function is strictly increasing. Still, it is easy to construct other examples where both the production function and wage are monotone in both the worker and firm types, and for this reason we believe that the special cases we highlight here are useful for understanding more general properties of the model.

## 3.4 Selection in the Wage Equation

To understand what drives the behavior of the average log wage, we contrast it with another wage equation. Equation (9) describes how wages depend on the worker's type $x$, the firm's type $y$, and the idiosyncratic shock $z$. This is a structural relationship which comes from assumption that wages are set by Nash bargaining. To see how this behaves, we use the threshold equation (6) to rewrite it as

$$W_{x,y}(z) = (1-\gamma)\bar{w}_x + \gamma z f_{x,y} - \gamma\bar{\pi}_y.$$

Assuming that $f_{x,y}$ is increasing in $y$, Lemma 1 tells us that $\bar{\pi}_y$ is increasing as well. Then for small $z$, $W_{x,y}(z)$ is necessarily decreasing in $y$. This may also be true at an average level of the idiosyncratic shock, $z = z^* \equiv \int_0^\infty zs(z)dz$, particularly for low worker types $x$ when there are complementarities in production, i.e. when $f_{x,y}$ is supermodular. Shimer and Smith (2000) and Lopes de Melo (2018) establish this in models where the idiosyncratic shock distribution is degenerate.

In contrast, equations (17) and (18) describe the expectation of the log wage conditional

on matching, i.e. conditional on $z \geq \bar{z}_{x,y}$. While no one has proposed a method for measuring $W_{x,y}(z)$, there are methods for recovering this reduced-form wage equation, notably those in Bonhomme, Lamadon and Manresa (2019). Propositions 2 and 3 illustrate that the reduced-form average wage (equations (17) and (18)) may look very different than the structural wage (equation (9)).

The source of this difference is selection, in particular the differential selection of wages across $(x, y)$ pairs. We may rewrite equation (9) as

$$W_{x,y}(z) = \bar{W}_{x,y} + \nu_{x,y}(z),$$

where $\bar{W}_{x,y} \equiv \bar{w}_x + \gamma f_{x,y}(z^* - \bar{z}_{x,y})$ and $\nu_{x,y}(z) \equiv \gamma f_{x,y}(z - z^*)$. By construction, the error term $\nu_{x,y}(z)$ has mean zero, and so $\bar{W}_{x,y}$ measures the average wage that a type $y$ firm would pay a type $x$ worker if it paid the wage given in eqaution (9) at any value of the match-specific shock. The main difference between $\bar{W}_{x,y}$ and $w^*_{x,y}$ then lies in the fact that the latter is selected, throwing away all observations where $z < \bar{z}_{x,y}$, or equivalently where $\nu_{x,y}(z) < \gamma f_{x,y}(\bar{z}_{x,y} - z^*)$.[4]

Now if all types had the same acceptance threshold $\bar{z}$, it would be as if we drew the match-specific shock from a truncated distribution. Selection would not change the shape of the average log wage. But importantly, our model features differential selection across $(x, y)$ pairs: $\bar{z}_{x,y}$ differs systematically with the worker's and firm's type. We argue in the next section that this differential selection is empirically reasonable, since it generates realistic assortative matching between workers and firms.

# 4 Assortative Matching

This section shows that our model can generate positively assortative matching between workers and firms. We define a strong notion of positively assortative matching: the matching density $\phi_{x,y}$ is strictly log-supermodular, meaning for all $x_1 < x_2$ and $y_1 < y_2$, $\phi_{x_1,y_1}\phi_{x_2,y_2} > \phi_{x_1,y_2}\phi_{x_2,y_1}$. An implication of this is that higher type workers and firms have a better distribution of match partners in the sense of first order stochastic dominance and that the correlation between the types of matched workers and firms is strictly positive (see Shimer, 2005, pp. 1013–1014).

[4] Another difference is that $w^*$ measures the average *log* wage. The proof of Propositions 2 and 3 shows how to obtain versions of equations (17) and (18) for the average wage and establishes that, under the same conditions as we use to prove $w^*$ is monotonic, the average wage is monotonic as well.

## 4.1 Assortative Matching in Productivity

We first find conditions for positively assortative matching for the Pareto case:

**Proposition 4** *Assume $S(z) = (z/z_0)^{-\theta}$ with $z_0 > 0$ and $\theta > 1$. Also assume $f_{x,y}$ is strictly increasing and weakly log-supermodular. Then the measure of matches $\phi_{x,y}$ is strictly log-supermodular.*

Weak log-supermodularity of $f$ is equivalent to $f_{x_1,y_1} f_{x_2,y_2} \geq f_{x_1,y_2} f_{x_2,y_1}$ for all $x_1 < x_2$ and $y_1 < y_2$.

The results for the exponential case impose somewhat stronger assumptions.

**Proposition 5** *Assume $S(z) = e^{-\theta z}$ with $\theta > 0$. Also assume $f_{x,y}$ is strictly increasing and either $-1/f_{x,y}$ is weakly supermodular or $f_{x,y}$ is multiplicatively separable. Then the measure of matches $\phi_{x,y}$ is strictly log-supermodular.*

For a strictly increasing function $f$, the assumption that $-1/f_{x,y}$ is weakly supermodular implies that $f_{x,y}$ is strictly log supermodular.

The combination of submodular average log wages (Propositions 2 and 3) and positively assortative matching (Propositions 4 and 5) may seem surprising. After all, if low productivity workers gain proportionately more from moving to high productivity firms, why do they work there less frequently? According to our model, the answer is selection. We use wage data as measured in a typical administrative data set, the wage paid by a firm to its employee. Such data sets do not have information about meetings that do not result in matches, i.e. about wage offers that are below the worker's reservation wage. Low wage workers rarely get acceptable wage offers from high productivity firms, and so rarely match there; but when such wage offers do materialize, our model predicts that the average log wage is higher than at a low productivity firm. We return to the role of selection in Section 8.

## 4.2 Assortative Matching in Wages

Define the average log wage paid to a type $x$ worker and the average log wage paid by a type $y$ firm:

$$\lambda_x \equiv \frac{\sum_{y=1}^{Y} w^*_{x,y} \phi_{x,y}}{\sum_{y=1}^{Y} \phi_{x,y}}, \tag{19}$$

$$\mu_y \equiv \frac{\sum_{x=1}^{X} w^*_{x,y} \phi_{x,y}}{\sum_{x=1}^{X} \phi_{x,y}} \tag{20}$$

The earlier results in this section give conditions for $\lambda_x$ and $\mu_y$ to be strictly increasing:

**Corollary 1** *1. Assume $S(z) = (z/z_0)^{-\theta}$ with $z_0 > 0$ and $\theta > 1$. Also assume $f_{x,y}$ is strictly increasing and weakly log-supermodular. Then $\lambda_x$ and $\mu_y$ are strictly increasing.*

*2. Assume $S(z) = e^{-\theta z}$ with $\theta > 0$. Also assume $f_{x,y}$ is strictly increasing and either $-1/f_{x,y}$ is weakly supermodular or $f_{x,y}$ is multiplicatively separable. Then $\lambda_x$ and $\mu_y$ are strictly increasing.*

The proof follows immediately from monotonicity of $w^*_{x,y}$ in both arguments (Propositions 2 and 3) and first order stochastic dominance of the match density $\phi$ (Propositions 4 and 5).[5]

This result implies that if $f$ is strictly increasing and satisfies the appropriate supermodularity condition, there is a one-to-one mapping between a worker's type $x$ their average log wage $\lambda_x$, and a one-to-one mapping between a firm's productivity $\pi_y$ and its average log wage $\mu_y$. Using the monotonicity of $\lambda$ and $\mu$, Propositions 4 and 5 then imply that the correlation between $\lambda_x$ and $\mu_y$ among matched workers and firms is strictly positive when $f$ is strictly increasing and satisifies the relevant supermodularity condition. Thus there is not only assortative matching of more productive workers to more productive firms, but also of high wage workers to high wage firms.

In Borovičková and Shimer (2020), we develop unbiased estimates of $\lambda$ for each worker and $\mu$ for each firm. We also show how to obtain consistent estimates of the variance of $\lambda_x$ across employed workers, the variance of $\mu_y$ across filled jobs, and the covariance between $\lambda_x$ and $\mu_y$ across matched workers and firms using a short panel with many workers and firms. Using administrative data from Austria, we verify that in fact the correlation between $\lambda$ and $\mu$ is strictly positive.

# 5 Numerical Example

In this section, we compare the numerical predictions of our model with data from Bonhomme, Lamadon and Manresa (2019). Using administrative data from Sweden, Bonhomme, Lamadon and Manresa (2019) estimate the average log wage $w^*_{x,y}$ in an $(x, y)$ match, as well as the share of such matches $\phi_{x,y}$. Their main findings are depicted in Figure 2 of their paper, which we summarize here. First, there is strong sorting between workers and firms, with low-type firms mostly employing low-type workers and high-type firms mostly employing high-type workers. Second, average log earnings are increasing in the worker type and firm type. Finally, the average log earnings of the lowest worker type is the most responsive to

[5]Monotonicity of $w^*_{x,y}$ and log-supermodularity of $\phi_{x,y}$ do not imply monotonicity of the Abowd, Kramarz and Margolis (1999) fixed effects $\bar{\alpha}_x$ and $\bar{\psi}_y$, defined in Appendix B and equations (46) and (47).

the firm type, consistent with submodular average log wages.[6]

We put numbers into our model to compare it to these findings. The purpose of this exercise is two-fold. First, we show that it is possible to match not only the qualitative wage and sorting patterns in the empirical literature, but also the quantitative behavior of the average log wage. Second, our efforts to match the quantitative behavior of the average log wage expose some limitations of the model. We address those in Section 6 when we introduce on-the-job search into the model.

We focus here on Pareto-distributed match quality and present results with exponentially-distributed match quality in Appendix C. We set the number of types at $X = Y = 10$ and assume equal numbers of each type, so $m_x = \frac{1}{X}$ and $n_y = \frac{1}{Y}$. The production function is CES with elasticity of substitution $\xi$,

$$f_{x,y} = \left( \tfrac{1}{2} p_x^{\frac{\xi-1}{\xi}} + \tfrac{1}{2} q_y^{\frac{\xi-1}{\xi}} \right)^{\frac{\xi}{\xi-1}},$$

where $p_x = (1+\Delta_p)^{x-1}$ and $q_y = (1+\Delta_q)^{y-1}$. We let $S(z) = (z/z_0)^{-\theta}$ and set the meeting rate $\rho$ so the average log wage in $(1,1)$ match is zero, $w^*_{1,1} = 0$.[7] We choose $\gamma = \frac{2}{3}$, $r = 0.05$, and $\delta = 0.5$, which implies expected duration of a match of 2 years. We choose the remaining parameters, $\Delta_p$, $\Delta_q$, $\xi$, and the Pareto tail parameter $\theta$ to match the variance decomposition in Bonhomme, Lamadon and Manresa (2019), as shown in Table 1.

Specifically we target the variance of worker's AKM fixed effects, the variance of firm's AKM fixed effects, the covariance between them, and the variance of the residual in an AKM regression. This gives us four moments to target with the four parameters.[8] In calculating the model-generated moments, we assume we have infinitely much data. This means we know each worker's type $x$ and each firm's type $y$. It also means that all distributions are deterministic functions of model parameters. Thus we abstract from important econometric

[6]Other papers present evidence consistent with submodular average log wages. Card, Heining and Kline (2013) regress log wages on worker and establishment (firm) fixed effects, as we discuss in Section B. They then sort workers and firms into deciles based on their estimated person and establishment effects and compute the average residual for each of the one hundred combinations of deciles. Figure VI in their paper shows that these residuals are on average positive when high types match with low types and negative when low types match with low types or high types match wit high types. This is consistent with submodular average log wages, so low worker-types increase their average log wage by more than the typical worker when they increase the firm type. On the other hand, Card, Cardoso and Kline (2016) show analogous calculations using Portuguese data in Figures B5 and B6 of their online appendix. They find that log wage residuals are positive for low-type workers working low-type firms and negative for low-type workers in high-type firms, which is consistent with average log wages being supermodular.

[7]We prove in Appendix D.2 that $\rho$ does not affect equilibrium allocations and has only a level effect on mean log wages in a more general version of this model, which allows for on-the-job search. We can apply the same normalizations to this simpler model.

[8]Since these targets are a non-linear function of the model parameters, there is no guarantee that we can match these moments. If we reduced workers' bargaining power to $\gamma = \frac{1}{2}$, we would not be able to do so.

| | $\text{var}(w)$ | $\text{var}(\alpha)$ | $\text{var}(\psi)$ | $2\text{cov}(\alpha,\psi)$ | $\text{var}(\varepsilon)$ | $\text{corr}(\alpha,\psi)$ |
|---|---|---|---|---|---|---|
| BLM | 0.124 | 0.074 | 0.003 | 0.015 | 0.031 | 0.491 |
| model | 0.124 | 0.074 | 0.003 | 0.015 | 0.031 | 0.491 |
| | | $\frac{\text{var}(\alpha)}{\text{var}(w)}$ | $\frac{\text{var}(\psi)}{\text{var}(w)}$ | $\frac{2\text{cov}(\alpha,\psi)}{\text{var}(w)}$ | $\frac{\text{var}(\varepsilon)}{\text{var}(w)}$ | |
| model | | 0.598 | 0.026 | 0.121 | 0.255 | |

Table 1: Decomposition of the variance of log wages into four components: variance of worker types, firm types, twice its covariance and the variance of the error term. The first row shows the decomposition reported in Bonhomme, Lamadon and Manresa (2019), the second row in our model. The third row shows what share of the variance is attributed to different components.

considerations that arise in real-world data sets where we observe each worker only for a few years and where we observe only a few workers at many firms.

The parameter values which allow us to match the target moments are $\Delta_p = 0.077$, $\Delta_q = 0.854$, $\xi = 0.713$ and $\theta = 7.146$. This implies that the most productive worker is on average 62 percent more productive than the least productive worker, and the most productive firm is on average 402 percent more productive than the least productive one [9]. The unemployment rate, which is primarily dictated by the parameter $\theta$, is 16.6 percent. We view both the high unemployment rate and the extreme cross-firm heterogeneity as limitations of this version of the model.

Still, the model can hit all four components of the variance decomposition, including the variance of error term. Our model provides a theory of it: the error term captures match-specific productivity which is not directly measurable. We note that the variance of the error term does not necessarily reflect the importance of match-specific shocks. In particular, the variance of the error term in our calibrated model is small, equal only to 0.03, yet the match-specific shocks play crucial role in generating average wage and sorting patterns.

Figure 1 is a direct analogue of Figure 2 in Bonhomme, Lamadon and Manresa (2019). The results are qualitatively very similar, though naturally results from our model are smoother than those from real-world data. The left panel of Figure 1 shows the average log wage as a function of firm type $y$. Each line corresponds to a different worker type $x$. As we proved in Proposition 2, the average log wage is increasing in $x$ and $y$ and submodular. Still, the lines are nearly parallel, which implies that the average log wage is almost

[9] This parametrization implies that worker heterogeneity, as measured by $f_{X,y}/f_{1,y} - 1$, lies between 0.37 and 0.82, depending on the value of $y$. Firm heterogeneity, as measured by $f_{x,Y}/f_{x,1} - 1$, lies between 3.35 and 4.78, depending on value of $x$.

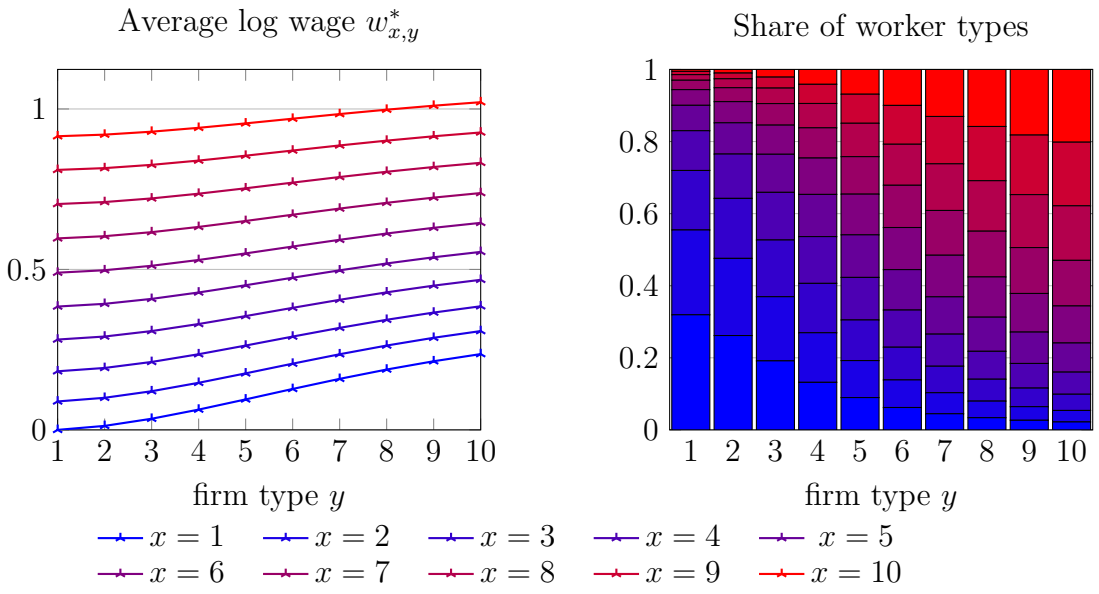



Figure 1: Average log wages and distribution of worker types conditional on firm type. The left panel shows the average log wage $w^*_{x,y}$ paid by different firms for different worker types. Each line represents one worker type. The right panel shows the distribution of worker types $x$ in firms with different firm types $y$.

additively separable in worker and firm types.

The right panel of Figure 1 shows the distribution of worker types in the different firm types, $\phi_{x,y}/\sum_{x'=1}^{X}\phi_{x',y}$. It is clear from the figure that high-type firms employ relatively more high-type workers than do low-type firms, as we proved in Proposition 4. That is, there is positive assortative matching.

# 6 On-the-Job Search

In this section, we extend our model to allow for on-the-job search, allowing employed workers to search for better jobs. This implies that vacant jobs meet both employed and unemployed workers, and that filled jobs sometimes end when the employee finds a better job opportunity.

We introduce on-the-job search for two reasons. First, the assumption is realistic. Employed workers often climb a job ladder, attaining higher wages as they move from job-to-job (Topel and Ward, 1992). It is important to see whether our findings on wages and assortative matching carry over to such an environment. Second, models with on-the-job search generate considerably more within-worker-type wage dispersion than models where only unemployed workers search (Hornstein, Krusell and Violante, 2011). We show how this affects

the variance decomposition of wages, in particular by reducing the amount of firm heterogeneity required to match the empirical variance of firm effects on wages. Additionally, by reducing the opportunity cost of taking a job, models with on-the-job search may potentially reduce the unemployment rate, which stood at 16.6 percent in our baseline model with only unemployed search.

## 6.1 Model with On the Job Search

The model is the same as the benchmark model but we now assume that workers can meet vacancies also when employed. An unemployed worker contacts a type-$y$ vacancy according to a Poisson process with arrival rate $\rho_0 v_y$, while an employed worker contacts a type-$y$ vacancy according to a Poisson process with arrival rate $\rho_1 v_y$, with $\rho_1 < \rho_0$. Following Bonhomme, Lamadon and Manresa (2019), we assume that a type-$x$ worker matched with a type-$y$ firm with a match-specific productivity shock $z$ earns the wage $W_{x,y}(z)$ which is such that the worker gets a fraction $\gamma$ of the joint match surplus. The wage does not depend on whether the worker has been hired from employment or unemployment, and is constant during the duration of the match, so there is no renegotiation. The rest of the model is the same as our benchmark model.

For a type-$x$ worker, the Bellman equation for the value of being unemployed is

$$rV_x^u = \rho_0 \sum_{y=1}^{Y} v_y \int_0^\infty \max\left\{V_{x,y}^e(z) - V_x^u, 0\right\} s(z) dz, \tag{21}$$

the same as equation (1). The value of a type $x$ worker employed at a type $y$ firm with a match-specific shock $z$ is

$$\begin{aligned} rV_{x,y}^e(z) = W_{x,y}(z) + \delta(V_x^u - V_{x,y}^e(z)) \\ + \rho_1 \sum_{y'=1}^{Y} v_{y'} \int_0^\infty \max\left\{V_{x,y'}^e(z') - V_{x,y}^e(z), 0\right\} s(z') dz'. \end{aligned} \tag{22}$$

The worker earns the wage $W_{x,y}(z)$ until the match ends. This happens either exogenously at rate $\delta$ or endogenously when the worker finds a better job. At the rate $\rho_1 v_{y'}$, the worker meets a type-$y'$ firm with a vacancy and draws a match-specific shock $z'$ according to distribution $s$. If the value of being employed at $y'$ with $z'$ exceeds the value of current employment, $V_{x,y'}^e(z') > V_{x,y}^e(z)$, the worker accepts the new job offer.

We next formulate Bellman equations for a firm. The Bellman equation for a vacant

type-$y$ job is

$$rV_y^v = \rho_0 \sum_{x=1}^{X} \int_0^\infty u_x \max\left\{V_{y,x}^f(z)) - V_y^v, 0\right\} s(z)dz + \rho_1 \sum_{x=1}^{X}\sum_{y'=1}^{Y} \int_0^\infty \phi_{x,y'}(z') \left( \int_0^\infty (V_{y,x}^f(z) - V_y^v) \mathbb{1}_{V_{x,y}^e(z) > V_{x,y'}^e(z')} s(z)dz \right) dz', \quad (23)$$

where $\phi_{x,y}(z)$ is the density of $(x, y, z)$ matches, and $\mathbb{1}$ is an indicator function. A vacant job meets an unemployed type-$x$ worker at the rate $\rho_0 u_x$ and decides to form a match if firm surplus is positive, $V_{y,x}^f(z) - V_y^v > 0$. At the rate $\rho_1 \phi_{x,y'}(z')$, the vacancy meets a type-$x$ worker employed in a type-$y'$ firm with match-specific shock $z'$. They draw a new match-specific shock $z$. If the worker prefers the offer from the poaching firm, i.e. $V_{x,y}^e(z) > V_{x,y'}^e(z')$, the worker accepts the new offer. Note that whenever the worker prefers the poaching offer, the match firm is willing to hire the worker.

Finally, we write the value of a type-$y$ firm employing a type-$x$ worker with match-specific shock $z$:

$$rV_{y,x}^f(z) = zf_{x,y} - W_{x,y}(z) + \left( \delta + \rho_1 \sum_{y'=1}^{Y} v_{y'} \int_0^\infty \mathbb{1}_{V_{x,y'}^e(z') > V_{x,y}^e(z)} s(z')dz' \right) \left( V_y^v - V_{y,x}^f(z)) \right). \quad (24)$$

The firm receives the profit flow $zf_{x,y} - W_{x,y}(z)$ during the duration of the match. The match can dissolve for one of two reasons: exogenously at the rate $\delta$ or endogenously due to a worker accepting an outside offer. This happens when the worker meets a type-$y'$ firm with a vacancy and draws a shock $z'$ such that the value of being employed there exceed worker's value of being in the current match, $V_{x,y'}^e(z') > V_{x,y}^e(z)$.

Define the match surplus $V_{x,y}^s(z) = V_{x,y}^e(z) + V_{y,x}^f(z) - V_x^u - V_y^v$. We assume that for all $(x, y, z)$, the worker keeps a share $\gamma$ of the match surplus, $V_{x,y}^e(z) - V_x^u = \gamma V_{x,y}^s(z)$, and the firm keeps the remainder, $V_{y,x}^f(z) - V_y^v = (1-\gamma)V_{x,y}^s(z)$. Using these as well as equations (21)–(24), we find the Bellman equation for the match surplus

$$(r+\delta)V_{x,y}^s(z) = zf_{x,y} - rV_x^u - rV_y^v + \rho_1 \sum_{y'=1}^{Y} v_{y'} \int_0^\infty \mathbb{1}_{V_{x,y'}^s(z') \geq V_{x,y}^s(z)} \left( \gamma V_{x,y'}^s(z') - V_{x,y}^s(z)) \right) s(z')dz'. \quad (25)$$

We can also rewrite the values of unemployed workers and vacant firms using equations (21)

and (23) as well as the surplus sharing rule:

$$rV_x^u = \gamma\rho_0 \sum_{y=1}^{Y} v_y \int_0^\infty \max\{V_{x,y}^s(z), 0\} s(z) dz, \tag{26}$$

$$\begin{aligned} rV_y^v = (1-\gamma)\rho_0 \sum_{x=1}^{X} \int_0^\infty u_x \max\{V_{x,y}^s(z)), 0\} s(z) dz \\ + (1-\gamma)\rho_1 \sum_{x=1}^{X}\sum_{y'=1}^{Y} \int_0^\infty \phi_{x,y'}(z') \left( \int_0^\infty V_{x,y}^s(z) \mathbb{1}_{V_{x,y}^s(z) > V_{x,y'}^s(z')} s(z) dz \right) dz'. \end{aligned} \tag{27}$$

Note that the last equation uses the fact that $V_{x,y}^e(z) > V_{x,y'}^e(z')$ if and only if $V_{x,y}^s(z) > V_{x,y'}^s(z')$.

To close the model, we characterize the distribution of unemployed workers $u_x$, vacant jobs $v_y$ and the density of matches $\phi_{x,y}(z)$. The steady-state measure of $(x, y, z)$ matches satisfies

$$\begin{aligned} \phi_{x,y}(z) \left( \delta + \rho_1 \sum_{y'} v_{y'} \int_0^\infty \mathbb{1}_{V_{x,y'}^s(z') \geq V_{x,y}^s(z)} s(z') dz' \right) \\ = v_y s(z) \left( \rho_0 u_x \mathbb{1}_{V_{x,y}^s(z) \geq 0} + \rho_1 \sum_{y'=1}^{Y} \int_0^\infty \mathbb{1}_{V_{x,y}^s(z) \geq V_{x,y'}^s(z')} \phi_{x,y'}(z') dz' \right). \end{aligned} \tag{28}$$

The left hand side is the measure of matches that are destroyed through exogenous separations and through endogenous on-the-job search. The right hand side counts matches which are created through a vacancy meeting an unemployed worker (first term) or an employed worker who prefers the new offer.

The measures of unemployed and vacant jobs satisfy

$$u_x = m_x - \sum_{y=1}^{Y} \int_0^\infty \phi_{x,y}(z) dz \tag{29}$$

$$v_y = n_y - \sum_{x=1}^{X} \int_0^\infty \phi_{x,y}(z) dz. \tag{30}$$

A steady state equilibrium is values $V_x^u$, $V_y^v$, and $V_{x,y}^s$, as well as matched and unmatched rates $\phi_{x,y}(z)$, $u_x$, and $v_y$ solving equations (25)–(30). In Appendix D, we describe how we solve the model.

Finally, using the surplus sharing rule and equation (22), we find an expression for the

wage

$$W_{x,y}(z) = rV_x^u + \gamma(r+\delta)V_{x,y}^s(z) - \gamma\rho_1 \sum_{y'=1}^{Y} v_{y'} \int_0^\infty \max\left\{V_{x,y'}^s(z') - V_{x,y}^s(z), 0\right\} s(z')dz'. \quad (31)$$

The wage equals worker's flow value of unemployment plus the $\gamma$-share of the flow match surplus, adjusted for the option value of on-the-job search.

## 6.2 Results

We now put numbers into the model, following the same empirical strategy as in Section 5. We focus throughout on the case of a Pareto-distributed match quality, $S(z) \propto z^{-\theta}$. We also assume a CES production technology

$$f_{x,y} = \left(\tfrac{1}{2}\left((1+\Delta_p)^{x-1}\right)^{\frac{\xi-1}{\xi}} + \tfrac{1}{2}\left((1+\Delta_q)^{x-1}\right)^{\frac{\xi-1}{\xi}}\right)^{\frac{\xi}{\xi-1}}.$$

There are ten types of workers and firms and equal numbers of each type. As before, we set $\gamma = 2/3$, $r = 0.05$, and $\delta = 0.5$.

Only the ratio $\rho_1/\rho_0$ affects the equilibrium allocation or relative wages, as we prove in Appendix D.2. We set $\rho_1 = 0.2\rho_0$, in line with relative intensities reported in the literature, and again set $\rho_0$ so as to normalize $w_{1,1}^* = 0$. Using French administrative data, Postel-Vinay and Robin (2002) estimate this ratio to be between 0.31 and 0.48, depending on worker occupation and skill category. Jolivet, Postel-Vinay and Robin (2006) estimate this ratio to be 0.03 and 0.19 for different European countries.

We again choose the four remaining parameters, $\Delta_p$, $\Delta_q$, $\theta$, and $\xi$, to match the variance decomposition of wages in Bonhomme, Lamadon and Manresa (2019), as shown in Table 1. In our calculations, we again assume we have infinitely much data, so we know each worker's type $x$ and each firm's type $y$ and so all distributions are deterministic functions of model parameters. But we note that the on-the-job search model satisfies the identification assumptions of the static Bonhomme, Lamadon and Manresa (2019) model, and hence one can apply their estimator to the model generated data to estimate average wages and sorting patterns. Since their estimator is consistent, the two procedures should yield similar results, with differences arising only due to finite sample in the model generated data.

The parameter values which allow us to match the target moments are $\Delta_p = 0.072$, $\Delta_q = 0.217$, $\xi = 0.688$ and $\theta = 16.261$. In this calibration, the most productive worker is on average 42 percent more productive than the least productive worker, and the most

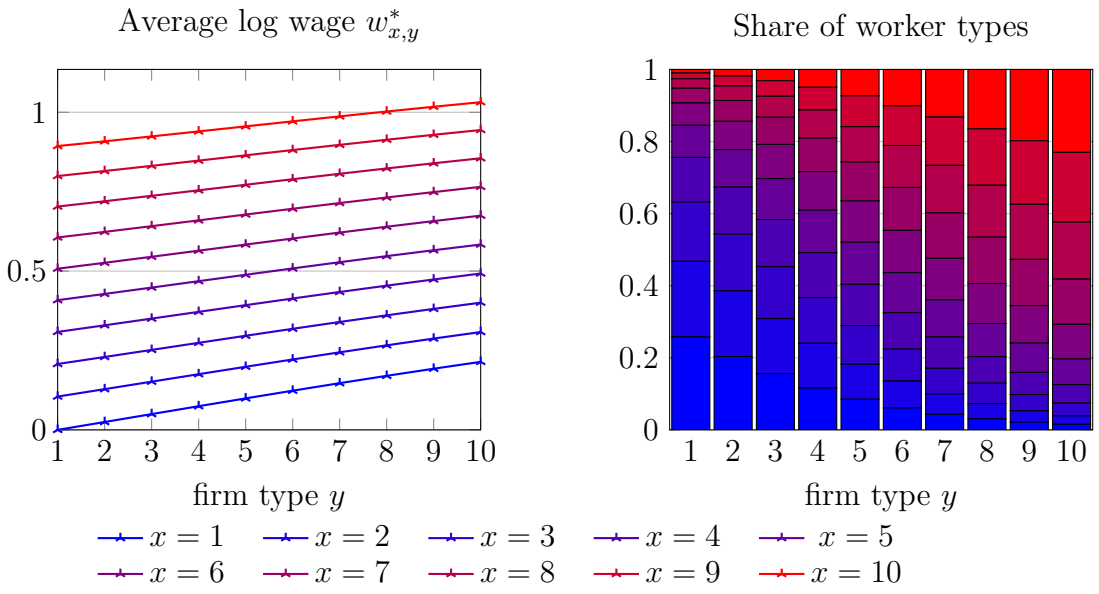



Figure 2: Average log wages and distribution of worker types conditional on firm type in the model with on the job . The left panel shows the average log wage $w^*_{x,y}$ paid by different firms for different worker types. Each line represents one worker type. The right panel shows the distribution of worker types $x$ in firms with different firm types $y$.

productive firm is on average 116 percent more productive than the least productive one.[10] Notably, with on-the-job search we have much less firm heterogeneity (lower $\Delta_q$) than we needed without on-the-job search, and also a much lower variance of the idiosyncratic shock (higher $\theta$). This reflects the fact that we can more easily generate within-worker-type wage dispersion in on-the-job search model (Hornstein, Krusell and Violante, 2011). We also find a bit less worker heterogeneity, while the elasticity of substitution in production scarcely changes.

These parameter values imply a much lower unemployment rate, 4.5 percent rather than 16.6 percent without on-the-job search. This is a consequence of two things. First, the opportunity cost of taking a job is lower when a worker can continue to search while employed, encouraging workers to accept marginal matches. Second, the calibration has much less firm heterogeneity and much less variation in the match-specific shock. Both of these reduce the option value of search as well.

Figure 2 replicates Figure 1 for the model with on-the-job search, and is again a direct analogue of Figure 2 in Bonhomme, Lamadon and Manresa (2019). The results are very

[10]This parametrization implies that worker heterogeneity, as measured by $f_{X,y}/f_{1,y} - 1$, lies between 0.34 and 0.51, depending on the value of $y$. Firm heterogeneity, as measured by $f_{x,Y}/f_{x,1} - 1$, lies between 1.04 and 1.30, depending on value of $x$.

similar. In particular, we proved without on-the-job search that the average log wage is increasing and submodular (Proposition 2). It is easy to see from the left panel that wages are still monotonic in both worker and firm type with on-the-job search. The lines are nearly parallel, implying that the average log wage is almost additively separable in worker and firm types, as in the Abowd, Kramarz and Margolis (1999) wage specification.

The right panel of Figure 2 shows the distribution of worker types in the different firm types, $\phi_{x,y} / \sum_{x'=1}^{X} \phi_{x',y}$. It is clear from the figure that high-type firms employ relatively more high-type workers than do low-type firms. That is, there is positive assortative matching. This is again consistent with the theoretical prediction without on-the-job search (Proposition 4).

Figure 7 in the Appendix shows results for a lower value of employed search intensity, $\rho_1 = 0.1\rho_0$. We recalibrate the four parameters to match the same variance decomposition. We find the same qualitative patterns for average log wages and sorting. As $\rho_1$ increases, the average log wage becomes less submodular and more additively separable in the worker and firm types. There is still strong positive sorting.

# 7 Testing Exogenous Mobility

The literature estimating the importance of workers and firms in wage determination is naturally concerned about whether the error in the wage equation has mean zero for all workers and firms. Following Abowd, Kramarz and Margolis (1999), the literature proposes regressing the log wage $w_{i,t}$ of worker $i$ employed by firm $j$ at time $t$ on fixed effects $\alpha_i$ for worker $i$, fixed effects $\psi_j$ for the employer $j$, and an error term, as in equation (16). This literature interprets $\alpha_i$ to be the average effect that worker $i$ would have on their wage at any firm, not just the firms that employ them. It seems natural to interpret this as the value of the worker's human capital. Similarly, $\psi_j$ is the average effect that firm $j$ would have on the wage of any worker, not just the workers that it employs. Card, Heining and Kline (2013) state that "such a premium could represent rent-sharing, an efficiency wage premium, or strategic wage posting behavior [p. 987]." In our model, firms matter for wages because of rent-sharing.

To estimate equation (16), one makes the usual assumption that the mean of the error term $\varepsilon_{i,t}$ is independent of the regressors $\alpha_i$ and $\psi_j$. Card, Heining and Kline (2013), Abowd, McKinney and Schmutte (2019), and others call this assumption "exogenous mobility." As we discussed in Section 3.4, exogenous mobility is violated for the structural wage equation (9) when we only have data on matches that are created. This is because the workers and firms that match are not a random subset of all workers and firms, but rather those

with a high value of the idiosyncratic shock $z$.

## 7.1 The Difficulty of Testing

As is well-known, the exogenous mobility assumption cannot be tested using regression approach and administrative wage data. One way to see this follows an insight from Flinn and Heckman (1982): we construct an alternative model in which mobility is clearly random yet which generates the same administrative wage data. For expositional simplicity, we assume that there is no on-the-job search.

In our alternative model, a type-$x$ worker meets a type-$y$ firm at rate $\rho S(\bar{z}_{x,y})v_y$ (rather than $\rho v_y$). In that event, they draw a match-specific productivity shock with survival function $S(z)/S(\bar{z}_{x,y})$ for $z \geq \bar{z}_{x,y}$ (rather than $S(z)$ for $z \geq z_0$), and the wage is given by $W_{x,y}(z)$ defined in equation (9). We call this the "exogenous sorting model."

By construction, all meetings result in matches in the exogenous sorting model and so there is no selection. It is fair to say that mobility is exogenous. Nevertheless, all observable outcomes in standard administrative data sets—the joint distribution of wages and match partners—are identical to our model of selection and endogenous sorting. We conclude that it is impossible to test whether there is selection once we allow for a bigger set of possible data generating processes.

To circumvent this, the literature has proposed proxy tests for exogenous mobility. Here we discuss how our model would fare with some of the better known tests for exogenous mobility. We note that in all cases, the exogenous sorting model would fare equally well, since the two models generate the same administrative data sets.

## 7.2 Event Studies

One well-known test, first introduced by Card, Heining and Kline (2013), looks at whether the wage gains of workers who move from one establishment to another are equal in magnitude and opposite in sign to the wage gains of workers who move in the opposite direction. They argue that this would not be the case in many models of on-the-job search, such as Burdett and Mortensen (1998), where workers always experience wage increases when switching jobs. Card, Heining and Kline (2013) and subsequent research by these authors find remarkable support for the "equal in magnitude and opposite in sign" prediction.

In implementing this test, Card, Heining and Kline (2013) look at workers who switch their main job between consecutive years. Some of these moves may be direct, without an intervening spell of unemployment, but some are indirect, with an unemployment spell between the jobs. To construct a similar event study in our model, we look at the typical

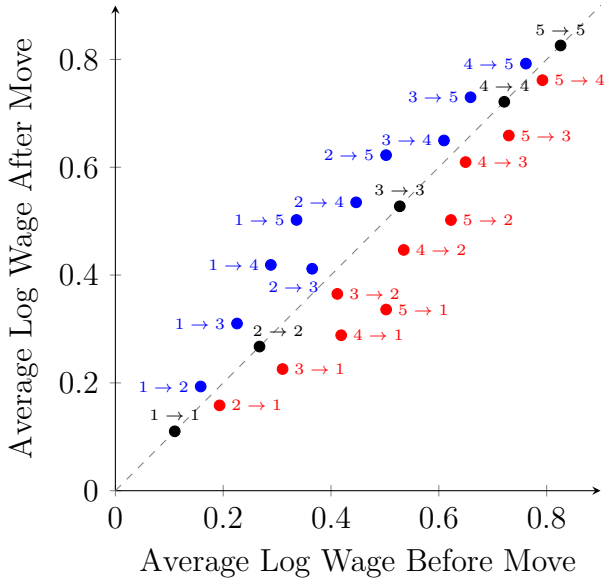



Figure 3: Average log wages before and after move as a function of the firm's type quintile. Calibrated model without on-the-job search. The dot labelled "$q_1 \to q_2$" shows the average log wage before and after the move for a worker going from a quintile $q_1$ to a quintile $q_2$ job.

experience of a worker who switches jobs, either because of an exogenous separation shock or because of a job-to-job movement. We then compare the worker's wage before and after the move. Finally, we group firms into five quintiles based on their type $y$ and calculate the average log wage of a worker moving from quintile $q_1$ to quintile $q_2$ both before and after the move. Appendix E provides more details on these calculations.

Figure 3 shows the results from our calibrated model without on-the-job search (Section 5). We comment on a few features of the figure and compare them to the data. First, a worker going from quintile $q_1$ to quintile $q_2$ earns more in both jobs if $q_1$ or $q_2$ is bigger. Card, Heining and Kline (2013) observe this in the data, writing on page 985 that "different mobility groups have different wage levels before and after a move."

The same forces that lead to a firm effect in wages in Section 3 explain the impact of $q_1$ on the first wage and $q_2$ on the second wage. To understand the impact of $q_2$ on the first wage or $q_1$ on the second wage, we note that wages are not conditioned on the worker's type in this event study. In the model, the worker's wage depends on their type and their employer's type, but not on future or past employers. However, future and past employers' types are correlated with the workers' type because of assortative matching, and thus help to predict the workers' wage.

Second, our model without on-the-job search features exact symmetry of the wages of a

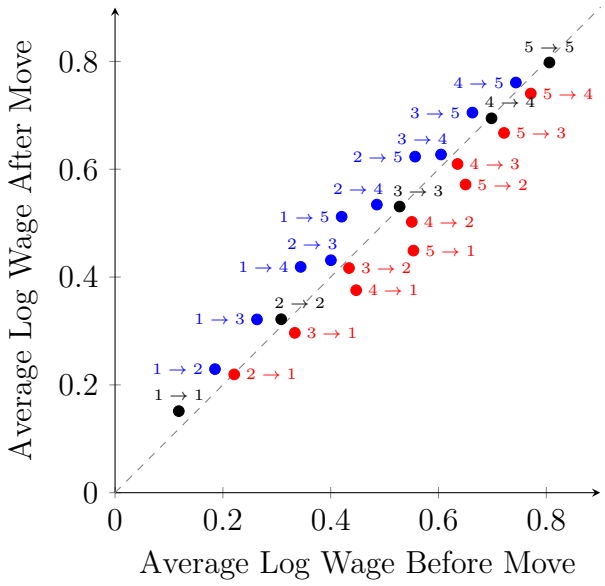



Figure 4: Average log wages before and after move as a function of the firm's type quintile. Calibrated model with on-the-job search. The dot labelled "$q_1 \to q_2$" shows the average log wage before and after the move for a worker going from a quintile $q_1$ to a quintile $q_2$ job.

worker going from quintile $q_1$ to quintile $q_2$ compared to a worker going from $q_2$ to $q_1$. Card, Heining and Kline (2013) argue that "This symmetry suggests that a simple model with additive worker and establishment effects may provide a reasonable characterization of the mean wages resulting from different pairings of workers to establishments (p. 985)." In our model, the symmetry reflects a more fundamental force: the distribution of worker types who move from $q_1$ to $q_2$ is identical to the distribution who move from $q_2$ to $q_1$, because each worker has the same probability of working first at $q_1$ and then at $q_2$ or of having the same two jobs in the reverse order. Moreover, the average log wage at each firm type depends on the worker and firm types for the reasons we discussed in Section 3.

In Figure 4, we show the corresponding figure for our model with on-the-job search. Now some moves are voluntary, in search of a higher wage, and some are involuntary, due to a separation shock. Most of the patterns are unchanged from Figure 3. First, a worker moving from a quintile $q_1$ to a quintile $q_2$ firm earns more in both jobs when either $q_1$ or $q_2$ is higher. This again reflects two factors: both worker and firm types predict the average log wage; and there is assortative matching, so past and future firm types predict the worker type.

With on-the-job search, the exact symmetry of upward and downward moves is broken. For example, a worker moving from the first to fifth quintile has an average log wage gain of 9.2 log points, while a worker moving in the opposite direction loses 10.4 log points. Still, the change in wage always has the "correct" sign—positive if $q_2 > q_1$ and negative if

$q_2 < q_1$—and generally a similar magnitude. The most asymmetric pairing is for movements between the bottom two quintiles, where an upward move gives a 4.4 log point gain, while a downward move gives a 0.2 log point loss.

There is one notable difference between these figures and the ones in Card, Heining and Kline (2013): the magnitude of the wage changes is much smaller here. We believe this is a consequence of our calibration, based on evidence in Bonhomme, Lamadon and Manresa (2019), which attributes little of the overall wage variation to firm effects.

## 7.3 Other Tests

In a separate test, Card, Heining and Kline (2013) look at the variance of the wage residual in equation (16). They write that a large "idiosyncratic match component of wages changes the interpretation of the estimated establishment effects, since different workers have different wage premiums at any given establishment, depending on the value of their match component (Card, Heining and Kline, 2013, p. 989)." In their empirical work, they find that the standard deviation of the match-specific wage residual is between 0.060 and 0.075, and argue that this is small enough to constitute evidence against endogenous mobility. According to our model, the wage residual in equation (16) depends on the variance of the idiosyncratic shock $z$ as well as any nonseparabilities in the average log wage. Importantly, we calibrated our model to match this wage residual. In Table 1, we showed that our model can simultaneously generate a realistic relationship between the average log wage and worker and firm types, a realistic amount of sorting between workers and firms, and a realistically wage residual. The model with on-the-job search hits the same targets. Thus we do not think that the size of the wage residual can rule out the possibility of endogenous mobility.

Finally, Abowd, McKinney and Schmutte (2019) look at whether the wage residual $\hat{\varepsilon}_{i,t}$ in a worker's current match forecasts a future employer's fixed effect $\hat{\psi}_{j'}$, where $j'$ is $i$'s employer at some future date $t' > t$, conditional on the worker's and current employer's fixed effects $\hat{\alpha}_i$ and $\hat{\psi}_j$. The null hypothesis, exogenous mobility, is that the estimated wage residual is not useful for forecasting. In our model without on-the-job search, the worker's type $x$ determines the distribution over future employers' fixed effects $\bar{\psi}_y$ and vice versa, and so we would not reject the null hypothesis. With on-the-job search, the wage residual may be useful for forecasting which type of job-to-job transitions are more likely, even conditional on the worker type $x$, and so we may reject the null hypothesis. The paper also explores this from the employer's perspective. In this case, even the on-the-job search model predicts that only the employer's type, not the current wage residual, predicts the type of future employees, so we should not reject the null hypothesis. In fact, Abowd, McKinney and

Schmutte (2019) reject both null hypotheses. Exploring the quantitative magnitude of these rejections remains a topic for future research.

# 8 The Role of Selection

## 8.1 The Value of Meetings and the Value of Matches

Bonhomme, Lamadon and Manresa (2019) view the combination of positive assortative matching and submodular average low wages as puzzling, writing "the results in this section are difficult to reconcile with models based on revealed preferences for wages only [p. 720]." Low-type workers have the most to gain from working in high-type firms yet they are predominantly employed by low-type firms. They propose several factors besides the average log wage that may drive sorting behavior. For example, they recognize that workers may care about other job attributes which differ across employers, such as amenities. For these to drive sorting patterns, they propose that high-wage workers value relatively more the amenities offered by high-wage firms, so these unobserved amenities drive sorting patterns. Alternatively, the propose that workers like to work with similar peers, and again this drives the empirical sorting patterns. Thus in their view, wage data give a misleading picture of the value that workers get from matching with different types of employers.

Positive assortative matching and submodular average low wages arise in our model as well. According to the numbers behind Figure 1, 26 percent of the lowest firm decile's workers are drawn from the lowest worker decile, compared to less than 1 percent for the highest firm decile. For the highest worker decile, these numbers are nearly flipped. They represent 23 percent of employment at the highest firm decile and less than 2 percent at the lowest firm decile. This is positively assortative matching. At the same time, the lowest worker decile has an average log wage of 0 in the lowest firm decile and 0.21 in the highest firm decile. The corresponding gain for the highest worker decile is slightly smaller, going from 0.89 to 1.03, consistent with submodular average log wages.

Why is there positively assortative matching when average log wages are submodular? According to our model, wages describe the value that workers derive from matching with different types of employers. However, they do not capture the value that workers obtain from merely *meeting* different types of employers. The value of a meeting is captured by the match surplus $V^s_{x,y}(z)$ in equation (5) or (25). For illustrative purposes, we set the lower bound on the idiosyncratic shock distribution, $z_0$, equal to the minimum acceptance

threshold across $(x, y)$ pairs. We then define the average surplus in meetings as

$$\bar{V}^{s,m}_{x,y} \equiv \int_{z_0}^{\infty} V^s_{x,y}(z)s(z)dz. \tag{32}$$

We also define the average surplus conditional on matching,

$$\bar{V}^{s}_{x,y} \equiv \frac{1}{S(\bar{z}_{x,y})} \int_{\bar{z}_{x,y}}^{\infty} V^s_{x,y}(z)s(z)dz. \tag{33}$$

In Figure 5, we plot the average surplus in an $(x, y)$ match, $\bar{V}^s_{x,y}$, in the left panel, and the average surplus in an $(x, y)$ meeting, $\bar{V}^{s,m}_{x,y}$, in the right panel. The top row shows the model without on-the-job search, while the bottom row shows the model with $\rho_1 = 0.2\rho_0$.

The left panels reflect the results in Figures 1 and 2. The average surplus in matches is increasing in firm type for each worker type and is log submodular. Looking at this panel, one might be tempted to conclude that low-type workers benefit at least as much from meeting a high-type firm than do high-type workers. The right panels of the same figure, where we show average surplus in meetings, contradicts this conclusion. We observe that in the model with on-the-job search, the lowest-type worker has the highest average surplus in a meeting with a type-1 firm and gets almost 83 percent less surplus from meeting a type-10 firm.

The difference between the left and right figures is selection: while the left panel shows the average surplus in meetings that result in matches, the right panel shows the average surplus in all meetings, keeping the support of the $z$ distribution the same for every pair. In other words, the right panel is not affected by selection of worker types into firms based on the realized value of $z$. Based on the right hand panel of Figure 5, it should no longer be surprising that low-type workers rarely take high-type jobs, as illustrated in the right panel of Figures 1 and 2. Average surplus in meetings, not in observed matches, drive sorting.

## 8.2 Identification of Meeting Surplus with Random Search

In this section, we prove that it is possible to identify the extent of log-supermodularity of the average surplus in meetings using the structure of our model and the type of information in standard administrative data sets. More precisely, we extend our baseline model without on-the-job search in several directions, relaxing the assumption that wages are determined by Nash bargaining and allowing for a more general matching technology. Importantly, we maintain the assumption that search is random.

In our more general model, an unemployed type $x$ worker meets a type $y$ firm at rate $\rho^u_x \rho^v_y$, a multiplicatively separable function of the worker and firm types. This allows for the

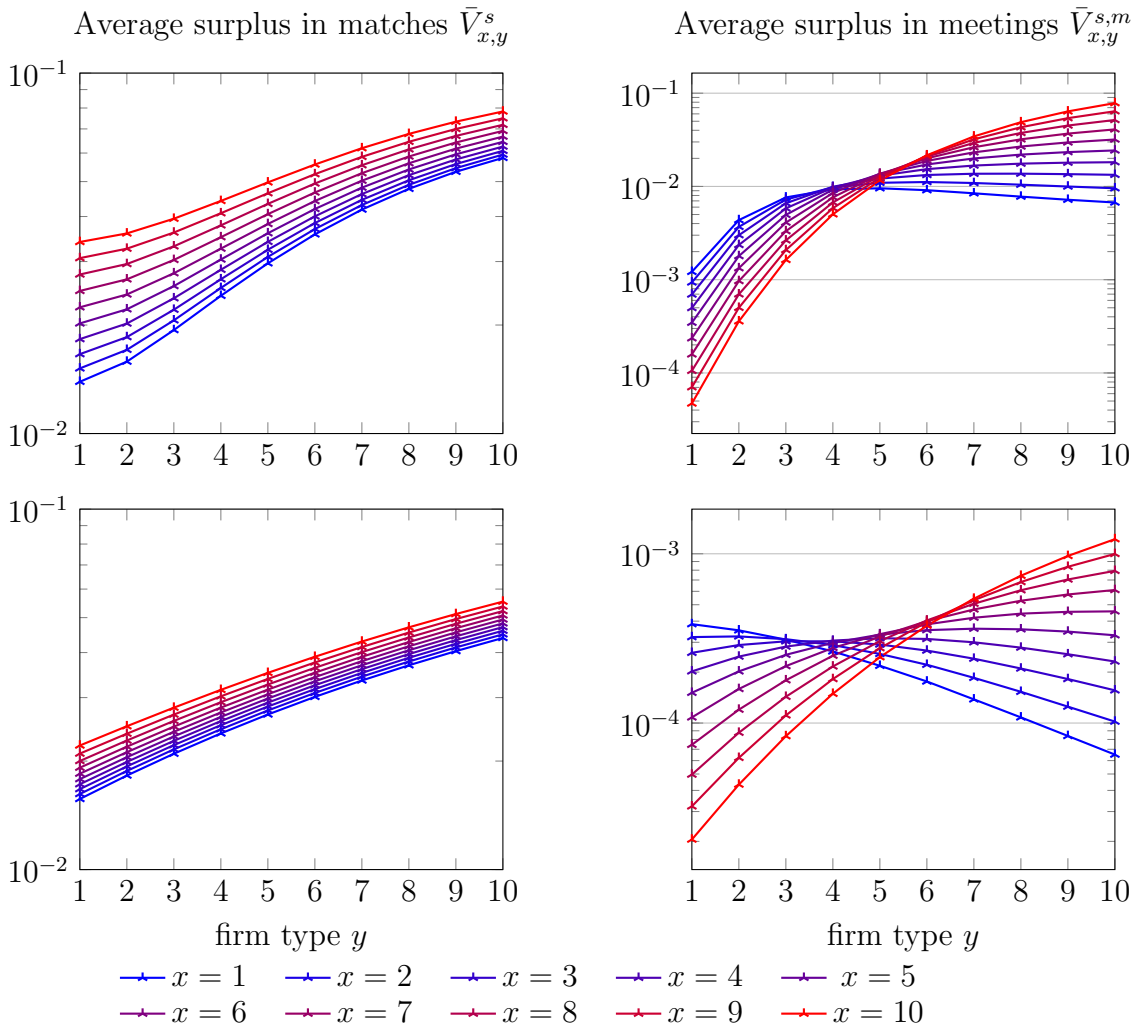



Figure 5: Average surplus in matches (left) and meetings when $z_0$ is equal to the minimum threshold across $(x, y)$ pairs. Top row shows the model without on-the-job search. Bottom row shows the model with on-the-job search. Each line represents one worker type and each point one firm type. Vertical scales are logarithmic.

possibility that some workers and firms are better at searching, but maintains the assumption that search is random rather than directed. Upon meeting, the worker samples from a wage distribution with survival function $\tilde{S}_{x,y}(w)$ and density $\tilde{s}_{x,y}(w)$, and chooses whether to accept or reject the job. This may be the reduced form coming from the Nash wage equation (9), the production technology $f$, and the distribution of match-specific shocks, but we do not need this interpretation for our results in this section. Finally, the job ends at rate $\delta_x$ and the worker discounts the future at rate $r$.

Under these assumptions, we have the value functions for type $x$ workers both when they are unemployed and when they are employed at a wage $w$:

$$rV_x^u = \rho_x^u \sum_y \rho_y^v \int \max\{V_x^e(w) - V_x^u, 0\} \tilde{s}_{x,y}(w) dw, \tag{34}$$

$$rV_x^e(w) = w + \delta_x(V_x^u - V_x^e(w)). \tag{35}$$

Note that the value of an employed worker depends on the wage but not the firm type. Moreover, from equation (35), we get that the worker accepts any job with $w > \bar{w}_x = rV_x^u$, the worker's reservation wage. Additionally, we use the fact that in steady state, the measure of $(x, y)$ matches solves

$$\delta_x \phi_{x,y} = \rho_x^u \rho_y^v u_x \tilde{S}_{x,y}(\bar{w}_x), \tag{36}$$

with $u_x = m_x - \sum_y \phi_{x,y}$. Finally, define the average wage (not log wage) in an $(x, y)$ match as

$$W_{x,y}^* \equiv \frac{\int_{\bar{w}_x}^{\infty} w \tilde{s}_{x,y}(w) dw}{\tilde{S}_{x,y}(\bar{w}_x)}. \tag{37}$$

Combining equations (34)–(37), we obtain

$$\bar{w}_x = \frac{\delta_x}{ru_x + \delta_x m_x} \sum_y W_{x,y}^* \phi_{x,y}. \tag{38}$$

Note that if $r = 0$, this is simply the average wage multiplied by the employment rate, while with discounting it is a bit lower.[11] Using the methodology in Bonhomme, Lamadon and Manresa (2019), the average wage $W_{x,y}^*$ and the share of type-$x$ workers matched to type-$y$ firms, $\phi_{x,y}/m_x$, are identified. Moreover, a slight extension to the methodology allows us to recover how many type-$x$ workers separate each period, $\delta_x m_x$, as well as the type-specific unemployment rate $u_x$ (Lamadon, Lise, Meghir and Robin, 2024). Thus the reservation wage is identified if we use other information to recover the discount rate $r$.

[11] In practice, the distinction is small because $r$ and $u_x$ are each an order of magnitude smaller than $\delta_x$ and $m_x$, respective.

The average surplus from an $(x, y)$ meeting, defined in equation (32), is

$$\bar{V}_{x,y}^{s,m} = \int \max\{V_x^e(w) - V_x^u, 0\}\tilde{s}_{x,y}(w)dw. \tag{39}$$

We simplify this using equations (35), (36), and (37) to obtain

$$\bar{V}_{x,y}^{s,m} = \frac{(W_{x,y}^* - \bar{w}_x)\delta_x\phi_{x,y}}{(r+\delta_x)\rho_x^u\rho_y^v u_x}. \tag{40}$$

Note that the average surplus from meetings, $\bar{V}_{x,y}^{s,m}$, is log-supermodular if and only if $(W_{x,y}^* - \bar{w}_x)\phi_{x,y}$ is log-supermodular. Each element of this expression is identified using the methodology in Bonhomme, Lamadon and Manresa (2019), and so we can identify the extent of log-supermodularity in the gains meetings without observing meetings that do not result in matches.

# 9 Conclusion

We have developed a random search model of the labor market with ex ante heterogeneous workers and firms and ex post match-specific productivity shocks. When the distribution of match-specific shocks is Pareto or exponential, we obtain simple proofs of three results. First, we show that there is positively assortative matching when the production function is sufficiently complementary. Second, we show that on average, more productive workers earn higher wages at any type of firm and more productive firms pay higher wages to any type of worker. And finally, we show that more productive workers earn more on average and more productive firms pay more on average, so high productivity workers and firms are high wage workers and firms.

We also argue that our model is consistent with a variety of facts in the empirical labor economics literature built around the Abowd, Kramarz and Margolis (1999) wage equation and the non-additive specification in Bonhomme, Lamadon and Manresa (2019). In particular, we are able to replicate the decomposition of the cross-sectional variance in log wages into the component due to ex ante worker heterogeneity, the component due to ex ante firm heterogeneity, the component due to the sorting of workers and firms, and a residual due to the match-specific productivity shocks.

Additionally, our model predicts that average log wages may be submodular—low-wage workers gain proportionately more than high-wage workers when they move from low-wage to high-wage firms—while simultaneously low-wage workers may be disproportionately employed at low-wage firms. This is an empirically plausible pattern. In our model, this reflects

a selection mechanism. Low-wage workers rarely accept high-wage jobs, but when they do, productivity is so high that their wage is also very high.

Our model has implications for whether the empirical literature stemming from the Abowd, Kramarz and Margolis (1999) wage equation can tell us something about how wages are determined. To continue the previous paragraph, take the empirical finding that average log wages are submodular. This result, viewed through a lens that ignores selection issues, implies that low-wage workers would particularly benefit from meeting high-wage firms. That they don't often match with such firms is then a puzzle. Bonhomme, Lamadon and Manresa (2019) write, "The coexistence of strong sorting and weak complementarities may be surprising in the perspective of the search matching models inspired by Becker (1973).... This may indicate that workers or firms care about other job attributes (Hwang, Mortensen and Reed, 1998), or that workers of a similar type enjoy working together as in the peer effects literature [p. 720]." Lamadon, Lise, Meghir and Robin (2024) resolve this through an assumption that low-wage workers enjoy substantially higher amenities when employed by low-wage firms. Our model offers an alternative mechanism, that observed wages are a selected subset of all (latent) wage offers, and that selection shapes the average log wage. A corollary of this is that a regression of log wages on worker and firm fixed effects need not shed light on firms' role in wage determination because of sample selection issues.

In a similar vein, recent papers by Engbom, Moser and Sauermann (2023) and Lachowska, Mas, Saggio and Woodbury (2023) estimate the AKM regression with time-varying firm types, $\psi_{j,t}$, and interpret changes in $\psi_{j,t}$ as changes in firm pay policies. Our model implies that time variation in firm fixed effects can capture other things, such as changes in selection due to time-varying variation in the distribution of match-specific shocks. Hence time-variation in estimated firm fixed effects again may not have a structural interpretation as a change in how firms set wages.

Our analysis also cautions against attempts to learn about the production function using only administrative data on average log wages. Bonhomme, Lamadon and Manresa (2019) write "As a first way to quantify the economic magnitude of complementarities, we next assess the explanatory power of worker types and firm classes when those enter the regression interactively as opposed to additively. The $R^2$ coefficient in the linear regression is 74.8%, while in the regression that includes all interactions between worker type indicators and firm class indicators, the $R^2$ is 75.8%. Hence, while the left panel of Figure 2 suggests the presence of some complementarity between firms and lower-type workers, those complementarities explain only a small part of the variance of log-earnings [pp. 718–719]." Proposition 2 shows that when the distribution of idiosyncratic shocks is Pareto, the production function does not directly enter the expression for the average log wage $w^*_{x,y}$, and thus this conclusion may

be unwarranted in our model.

# References


**Abowd, John M., Francis Kramarz, and David N. Margolis**, "High Wage Workers and High Wage Firms," *Econometrica*, 1999, *67* (2), 251–333.

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

**Abowd, John M., Robert H. Creecy, and Francis Kramarz**, "Computing person and firm effects using linked longitudinal employer-employee data," 2002. Center for Economic Studies, US Census Bureau.

**Andrews, Martyn J., Leonard Gill, Thorsten Schank, and Richard Upward**, "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias?," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2008, *171* (3), 673–697.

**Bagger, Jesper and Rasmus Lentz**, "An Empirical Model of Wage Dispersion with Sorting," *Review of Economic Studies*, 2019, *86* (1), 153–190.

**Becker, Gary S.**, "A Theory of Marriage: Part I," *Journal of Political Economy*, 1973, *81* (4), 813–846.

**Bonhomme, Stéphane, Kerstin Holzheu, Thibaut Lamadon, Elena Manresa, Magne Mogstad, and Bradley Setzler**, "How much should we trust estimates of firm effects and worker sorting?," *Journal of Labor Economics*, 2023, *41* (2), 291–322.

**Bonhomme, Stéphane, Thibaut Lamadon, and Elena Manresa**, "A Distributional Framework for Matched Employer Employee Data," *Econometrica*, 2019, *87* (3), 699–738.

**Bonneton, Nicolas and Christopher Sandmann**, "Non-Stationary Search and Assortative Matching," 2023. mimeo.

**Borovičková, Katarína and Robert Shimer**, "High Wage Workers Work for High Wage Firms," 2020. Mimeo.

**Buera, Francisco J and Ezra Oberfield**, "The Global Diffusion of Ideas," *Econometrica*, 2020, *88* (1), 83–114.

**Burdett, Ken and Melvyn G Coles**, "Marriage and Class," *Quarterly Journal of Economics*, 1997, *112* (1), 141–168.

**Burdett, Kenneth and Dale Mortensen**, "Wage Differentials, Employer Size, and Unemployment," *International Economic Review*, 1998, *39* (2), 257–73.

**Card, David, Ana Rute Cardoso, and Patrick Kline**, "Bargaining, sorting, and the gender wage gap: Quantifying the impact of firms on the relative pay of women," *Quarterly Journal of Economics*, 2016, *131* (2), 633–686.

**_ , _ , Joerg Heining, and Patrick Kline**, "Firms and Labor Market Inequality: Evidence and Some Theory," *Journal of Labor Economics*, 2018, *36* (S1), S13–S70.

**_ , Jörg Heining, and Patrick Kline**, "Workplace Heterogeneity and the Rise of West German Wage Inequality," *Quarterly Journal of Economics*, 2013, *128* (3), 967–1015.

**Diamond, Peter A. and Eric Maskin**, "An Eequilibrium Analysis of Search and Breach of Contract, I: Steady States," *Bell Journal of Economics*, 1979, *10* (1), 282–316.

**Eeckhout, Jan and Philipp Kircher**, "Sorting and Decentralized Price Competition," *Econometrica*, 2010, *78* (2), 539–574.

**Engbom, Niklas, Christian Moser, and Jan Sauermann**, "Firm pay dynamics," *Journal of Econometrics*, 2023, *233* (2), 396–423.

**Flinn, Christopher and James Heckman**, "New Methods for Analyzing Structural Models of Labor Force Dynamics," *Journal of Econometrics*, 1982, *18* (1), 115–168.

**Goussé, Marion, Nicolas Jacquemet, and Jean-Marc Robin**, "Marriage, Labor Supply, and Home Production," *Econometrica*, 2017, *85* (6), 1873–1919.

**Hagedorn, Marcus, Tzuo Hann Law, and Iourii Manovskii**, "Identifying Equilibrium Models of Labor Market Sorting," *Econometrica*, 2017, *85* (1), 29–65.

**Hornstein, Andreas, Per Krusell, and Giovanni L. Violante**, "Frictional Wage Dispersion in Search Models: A Quantitative Assessment," *American Economic Review*, 2011, *101* (7), 2873–2898.

**Hwang, Hae-shin, Dale T Mortensen, and W Robert Reed**, "Hedonic wages and labor market search," *Journal of Labor Economics*, 1998, *16* (4), 815–847.

**Jolivet, Gregory, Fabien Postel-Vinay, and Jean-Marc Robin**, "The empirical con-

tent of the job search model: Labor mobility and wage distributions in Europe and the US," *Contributions to Economic Analysis*, 2006, *275*, 269–308.

**Kline, Patrick, Raffaele Saggio, and Mikkel Sølvsten**, "Leave-out estimation of variance components," *Econometrica*, 2020, *88* (5), 1859–1898.

**Lachowska, Marta, Alexandre Mas, Raffaele Saggio, and Stephen A Woodbury**, "Do firm effects drift? Evidence from Washington administrative data," *Journal of Econometrics*, 2023, *233* (2), 375–395.

**Lamadon, Thibaut, Jeremy Lise, Costas Meghir, and Jean-Marc Robin**, "Labor Market Matching, Wages, and Amenities," 2024. NBER Working Paper 32687.

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

**Oberfield, Ezra**, "A Theory of Input-Output Architecture," *Econometrica*, 2018, *86* (2), 559–589.

**Postel-Vinay, Fabien and Jean-Marc Robin**, "Equilibrium Wage Dispersion with Worker and Employer Heterogeneity," *Econometrica*, 2002, *70* (6), 2295–2350.

**Shi, Shouyong**, "Frictional Assignment. I. Efficiency," *Journal of Economic Theory*, 2001, *98* (2), 232–260.

_ , "Frictional Assignment, Part II: Infinite Horizon and Unequality," *Review of Economic Dynamics*, 2005, *8* (1), 106–137.

**Shimer, Robert**, "The Assignment of Workers to Jobs in an Economy with Coordination Frictions," *Journal of Political Economy*, 2005, *113* (5), 996–1025.

_ **and Lones Smith**, "Assortative Matching and Search," *Econometrica*, 2000, *68* (2), 343–369.

**Smith, Lones**, "The Marriage Model with Search Frictions," *Journal of political Economy*, 2006, *114* (6), 1124–1144.

**Song, Jae, David J Price, Fatih Guvenen, Nicholas Bloom, and Till Von Wachter**, "Firming up Inequality," *Quarterly Journal of Economics*, 2019, *134* (1), 1–50.

**Topel, Robert H. and Michael P. Ward**, "Job Mobility and the Careers of Young Men," *Quarterly Journal of Economics*, 1992, *107* (2), 439–479.

# Appendix

## A Proofs

### A.1 Existence and Monotonicity

**Proof of Proposition 1.** Consider the following $2(X+Y)$ functions of $(\bar{w}_x, u_x)_{x=1}^X$ and $(\bar{\pi}_y, v_y)_{y=1}^Y$:

$$
\begin{aligned}
T_{1,x}(\bar{w},\bar{\pi},u,v) &= \frac{\gamma\rho}{r+\delta}\sum_{y=1}^{Y} v_y f_{x,y}\int_{\bar{z}_{x,y}}^{\infty}(z-\bar{z}_{x,y})s(z)dz,\\
T_{2,y}(\bar{w},\bar{\pi},u,v) &= \frac{(1-\gamma\rho)}{r+\delta}\sum_{x=1}^{X} u_x f_{x,y}\int_{\bar{z}_{x,y}}^{\infty}(z-\bar{z}_{x,y})s(z)dz,\\
T_{3,x}(\bar{w},\bar{\pi},u,v) &= m_x - \sum_{y=1}^{Y}\frac{\rho u_x v_y S(\bar{z}_{x,y})}{\delta},\\
T_{4,y}(\bar{w},\bar{\pi},u,v) &= n_y - \sum_{x=1}^{X}\frac{\rho u_x v_y S(\bar{z}_{x,y})}{\delta},
\end{aligned}
$$

where $\bar{z}_{x,y} = (\bar{w}_x + \bar{v}_y)/f_{x,y}$, as in equation (6). We refer to this mapping collectively as $T = (T_1, T_2, T_3, T_4)$. It is immediate that any vector $(\bar{w}, \bar{\pi}, u, v)$, together with $\bar{z}$ solving equation (6) and $\phi$ solving equation (12), is an equilibrium if and only if it is a fixed point of $T$, i.e. $T(\bar{w},\bar{\pi},u,v) = (\bar{w},\bar{\pi},u,v)$.

Now define the mapping $\tilde{T}$:

$$
\begin{aligned}
\tilde{T}_{1,x}(\bar{w},\bar{\pi},u,v) &= \min\left\{T_{1,x}(\bar{w},\bar{\pi},u,v), \bar{\bar{w}}_x\right\},\\
\tilde{T}_{2,y}(\bar{w},\bar{\pi},u,v) &= \min\left\{T_{2,y}(\bar{w},\bar{\pi},u,v), \bar{\bar{\pi}}_y\right\},\\
\tilde{T}_{3,x}(\bar{w},\bar{\pi},u,v) &= \max\left\{T_{3,x}(\bar{w},\bar{\pi},u,v), 0\right\},\\
\tilde{T}_{4,y}(\bar{w},\bar{\pi},u,v) &= \max\left\{T_{4,y}(\bar{w},\bar{\pi},u,v), 0\right\},
\end{aligned}
$$

where

$$
\begin{aligned}
\bar{\bar{w}}_x &= \frac{\gamma\rho}{r+\delta}\sum_{y=1}^{Y} n_y f_{x,y}\int_{\bar{\bar{w}}_x/f_{x,y}}^{\infty}(z-\bar{\bar{w}}_x/f_{x,y})s(z)dz,\\
\bar{\bar{\pi}}_y &= \frac{(1-\gamma)\rho}{r+\delta}\sum_{x=1}^{X} m_x f_{x,y}\int_{\bar{\bar{\pi}}_y/f_{x,y}}^{\infty}(z-\bar{\bar{\pi}}_y/f_{x,y})s(z)dz.
\end{aligned}
$$

The mapping $\tilde{T}$ is continuous. Moreover, it maps points satisfying $\bar{w}_x \in [0, \bar{\bar{w}}_x]$, $\bar{\pi}_y \in [0, \bar{\bar{\pi}}_y]$, $u_x \in [0, m_x]$, and $v_y \in [0, n_y]$ for all $x \in \{1, \ldots, X\}$ and $y \in \{1, \ldots, Y\}$ into itself. Therefore $\tilde{T}$ has a fixed point by Brouwer's fixed point theorem.

We prove that at any $(\bar{w}, \bar{\pi}, u, v)$ which is a fixed point of $\tilde{T}$, $T(\bar{w}, \bar{\pi}, u, v) = \tilde{T}(\bar{w}, \bar{\pi}, u, v)$, and thus $(\bar{w}, \bar{\pi}, u, v)$ is also a fixed point of $T$. In the first step we prove that $u_x > 0$ for all $x$. This is because if $u_x = 0$, $\tilde{T}_{3,x}(\bar{w}, \bar{\pi}, u, v) = m_x$, contradicting $(\bar{w}, \bar{\pi}, u, v)$ being a fixed point. This implies $\tilde{T}_{3,x}(\bar{w}, \bar{\pi}, u, v) = T_{3,x}(\bar{w}, \bar{\pi}, u, v)$ for all $x$ at any fixed point. Similarly $v_y > 0$ at any fixed point of $\tilde{T}$ and so $\tilde{T}_{4,y}(\bar{w}, \bar{\pi}, u, v) = T_{4,y}(\bar{w}, \bar{\pi}, u, v)$ for all $y$ at any fixed point.

Next, any fixed point has $\bar{w}_x > 0$ for all $x$: $\tilde{T}_{1,x}(\bar{w}, \bar{\pi}, u, v)$ is continuous and decreasing in $\bar{w}_x$ and is strictly positive at $\bar{w}_x = 0$ since $v_y > 0$ for all $y$. Similarly any fixed point has $\bar{\pi}_y > 0$ for all $y$.

Finally, in any fixed point of $\tilde{T}$, any solution to $\tilde{T}_{1,x}(\bar{w}, \bar{\pi}, u, v) = \bar{w}_x$ has $\bar{w}_x < \bar{\bar{w}}_x$ because $\bar{\pi}_y > 0$. This implies that the fixed point of $\tilde{T}$ also solves $\tilde{T}_{1,x}(\bar{w}, \bar{\pi}, u, v) = T_{1,x}(\bar{w}, \bar{\pi}, u, v)$ for all $x$. The same logic implies $\tilde{T}_{2,y}(\bar{w}, \bar{\pi}, u, v) = T_{2,y}(\bar{w}, \bar{\pi}, u, v)$ for all $y$ in any fixed point.

In summary, we have proved that there exists a vector $(\bar{w}, \bar{\pi}, u, v)$ with $\tilde{T}(\bar{w}, \bar{\pi}, u, v) = (\bar{w}, \bar{\pi}, u, v)$, that at any such vector $T(\bar{w}, \bar{\pi}, u, v) = \tilde{T}(\bar{w}, \bar{\pi}, u, v)$ and so $(\bar{w}, \bar{\pi}, u, v)$ is a fixed point of $T$, and that any fixed point of $T$ is an equilibrium. This proves an equilibrium exists.

Along the way we also proved that $\bar{w}_x > 0$ for all $x$ and $\bar{\pi}_y > 0$ for all $y$ at any fixed point of $\tilde{T}$, and hence in any equilibrium. ■

**Proof of Lemma 1.** We prove that $\bar{w}_x$ is strictly increasing. The proof that $\bar{\pi}_y$ is strictly increasing is analogous.

To find a contradiction, suppose there exists an $x_1 < x_2$ with $\bar{w}_{x_1} \geq \bar{w}_{x_2}$. Since $f$ is monotonic, $f_{x_2,y} > f_{x_1,y}$ for all $y$. Then using equation (6), we have $\bar{z}_{x_1,y} > \bar{z}_{x_2,y}$ for all $y$. Additionally, observe that $\int_{\bar{z}}^{\infty}(z - \bar{z})s(z)dz$ is strictly positive for all $\bar{z}$, since $S(z) > 0$ for all $z$. Additionally, the integral is decreasing in $\bar{z}$, as can be confirmed directly. This means that $\int_{\bar{z}_{x_2,y}}^{\infty}(z - \bar{z}_{x_2,y})s(z)dz > \int_{\bar{z}_{x_1,y}}^{\infty}(z - \bar{z}_{x_1,y})s(z)dz > 0$.

Putting this together, if there exists an $x_1 < x_2$ with $\bar{w}_{x_1} \geq \bar{w}_{x_2}$,

$$\begin{aligned}\bar{w}_{x_2} &= \frac{\gamma\rho}{r+\delta}\sum_{y=1}^{Y} v_y f_{x_2,y}\int_{\bar{z}_{x_2,y}}^{\infty}(z - \bar{z}_{x_2,y})s(z)dz \\ &> \frac{\gamma\rho}{r+\delta}\sum_{y=1}^{Y} v_y f_{x_1,y}\int_{\bar{z}_{x_1,y}}^{\infty}(z - \bar{z}_{x_1,y})s(z)dz = \bar{w}_{x_1},\end{aligned}$$

where the two equations use the value function (10) and the inequality uses $f_{x_2,y} > f_{x_1,y}$ and $\int_{\bar{z}_{x_2,y}}^{\infty}(z-\bar{z}_{x_2,y})s(z)dz > \int_{\bar{z}_{x_1,y}}^{\infty}(z-\bar{z}_{x_1,y})s(z)dz > 0$, together with strict positivity of the remaining terms. But this is a contradiction, proving $\bar{w}_{x_1} < \bar{w}_{x_2}$. ■

## A.2 Average Log Wage

**Proof of Proposition 2.** We prove a more general version of this proposition. For any strictly increasing function $G : \mathbb{R}_+ \to \mathbb{R}$, define

$$w_{x,y}^G = \frac{\int_{\bar{z}_{x,y}}^{\infty} G(W_{x,y}(z))s(z)dz}{S(\bar{z}_{x,y})}. \tag{41}$$

From equations (6) and (9), we have

$$W_{x,y}(z) = \bar{w}_x + \gamma(\bar{w}_x + \bar{\pi}_y)\left(\frac{z}{\bar{z}_{x,y}} - 1\right).$$

Using equation (41) and the functional form of the Pareto distribution, we obtain

$$w_{x,y}^G = \frac{\int_{\bar{z}_{x,y}}^{\infty} G\Big(\bar{w}_x + \gamma(\bar{w}_x + \bar{\pi}_y)\big(z/\bar{z}_{x,y} - 1\big)\Big)\theta z^{-\theta-1}dz}{\bar{z}_{x,y}^{-\theta}}.$$

Now let $q = z/\bar{z}_{x,y} - 1$ and perform a change in the variable of integration to obtain

$$w_{x,y}^G \equiv \int_0^{\infty} G\big(\bar{w}_x + \gamma(\bar{w}_x + \bar{\pi}_y)q\big)\theta(1+q)^{-\theta-1}dq.$$

Setting $G(W) = \log(W)$ for all $W$ gives us equation (17). This equation implies that $w_{x,y}^G$ is simply a weighted average of $G\big(\bar{w}_x + \gamma(\bar{w}_x + \bar{\pi}_y)q\big)$, with the same weights for all $(x, y)$. This means that $x$ and $y$ only affect $w_{x,y}^G$ through $\bar{w}_x$ and $\bar{\pi}_y$.

Now since $G$ is strictly increasing, it is straightforward to verify that increasing either $\bar{w}_x$ or $\bar{\pi}_y$ raises the integrand in equation (17) at all $q > 0$, and hence raises $w_{x,y}^G$. This establishes the first two enumerated points.

We next prove the third point if $G$ is strictly concave. If $G$ is strictly convex, we prove $w_{x_1,y_2}^G + w_{x_2,y_1}^G < w_{x_1,y_1}^G + w_{x_2,y_2}^G$. If $G$ is linear, this is an equality. Take $x_1$ and $x_2$ with $\bar{w}_{x_1} < \bar{w}_{x_2}$; and $y_1$ and $y_2$ with $\bar{\pi}_{y_1} < \bar{\pi}_{y_2}$. Let

$$\lambda \equiv \frac{(1+\gamma q)(\bar{w}_{x_2} - \bar{w}_{x_1})}{(1+\gamma q)(\bar{w}_{x_2} - \bar{w}_{x_1}) + \gamma q(\bar{\pi}_{y_2} - \bar{\pi}_{y_1})}.$$

The assumptions on $x_1$, $x_2$, $y_1$, and $y_2$ ensure that $\lambda \in (0,1)$ for all $q > 0$. Then verify algebraically that

$$\bar{w}_{x_1} + \gamma(\bar{w}_{x_1} + \bar{\pi}_{y_2})q = \lambda(\bar{w}_{x_1} + \gamma(\bar{w}_{x_1} + \bar{\pi}_{y_1})q) + (1-\lambda)(\bar{w}_{x_2} + \gamma(\bar{w}_{x_2} + \bar{\pi}_{y_2})q), \text{ and}$$
$$\bar{w}_{x_2} + \gamma(\bar{w}_{x_2} + \bar{\pi}_{y_1})q = (1-\lambda)(\bar{w}_{x_1} + \gamma(\bar{w}_{x_1} + \bar{\pi}_{y_1})q) + \lambda(\bar{w}_{x_2} + \gamma(\bar{w}_{x_2} + \bar{\pi}_{y_2})q).$$

Thus $G$ concave (convex) implies

$$G(\bar{w}_{x_1} + \gamma(\bar{w}_{x_1} + \bar{\pi}_{y_2})q) > (<)\lambda G(\bar{w}_{x_1} + \gamma(\bar{w}_{x_1} + \bar{\pi}_{y_1})q) + (1-\lambda)G(\bar{w}_{x_2} + \gamma(\bar{w}_{x_2} + \bar{\pi}_{y_2})q), \text{ and}$$
$$G(\bar{w}_{x_2} + \gamma(\bar{w}_{x_2} + \bar{\pi}_{y_1})q) > (<)(1-\lambda)G(\bar{w}_{x_1} + \gamma(\bar{w}_{x_1} + \bar{\pi}_{y_1})q) + \lambda G(\bar{w}_{x_2} + \gamma(\bar{w}_{x_2} + \bar{\pi}_{y_2})q).$$

Summing these gives

$$G\big(\bar{w}_{x_1} + \gamma(\bar{w}_{x_1} + \bar{\pi}_{y_2})q\big) + G\big(\bar{w}_{x_2} + \gamma(\bar{w}_{x_2} + \bar{\pi}_{y_1})q\big) \\ > (<)G\big(\bar{w}_{x_1} + \gamma(\bar{w}_{x_1} + \bar{\pi}_{y_1})q\big) + G\big(\bar{w}_{x_2} + \gamma(\bar{w}_{x_2} + \bar{\pi}_{y_2})q\big)$$

when $G$ is concave (convex). Integrating over the density $\theta(1+q)^{-\theta-1}$ for $q > 0$ delivers the third bullet point.

Finally, we note that if $f_{x,y}$ is strictly increasing in $x$ and $y$, Lemma 1 implies $\bar{w}_x$ and $\bar{\pi}_y$ are strictly increasing. Then the three numbered conditions in the statement of the proposition imply that $w^G_{x,y}$ is strictly increasing and strictly submodular (supermodular) in $x$ and $y$ when $G$ is strictly concave (convex). ■

Before proving Proposition 3, we establish a preliminary result:

**Lemma 2** *Assume $f_{x,y} = f^1_x f^2_y$ for strictly positive and strictly increasing functions $f^1$ and $f^2$. Then $\bar{w}_x/f^1_x$ and $\bar{\pi}_y/f^2_y$ are strictly increasing.*

**Proof.** We prove that $\bar{w}_x/f^1_x$ is strictly increasing. To find a contradiction, assume there is an $x_1 < x_2$ with $\bar{w}_{x_1}/f^1_{x_1} \geq \bar{w}_{x_2}/f^1_{x_2}$. Note $x_1 < x_2$ implies $1/f^1_{x_1} > 1/f^1_{x_2}$. Equation (6) states

$$\bar{z}_{x,y} = \frac{\frac{\bar{w}_x}{f^1_x} + \frac{\bar{\pi}_y}{f^1_x}}{f^2_y},$$

and so $\bar{z}_{x_1,y} > \bar{z}_{x_2,y}$ for all $y$. Moreover, equation (10) implies

$$\frac{\bar{w}_x}{f^1_x} = \frac{\gamma\rho}{r+\delta}\sum_{y=1}^{Y} v_y f^2_y \int_{\bar{z}_{x,y}}^{\infty} (z - \bar{z}_{x,y}) s(z) dz.$$

Since the integral is decreasing in $\bar{z}_{x,y}$, we have $\frac{\bar{w}_{x_1}}{f^1_{x_1}} < \frac{\bar{w}_{x_2}}{f^1_{x_2}}$, a contradiction.

The proof that $\bar{\pi}_y/f^2_y$ is strictly increasing is identical, manipulating equation (11) instead of equation (10). ■

**Proof of Proposition 3.** For the first part of the Proposition, we again work with the more general version, using $w^G$ defined in equation (41) for an arbitrary increasing function $G$. From the wage equation (9), and the exponential distribution, we have

$$w^G_{x,y} = \theta \int_{\bar{z}_{x,y}}^{\infty} G(\bar{w}_x + \gamma(zf_{x,y} - \bar{w}_x - \bar{\pi}_y))e^{-\theta(z-\bar{z}_{x,y})}dz.$$

Let $q = z - \bar{z}_{x,y}$ to get

$$w^G_{x,y} = \theta \int_0^{\infty} G(\bar{w}_x + \gamma((q + \bar{z}_{x,y})f_{x,y} - \bar{w}_x - \bar{\pi}_y))e^{-\theta q}dq.$$

From equation (6), we can reduce this to

$$w^G_{x,y} = \theta \int_0^{\infty} G(\bar{w}_x + \gamma q f_{x,y})e^{-\theta q}dq,$$

which is equivalent to equation (18) when $G(W) = \log W$ for all $W$. If $f$ is strictly increasing, so is $\bar{w}_x$ (Lemma 1). And since $G$ is strictly increasing, then the integrand in equation (18) is strictly increasing in both $x$ and $y$ for all $q$. Thus $w^G$ is strictly increasing.

Finally we turn to the multiplicatively separable case, $f_{x,y} = f^1_x f^2_y$. Here we assume $G(W) = \log W$. We use equation (18). $w^*$ is strictly submodular if

$$\log(\bar{w}_x + \gamma q f^1_x f^2_y)$$

is strictly submodular for all $q > 0$. So take $x_1 < x_2$ and $y_1 < y_2$. Strict submodularity requires

$$\log(\bar{w}_{x_1} + \gamma q f^1_{x_1} f^2_{y_1}) + \log(\bar{w}_{x_2} + \gamma q f^1_{x_2} f^2_{y_2}) < \log(\bar{w}_{x_1} + \gamma q f^1_{x_1} f^2_{y_2}) + \log(\bar{w}_{x_2} + \gamma q f^1_{x_2} f^2_{y_1}),$$

or equivalently

$$(\bar{w}_{x_1} + \gamma q f^1_{x_1} f^2_{y_1})(\bar{w}_{x_2} + \gamma q f^1_{x_2} f^2_{y_2}) < (\bar{w}_{x_1} + \gamma q f^1_{x_1} f^2_{y_2})(\bar{w}_{x_2} + \gamma q f^1_{x_2} f^2_{y_1}).$$

Expand the products and cancel common terms to get

$$\bar{w}_{x_1} f^1_{x_2} f^2_{y_2} + \bar{w}_{x_2} f^1_{x_1} f^2_{y_1} < \bar{w}_{x_1} f^1_{x_2} f^2_{y_1} + \bar{w}_{x_2} f^1_{x_1} f^2_{y_2}$$

or

$$\left(\frac{\bar{w}_{x_2}}{f^1_{x_2}} - \frac{\bar{w}_{x_1}}{f^1_{x_1}}\right)(f^2_{y_2} - f^2_{y_1}) > 0.$$

Lemma 2 implies the first term is positive for all $x_1 < x_2$ and the second term is positive by assumption. ■

## A.3 Assortative Matching

**Proof of Proposition 4.** We first prove that $\bar{z}_{x,y}$ is strictly log-submodular. Substituting equation (6) for $\bar{z}$, we must prove that for $x_1 < x_2$ and $y_1 < y_2$,

$$\left(\frac{\bar{w}_{x_1} + \bar{\pi}_{y_1}}{f_{x_1,y_1}}\right)\left(\frac{\bar{w}_{x_2} + \bar{\pi}_{y_2}}{f_{x_2,y_2}}\right) < \left(\frac{\bar{w}_{x_1} + \bar{\pi}_{y_2}}{f_{x_1,y_2}}\right)\left(\frac{\bar{w}_{x_2} + \bar{\pi}_{y_1}}{f_{x_2,y_1}}\right).$$

Weak log supermodularity of $f$ implies $f_{x_1,y_1} f_{x_2,y_2} \geq f_{x_1,y_2} f_{x_2,y_1}$. And we see that the product of the numerators on the left hand side is smaller than the product of the numerators on the right hand side if and only if

$$\bar{w}_{x_1}\bar{\pi}_{y_2} + \bar{w}_{x_2}\bar{\pi}_{y_1} < \bar{w}_{x_1}\bar{\pi}_{y_1} + \bar{w}_{x_2}\bar{\pi}_{y_2} \Leftrightarrow (\bar{w}_{x_2} - \bar{w}_{x_1})(\bar{\pi}_{y_2} - \bar{\pi}_{y_1}) > 0.$$

This is immediate because $\bar{w}$ and $\bar{\pi}$ are strictly increasing (Lemma 1).

We now use the assumption that $S(z) = (z/z_0)^{-\theta}$. Then since $\log \bar{z}_{x,y}$ is strictly submodular, $\log S(\bar{z}_{x,y}) = \theta \log z_0 - \theta \log \bar{z}_{x,y}$ is strictly supermodular. Finally, equation (12) implies that $\log \phi_{x,y}$ inherits the strict supermodularity of $\log S(\bar{z}_{x,y})$. ■

**Proof of Proposition 5.** We split the proof into two pieces.

**$-1/f_{x,y}$ weakly supermodular.** Take any $x_1 < x_2$ and $y_1 < y_2$. Then

$$\left(\frac{1}{f_{x_2,y_2}} - \frac{1}{f_{x_2,y_1}}\right)(\bar{w}_{x_2} - \bar{w}_{x_1}) < 0,$$
$$\left(\frac{1}{f_{x_2,y_2}} - \frac{1}{f_{x_1,y_2}}\right)(\bar{\pi}_{y_2} - \bar{\pi}_{y_1}) < 0,$$

since in both cases the first term is negative (since $f$ is positive and strictly increasing) and the second term is positive by Lemma 1. Additionally, $1/f$ submodular implies

$$\frac{1}{f_{x_1,y_1}} + \frac{1}{f_{x_2,y_2}} \leq \frac{1}{f_{x_1,y_2}} + \frac{1}{f_{x_2,y_1}}.$$

Multiply each term in the last inequality by $\bar{w}_{x_1} + \bar{\pi}_{y_1}$ and add to the preceding inequalities to get

$$\frac{\bar{w}_{x_1} + \bar{\pi}_{y_1}}{f_{x_1,y_1}} + \frac{\bar{w}_{x_2} + \bar{\pi}_{y_2}}{f_{x_2,y_2}} < \frac{\bar{w}_{x_1} + \bar{\pi}_{y_2}}{f_{x_1,y_2}} + \frac{\bar{w}_{x_2} + \bar{\pi}_{y_1}}{f_{x_2,y_1}},$$

From equation (6), this implies that $\bar{z}_{x,y}$ is strictly submodular.

Next, in steady state with an exponential distribution, we have from equation (12) that

$$\log \phi_{x,y} = \log(\rho/\delta) + \log u_x + \log v_y - \theta \bar{z}_{x,y}.$$

Since $\bar{z}_{x,y}$ is strictly submodular and the other terms are amodular, this proves that $\log \phi_{x,y}$ is strictly supermodular, i.e. that $\phi$ is strictly log-supermodular.

$f_{x,y}$ **multiplicatively separable.** Assume $f_{x,y} = f^1_x f^2_y$ for some strictly positive and strictly increasing functions $f^1$ and $f^2$. As in the other case, we must prove that for all $x_2 > x_1$ and $y_2 > y_1$, $\bar{z}_{x,y}$ is strictly submodular:

$$\frac{\bar{w}_{x_1} + \bar{\pi}_{y_1}}{f^1_{x_1} f^2_{y_1}} + \frac{\bar{w}_{x_2} + \bar{\pi}_{y_2}}{f^1_{x_2} f^2_{y_2}} < \frac{\bar{w}_{x_1} + \bar{\pi}_{y_2}}{f^1_{x_1} f^2_{y_2}} + \frac{\bar{w}_{x_2} + \bar{\pi}_{y_1}}{f^1_{x_2} f^2_{y_1}}.$$

Regroup terms to write this as

$$\left(\frac{\bar{w}_{x_1}}{f^1_{x_1}} - \frac{\bar{w}_{x_2}}{f^1_{x_2}}\right)\left(\frac{1}{f^2_{y_1}} - \frac{1}{f^2_{y_2}}\right) + \left(\frac{\bar{\pi}_{y_1}}{f^2_{y_1}} - \frac{\bar{\pi}_{y_2}}{f^2_{y_2}}\right)\left(\frac{1}{f^1_{x_1}} - \frac{1}{f^1_{x_2}}\right) < 0.$$

Lemma 2 establishes that of $\bar{w}_x/f^1_x$ is increasing. A parallel (omitted) proof establishes monotonicity of $\bar{\pi}_y/f^2_y$. Thus both of the terms on the left hand side are negative, proving the result. ■

# B Log-Linear Wage Equation

Most of the empirical literature does not aim to estimate the average log wage, $w^*_{x,y}$. Instead, following Abowd, Kramarz and Margolis (1999), authors impose a log-linear wage structure to estimate worker and firm fixed effects. In this section we analyze what that procedure recovers if our model is data-generating process, and in particular the relationship between a log-linear wage equation and the average log wage $w^*_{x,y}$.

## B.1 Econometric Framework

Consider a panel data set containing the wage $W_{i,t}$ of worker $i \in \{1, \ldots, I\}$ at time $t \in \{1, 2, \ldots, T\}$ as well as the employer identifier $\mathcal{J}_{i,t} \in \{1, \ldots, J\}$. Since the worker may not always be employed, we let $\mathcal{T}_i \subseteq \{1, 2, \ldots, T\}$ denote the periods when the worker earns a wage. We assume, in line with the literature, that we observe neither a wage nor a wage offer for $i$ when they are not employed, at $t \in \mathcal{T}_i^c \equiv \{1, 2, \ldots, T\} \backslash \mathcal{T}_i$. For notational simplicity and following the literature, we impose that each worker only works for one firm at each point in time, for example by focusing on their main job in each period that they are employed.

Following Abowd, Kramarz and Margolis (1999), we could regress the log wage on a full set of worker and firm fixed effects and an error term: For all $i \in \{1, \ldots, I\}$ and $t \in \mathcal{T}_i$,

$$\log W_{i,t} = \alpha_i + \psi_{\mathcal{J}_{i,t}} + \varepsilon_{i,t}. \tag{42}$$

We are interested in the coefficient estimates $\hat{\alpha}_i$ and $\hat{\psi}_j$ when estimating equation (42) using ordinary least squares (OLS). Regardless of the economic model and the data set, OLS is a statistical procedure which minimizes the sum of squared errors:

$$\{\hat{\alpha}_i, \hat{\psi}_j\} = \arg \min_{\{\alpha_i, \psi_j\}} \sum_{i=1}^{I} \sum_{t \in \mathcal{T}_i} (\log W_{i,t} - \alpha_i - \psi_{\mathcal{J}_{i,t}})^2. \tag{43}$$

The first order condition for $\alpha_i$ from equation (43) is

$$\sum_{t \in \mathcal{T}_i} (\log W_{i,t} - \hat{\alpha}_i - \hat{\psi}_{\mathcal{J}_{i,t}}) = 0. \tag{44}$$

Symmetrically, define $\mathcal{I}_{j,t}$ as the set of workers whom $j$ employs at $t$, so $j = \mathcal{J}_{i,t}$ if and only if $i \in \mathcal{I}_{j,t}$.[12] Then the first order condition for $\psi_j$ is

$$\sum_{t=1}^{T} \sum_{i \in \mathcal{I}_{j,t}} (\log W_{i,t} - \hat{\alpha}_i - \hat{\psi}_j) = 0. \tag{45}$$

Under the assumption that all workers and firms are connected though the matching graph,[13] Abowd, Creecy and Kramarz (2002) establish that equations (44) and (45) pin down $\hat{\alpha}$ and $\hat{\psi}$ up to an additive constant. That is, we can increase all the worker fixed effects by $k$ and

[12] In our model, $\mathcal{I}_{j,t}$ has either zero or one element, depending on whether the job is filled or vacant. In real world data, firms can employ multiple workers and so $\mathcal{I}_{j,t}$ typically has multiple elements.

[13] Formally, we require that any worker $i_0$ can be linked to any firm $j$ through a finite sequence of steps $t_0, t_1, \ldots, t_n$: $j_t = \mathcal{J}_{i_{t-1},t}$ for $t$ odd and $i_t \in \mathcal{I}_{j_{t-1},t}$ for $t$ even, with $j = \mathcal{J}_{t_n}$.

decrease all the firm fixed effects by $k$ without changing the fit of equation (43).

## B.2 Estimates in Model-Generated Data

Next consider estimating equation (42) using an ideal data set generated by our model. We assume that there is a large number of workers $I$ and a large number of firms $J$. Each worker $i$ has an unobserved type $x_i$, and similarly each firm $j$ has an unobserved type $y_j$. We assume $i$ and $j$ behave according to the decision rules in our model. That is, when $i$ is unemployed, they meet a type-$y$ vacant job in a match with productivity at least $z$ at rate $S(z)v_y$, and they accept the job and earn a wage $W_{x_i,y}(z)$ if and only if $z \geq \bar{z}_{x_i,y}$. Symmetrically, when $j$ has a vacant job, it meets a type-$x$ unemployed worker in a match with productivity at least $z$ at rate $S(z)u_x$, and it hires the worker and earns profits $zf_{x,y_j} - W_{x,y_j}(z)$ if and only if $z \geq \bar{z}_{x,y_j}$.

We are interested in an environment where there is a large but finite number of workers and jobs and where we observe each worker and job for a very long time, $T \to \infty$.[14] In this case, worker $i$ with type $x_i$ will spend a fraction $u_{x_i}/m_{x_i}$ of their time unemployed and fraction $\phi_{x_i,y}/m_{x_i}$ of their time matched to a type-$y$ firm. In such matches, the density of match productivity will be $s(z)/S(\bar{z}_{x_i,y})$ for $z \geq \bar{z}_{x_i,y}$ and the wage will be $W_{x_i,y}(z)$. Similarly, the relative likelihood of firm $j$ with type $y_j$ matching with a type-$x$ worker is proportional to $\phi_{x,y_j}$. Again, in such matches, the density of match productivity will be $s(z)/S(\bar{z}_{x,y_j})$ for $z \geq \bar{z}_{x,y_j}$ and the wage will be $W_{x,y_j}(z)$. Since there is no uncertainty about these long-run distributions, $\hat{\alpha}_i$ and $\hat{\psi}_j$ have well-behaved limits in the limit as $T \to \infty$. These limits depend only on the worker's and firm's type, since the distribution of partners and wages only depend on types. We let $\bar{\alpha}_{x_i}$ denote the limiting value of $\hat{\alpha}_i$ and $\bar{\psi}_{y_j}$ denote the limiting value of $\hat{\psi}_j$ when $T \to \infty$.[15] We are interested in characterizing and interpreting those values.

Using equations (44) and (45) and the model structure, we obtain

$$\bar{\alpha}_x = \frac{\sum_{y=1}^{Y}(w^*_{x,y} - \bar{\psi}_y)\phi_{x,y}}{\sum_{y=1}^{Y}\phi_{x,y}} = \lambda_x - \frac{\sum_{y=1}^{Y}\bar{\psi}_y\phi_{x,y}}{\sum_{y=1}^{Y}\phi_{x,y}}, \tag{46}$$

$$\bar{\psi}_y = \frac{\sum_{x=1}^{X}(w^*_{x,y} - \bar{\alpha}_x)\phi_{x,y}}{\sum_{x=1}^{X}\phi_{x,y}} = \mu_y - \frac{\sum_{x=1}^{X}\bar{\alpha}_x\phi_{x,y}}{\sum_{x=1}^{X}\phi_{x,y}}, \tag{47}$$

[14] In the real world, $T$ is finite, so the OLS estimates $\hat{\alpha}_i$ and $\hat{\psi}_j$ are unbiased but noisy estimates of $\bar{\alpha}_{x_i}$ and $\bar{\psi}_{y_j}$, respectively. This creates econometric issues which we sidestep in this paper through our idealized "large $T$" assumption.

[15] In practice, the literature estimating equation (42) looks only at one job per worker per year. In Abowd, Kramarz and Margolis (1999), this is the job where the individual works the most days. In Card, Cardoso and Kline (2016), this is the job where the worker has the most hours during a reference week. In

where $w^*_{x,y}$ is the average log wage in an $(x, y)$ match, defined in equation (15), $\phi_{x,y}$ is the measure of $(x, y)$ matches, defined in equation (12), and $\lambda_x$ and $\mu_y$ are defined in equations (19) and (20). Our model ensures that all workers and firms are connected when $T \to \infty$, since $S(\bar{z}_{x,y}) > 0$ for all $(x, y)$. That means that we can solve equations (46) and (47) for $(\bar{\alpha}_x, \bar{\psi}_y)$ up to the additive constant $k$ mentioned before.

If the average log wage were an additively separable function of the worker's type and the firm's type, $w^*_{x,y} = a_x + b_y$, then equations (46) and (47) would imply $\bar{\alpha}_x = a_x + k$ and $\bar{\psi}_y = b_y - k$, where $k$ is the irrelevant additive constant discussed before. In this case, $\bar{\psi}_{y_2} - \bar{\psi}_{y_1}$ is the average difference in the log wage that any worker earns at a type-$y_2$ firm compared to a type-$y_1$ firm when following the equilibrium decision rules. Our model implies that $w^*_{x,y}$ is increasing in $x$ and $y$ but is submodular rather than linear.

Finally, we follow Abowd, Kramarz and Margolis (1999), Andrews, Gill, Schank and Upward (2008), Card, Heining and Kline (2013), and others in focusing on the cross-sectional variance of $\bar{\alpha}$ and $\bar{\psi}$ as well as the covariance between $\alpha$ and $\psi$ in matched pairs. Define

$$\mathbb{E}(g_{x,y}) \equiv \frac{\sum_{x=1}^{X} \sum_{y=1}^{Y} g_{x,y} \phi_{x,y}}{\sum_{x=1}^{X} \sum_{y=1}^{Y} \phi_{x,y}}$$

for any $g$. We let $\sigma^2_{\bar{\alpha}} \equiv \mathbb{E}\big((\bar{\alpha}_x - \mathbb{E}(\bar{\alpha}_x))^2\big)$ and $\sigma^2_{\bar{\psi}} \equiv \mathbb{E}\big((\bar{\psi}_y - \mathbb{E}(\bar{\psi}_y)))^2\big)$ denote the variance of $\alpha_{x_i}$ and $\psi_{y_j}$ across employed workers $i$ and filled jobs $j$. Also let $\text{cov}_{\bar{\alpha},\bar{\psi}} \equiv \mathbb{E}\big((\bar{\alpha}_x - \mathbb{E}(\bar{\alpha}_x))(\bar{\psi}_y - \mathbb{E}(\bar{\psi}_y))\big)$ denote the covariance between $\bar{\alpha}_{x_i}$ and $\bar{\psi}_{y_j}$ across matched pairs $(i, j)$ with $j = J_{i,t}$. The unidentified additive constant $k$ does not affect any of these moments.

If we estimate $\hat{\alpha}_i$ and $\hat{\psi}_j$ using OLS, the well-known issue of limited mobility (Andrews, Gill, Schank and Upward, 2008) biases estimates of the variances and covariances when $T$ is finite. Throughout this paper, we assume an idealized environment, either one where $T \to \infty$, or alternatively a statistical procedure that gives unbiased estimates of the variance and covariance, as proposed by Andrews, Gill, Schank and Upward (2008), Bonhomme, Lamadon and Manresa (2019), or Kline, Saggio and Sølvsten (2020).

Using the parametrization described in Section 8, we conduct the variance decomposition and report the results in Table 1. By construction, we exactly match the empirical moments.

# C Exponentially-Distributed Match Quality

We present results for the model with exponentially distributed match quality. We proceed as in the case of Pareto distribution. We set the number of types at $X = Y = 10$ and assume a uniform type distribution, $m_x = \frac{1}{X}$ and $n_y = \frac{1}{Y}$. The production function is

| | $Var(w)$ | $Var(\alpha)$ | $Var(\psi)$ | $2Cov(\alpha,\psi)$ | $Var(\varepsilon)$ | $Corr(\alpha,\psi)$ |
|---|---|---|---|---|---|---|
| BLM | 0.124 | 0.074 | 0.003 | 0.015 | 0.031 | 0.491 |
| model | 0.124 | 0.078 | 0.003 | 0.014 | 0.029 | 0.442 |

Table 2: Decomposition of variance of log wages in the model with exponentially distributed match quality shocks.

CES with elasticity of substitution $\xi$, $f_{x,y} = (0.5p_x^{\frac{\xi-1}{\xi}} + 0.5q_y^{\frac{\xi-1}{\xi}})^{\frac{\xi}{\xi-1}}$ where $p_1 = q_1 = 1$, $p_x = (1+\Delta_p)p_{x-1}$ for $x = 2, \dots X$, and $q_y = (1+\Delta_q)q_{y-1}$ for $y = 2, \dots Y$. We let $S(z) = e^{-z}$ because we normalize $\theta = 1$. We choose $\gamma = \frac{2}{3}$, and $r = 0.05$, $\delta = 0.5$, which implies expected duration of a match of 2 years. We choose the remaining parameters, $\Delta_p, \Delta_q, \rho, \xi$ to match the variance decomposition in Bonhomme, Lamadon and Manresa (2019). As before, we target the variance of worker type, firm types, the covariance and the variance of the residual, and we assume that we have infinitely much data, so we know each worker's type $x$ and each firm's type $y$ and so all distributions are deterministic functions of model parameters. The parameter values which allow us to match the target moments are $\Delta_p = 0.065$, $\Delta_q = 1.962$, $\xi = 0.694$ and $\rho = 2.9 \times 10^5$. Moments are summarized in Table 2.

This parametrization implies that the most productive worker is on average 59% more productive than the least productive worker, and the most productive firm is on average 443% more productive than the least productive one. In particular, $f_{Xy}/f_{1y} - 1$ lies between 0.30 and 0.75, depending on the value of $y$. For firms, $f_{xY}/f_{xy_1} - 1$ lies between 3.68 and 5.27, depending on value of $x$. The unemployment rate is 14.0%.

Figure 6 shows the results. In the first row we show the average log wage in matches and the distribution of worker types conditional on firm type. The log wage is monotone in firm type for each worker type, so again, one is tempted to conclude that low-type workers gain a lot from meeting a very productive firms. The top right panel shows that there is strong sorting, so low productive workers are only rarely employed by firms with high productivity. The bottom row shows that the reason behind this pattern is selection. While the bottom left figure shows a large increase in the mean surplus in matches as one moves from low to high productivity firm, the right panel of the bottom row shows that this is not the case for the average surplus in meetings.

# D On-the-Job Search Model

We show how to reformulate the model so that it is easier to solve, and then report some additional quantitative results.

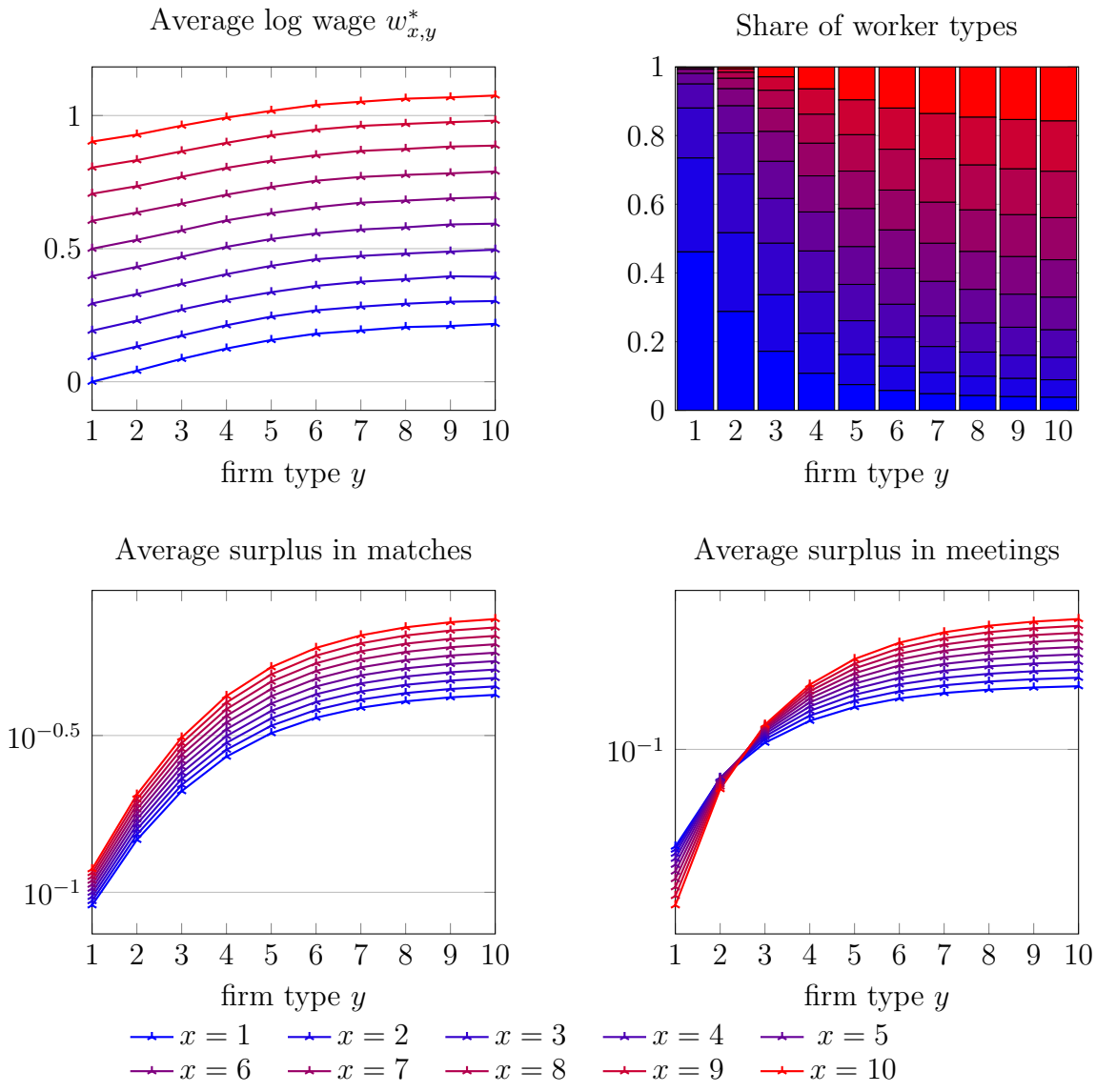



Figure 6: Results for model with exponentially distributed match quality shocks.

## D.1 Reformulation of the Model

Define $Z_{x,y}(\omega)$ to be the value of $z$ such that the match surplus equals $\omega$: $V^s_{x,y}(Z_{x,y}(\omega)) = \omega$. Then $z = Z_{x,y}(V^s_{x,y}(z))$ so $dz = Z'_{x,y}(\omega)d\omega$. We change variables in (25):

$$(r+\delta)\omega = Z_{x,y}(\omega)f_{x,y} - rV^u_x - rV^v_y + \rho_1 \sum_{y_1} v_{y_1} \int_\omega^\infty (\gamma\omega' - \omega)s(Z_{x,y_1}(\omega'))Z'_{x,y_1}(\omega')d\omega'. \quad (48)$$

We differentiate (48) with respect $\omega$ to obtain an ODE for $Z_{x,y}(\omega)$

$$r+\delta = Z'_{x,y}(\omega)f_{x,y} - \rho_1 \sum_{y_1} v_{y_1} S(Z_{x,y_1}(\omega')) + \rho_1(1-\gamma)\sum_{y_1} v_{y_1}\omega s(Z_{x,y_1}(\omega))Z'_{x,y_1}(\omega). \quad (49)$$

For large $\omega$, the on-the-job part is not important, giving terminal condition $\lim_{\omega\to\infty} Z_{x,y}(\omega) = \frac{(r+\delta)\omega + rV^u_x + rV^v_y}{f_{x,y}}$.

Next, define $\tilde{\phi}_{x,y}(\omega)$ be the density of the type $x$ workers matched with type $y$ firms with the match surplus $\omega$. Let $\tilde{\Phi}_x(\omega)$ be the measure of worker type $x$ who are employed in a match with surplus less than $\omega$, $\tilde{\Phi}_x(\omega) = \sum_{y=1}^Y \int_0^\omega \tilde{\phi}_{x,y}(\omega')d\omega'$. Then it holds,

$$\left(\delta + \rho_1 \sum_{y=1}^Y v_y S(Z_{x,y}(\omega))\right)\tilde{\Phi}_x(\omega) = \rho_0 u_x \sum_{y=1}^Y v_y \left(S(Z_{x,y}(0)) - S(Z_{x,y}(\omega))\right). \quad (50)$$

The outflow from $\tilde{\Phi}_x(\omega)$, the left-hand side of (50), is either due to exogenous separations or workers accepting offers with surplus higher than $\omega$. The inflow into $\tilde{\Phi}_X(\omega)$ is only from unemployment, and it is given by the rate at which workers receive offers with surplus between 0 and $\omega$.

It is now straightforward to characterize density $\phi_{x,y}(\omega)$:

$$\left(\delta + \rho_1 \sum_{y_1=1}^Y v_{y_1} S(Z_{x,y_1}(\omega))\right)\tilde{\phi}_{x,y}(\omega) = (\rho_0 u_x + \rho_1 \tilde{\Phi}_x(\omega))v_y s(Z_{x,y}(\omega))Z'_{x,y}(\omega). \quad (51)$$

The left-hand side shows the outflow from $\tilde{\phi}_{x,y}(\omega)$, which is either due to exogenous separations or workers moving to higher-surplus matches. The right-hand side shows the inflow. Workers come from unemployment by receiving an offer at the rate $\rho_0 u_x$, or from employment at the rate $\rho_1\tilde{\Phi}_x(\omega)$ since $\tilde{\Phi}_x(\omega)$ is the measure of workers with surplus below $\omega$.

Bonhomme, Lamadon and Manresa (2019), a worker is only included in the sample if they are employed by a single firm in all twelve months. Since the duration distribution of all jobs is the same in our model, none of these selection criterion affect the asymptotic values $\bar{\alpha}_x$ and $\bar{\psi}_y$.

The measures of type-$x$ unemployed and type-$y$ vacancies are

$$u_x = m_x - \sum_{y=1}^{Y} \int_0^\infty \tilde{\phi}_{x,y}(\omega) d\omega, \tag{52}$$

$$v_y = n_y - \sum_{x=1}^{X} \int_0^\infty \tilde{\phi}_{x,y}(\omega) d\omega. \tag{53}$$

Finally, we write equations (26) and (27) in terms of $Z_{x,y}(\omega)$:

$$rV_x^u = \rho_0 \sum_{y=1}^{Y} v_y \int_0^\infty \omega s(Z_{x,y}(\omega)) Z'_{x,y}(\omega) d\omega \tag{54}$$

$$rV_y^v = \rho_0 \sum_{x=1}^{X} u_x \int_0^\infty \omega s(Z_{x,y}(\omega)) Z'_{x,y}(\omega) d\omega + \rho_1 \sum_{x=1}^{X} \int_0^\infty (1-\gamma)\omega \tilde{\Phi}_x(\omega) s(Z_{x,y}(\omega)) Z'_{x,y}(\omega) d\omega. \tag{55}$$

To recover wages, we rewrite (31) in terms of $Z_{x,y}(\omega)$,

$$\tilde{W}_{x,y}(\omega) = rV_x^u + \gamma(r+\delta)\omega - \rho_1 \gamma \sum_{y'=1}^{Y} v_{y'} \int_\omega^\infty (\omega' - \omega) s(Z_{x,y}(\omega')) Z'_{x,y}(\omega') d\omega'. \tag{56}$$

To solve the model, we proceed as follows:

1. Given $V_x^u$, $V_y^v$, and $v_y$, we solve ODE (49) with the terminal condition $Z_{x,y}(\bar{\omega}) = \frac{(r+\delta)\bar{\omega} + rV_x^u + rV_y^v}{f_{x,y}}$ at large $\bar{\omega}$.
2. Given $Z_{x,y}(\omega)$, use equations (50)—(53) to update $u_x$ and $v_y$.
3. Finally, use (54) and (55) to update $V_x^u$ and $V_y^v$.
4. Repeat until the changes in $u_x$, $v_y$, $V_x^u$, and $V_y^v$ are small enough.
5. Lastly recover the wage using equation (56).

## D.2 Normalizations

A steady state equilibrium is described by a solution to equations (48) and (50)–(56). Here we show how to normalize the wage in both the Pareto and Exponential cases.

**Pareto Case**

Assume $S(z) = z^{-\theta}$. Consider an equilibrium $(V_x^u, V_y^v, Z_{x,y}(\omega), \tilde{\Phi}_x(\omega), \tilde{\phi}_{x,y}(\omega), u_x, v_y, \tilde{W}_{x,y}(\omega))$ given one value of the contact rates $\rho_0$ and $\rho_1$. Let $\hat{\rho}_0 = \lambda\rho_0$ and $\hat{\rho}_1 = \lambda\rho_1$ for some $\lambda > 0$. Then we claim that with the contact rate $\hat{\rho}$, there is an equilibrium as follows:

$$\begin{gathered}
\hat{V}_x^u = \lambda^{1/\theta} V_x^u, \quad \hat{V}_y^v = \lambda^{1/\theta} V_y^v \\
\hat{Z}_{x,y}(\lambda^{1/\theta}\omega) = \lambda^{1/\theta} Z_{x,y}(\omega), \quad \hat{\tilde{W}}_{x,y}(\lambda^{1/\theta}\omega) = \lambda^{1/\theta} \tilde{W}_{x,y}(\omega) \\
\hat{\tilde{\Phi}}_x(\lambda^{1/\theta}\omega) = \tilde{\Phi}_x(\omega) \\
\lambda^{1/\theta} \hat{\tilde{\phi}}_{x,y}(\lambda^{1/\theta}\omega) = \tilde{\phi}_{x,y}(\omega) \\
\hat{u}_x = u_x, \hat{v}_y = v_y.
\end{gathered}$$

The proof involves writing the eight equations in terms of the hatted variables evaluated at $\lambda^{1/\theta}\omega$ and simplifying. For, equation (48) becomes

$$\begin{aligned}
(r+\delta)\lambda^{1/\theta}\omega = \hat{Z}_{x,y}(\lambda^{1/\theta}\omega) f_{x,y} - r\hat{V}_x^u - r\hat{V}_y^v \\
+ \hat{\rho}_1 \sum_{y_1} v_{y_1} \int_{\lambda^{1/\theta}\omega}^{\infty} (\gamma\omega' - \lambda^{1/\theta}\omega)\theta \hat{Z}_{x,y_1}(\omega')^{-\theta-1} \hat{Z}'_{x,y_1}(\omega') d\omega'.
\end{aligned}$$

In the integral, do a change of variables to $\omega'' \equiv \lambda^{-1/\theta}\omega'$, so $d\omega' = \lambda^{1/\theta} d\omega''$. Also change the "hatted" variables to their unhatted counterparts using our conjectured functional forms

$$\begin{aligned}
(r+\delta)\lambda^{1/\theta}\omega = \lambda^{1/\theta} Z_{x,y}(\omega) f_{x,y} - \lambda^{1/\theta} r V_x^u - \lambda^{1/\theta} r V_y^v \\
+ \lambda\rho_1 \sum_{y_1} v_{y_1} \int_{\omega}^{\infty} (\gamma\lambda^{1/\theta}\omega'' - \lambda^{1/\theta}\omega)\theta \big(\lambda^{1/\theta} Z_{x,y_1}(\omega'')\big)^{-\theta-1} Z'_{x,y_1}(\omega'') \lambda^{1/\theta} d\omega''.
\end{aligned}$$

All the terms involving $\lambda$ cancel, yielding equation (48). We omit the other steps in the argument, which use an identical logic.

We now show that the measures of type-$x$ workers employed by type-$y$ firms is unaffected by the scale parameter in the meeting rate, $\lambda$:

$$\hat{\phi}_{x,y} \equiv \int_0^{\infty} \hat{\tilde{\phi}}_{x,y}(\omega) d\omega = \int_0^{\infty} \hat{\tilde{\phi}}_{x,y}(\lambda^{1/\theta}\omega')\lambda^{1/\theta} d\omega' = \int_0^{\infty} \tilde{\phi}_{x,y}(\omega') d\omega' \equiv \phi_{x,y}.$$

We first define $\hat{\phi}_{x,y}$. We then use a change of variables with $\lambda^{1/\theta}\omega' = \omega$. The third equation then replaces $\hat{\tilde{\phi}}$ with our expression for $\tilde{\phi}$. Thus the right hand side of Figures 1 and 2 are unaffected by the scale of the meeting rate.

We also calculate the average log wage that a type-$x$ worker receives when employed by a type-$y$ firm:

$$\hat{w}^*_{x,y} \equiv \frac{\int_0^\infty \log \hat{\tilde{W}}_{x,y}(\omega)\hat{\tilde{\phi}}_{x,y}(\omega)d\omega}{\int_0^\infty \hat{\tilde{\phi}}_{x,y}(\omega)d\omega} = \frac{\int_0^\infty \log \hat{\tilde{W}}_{x,y}(\lambda^{1/\theta}\omega')\hat{\tilde{\phi}}_{x,y}(\lambda^{1/\theta}\omega')\lambda^{1/\theta}d\omega'}{\int_0^\infty \hat{\tilde{\phi}}_{x,y}(\lambda^{1/\theta}\omega')\lambda^{1/\theta}d\omega'}$$
$$= \frac{\int_0^\infty \log\left(\lambda^{1/\theta}\tilde{W}_{x,y}(\omega')\right)\tilde{\phi}_{x,y}(\omega')d\omega'}{\int_0^\infty \tilde{\phi}_{x,y}(\omega')d\omega'} = \frac{1}{\theta}\log\lambda + w^*_{x,y}.$$

The steps are the same: define $\hat{w}^*$, do the same change of variables, replace 'hatted' variables by the 'unhatted' counterparts, and simplify. This shows how we choose $\lambda$ to target an average log wage of 0 in $(1,1)$ match, and that this simply scales all average log wages by the same additive constant.

**Exponential Case**

Let $(V^u_x, V^v_y, Z_{x,y}(\omega), \tilde{\Phi}_x(\omega), \tilde{\phi}_{x,y}(\omega), u_x, v_y, \tilde{W}_{x,y}(\omega))$ denote an equilibrium when the distribution of match-specific shocks is exponential distribution with mean 1. We construct an equilibrium with an exponential distribution with mean $\theta$, i.e. $S(z) = e^{-z/\theta}$, as follows:

$$\hat{V}^u_x = \theta V^u_x, \quad \hat{V}^v_y = \theta V^v_y$$
$$\hat{Z}_{x,y}(\theta\omega) = \theta Z_{x,y}(\omega), \quad \hat{\tilde{W}}_{x,y}(\theta\omega) = \theta\tilde{W}_{x,y}(\omega)$$
$$\hat{\tilde{\Phi}}_x(\theta\omega) = \tilde{\Phi}_x(\omega)$$
$$\theta\hat{\tilde{\phi}}_{x,y}(\theta\omega) = \tilde{\phi}_{x,y}(\omega)$$
$$\hat{u}_x = u_x, \hat{v}_y = v_y.$$

Again, we prove this by writing the eight equations (48) and (50)–(56) in terms of the hatted variables evaluated at $\theta\omega$ and simplifying. Start with equation (48):

$$(r+\delta)\theta\omega = \hat{Z}_{x,y}(\theta\omega)f_{x,y} - r\hat{V}^u_x - r\hat{V}^v_y + \rho_1\sum_{y_1} v_{y_1}\int_{\theta\omega}^\infty(\gamma\omega' - \theta\omega)\frac{1}{\theta}e^{-\hat{Z}_{x,y_1}(\omega')/\theta}\hat{Z}'_{x,y_1}(\omega')d\omega'.$$

In the integral, do a change of variables to $\omega'' = \omega'/\theta$, so $d\omega' = \theta d\omega''$. Also change the hatted variables to their unhatted counterparts:

$$(r+\delta)\theta\omega = \hat{Z}_{x,y}(\theta\omega)f_{x,y} - \theta rV^u_x - \theta rV^v_y + \rho_1\sum_{y_1} v_{y_1}\int_{\omega}^\infty(\gamma\theta\omega'' - \theta\omega)\frac{1}{\theta}e^{-Z_{x,y_1}(\omega'')}Z'_{x,y_1}(\omega'')\theta d\omega''.$$

| | $\Delta_p$ | $\Delta_q$ | $\xi$ | $\theta$ |
|---|---|---|---|---|
| $\rho_1 = 0$ | 0.077 | 0.854 | 0.713 | 7.146 |
| $\rho_1 = 0.1\rho_0$ | 0.097 | 0.340 | 0.710 | 9.089 |
| $\rho_1 = 0.2\rho_0$ | 0.072 | 0.217 | 0.688 | 16.261 |

Table 3: Calibrated parameters for the case $\rho_1 = 0$, $\rho_1 = 0.1\rho_0$ and $\rho_1 = 0.3\rho_0$.

All the terms involving $\theta$ cancel, yielding equation (48). We again omit the remaining steps in the argument, which use the same logic.

The proof that the measure of type-$x$ workers employed by type-$y$ firms is unaffected by the mean of the exponential is the same as in the Pareto case. For the average log wage, we also have follow similar steps as we did with the Pareto:

$$\begin{aligned}\hat{w}^*_{x,y} &\equiv \frac{\int_0^\infty \log \hat{\tilde{W}}_{x,y}(\omega)\hat{\tilde{\phi}}_{x,y}(\omega)d\omega}{\int_0^\infty \hat{\tilde{\phi}}_{x,y}(\omega)d\omega} = \frac{\int_0^\infty \log \hat{\tilde{W}}_{x,y}(\theta\omega')\hat{\tilde{\phi}}_{x,y}(\theta\omega')\theta d\omega'}{\int_0^\infty \hat{\tilde{\phi}}_{x,y}(\theta\omega')\theta d\omega'} \\ &= \frac{\int_0^\infty \log\left(\theta\tilde{W}_{x,y}(\omega')\right)\tilde{\phi}_{x,y}(\omega')d\omega'}{\int_0^\infty \tilde{\phi}_{x,y}(\omega')d\omega'} = \log\theta + w^*_{x,y}.\end{aligned}$$

This shows how we choose $\theta$ to target an average log wage of 0 in $(1,1)$ match, and that this simply scales all average log wages by the same additive constant.

## D.3 Quantitative Results $\rho_1 = 0.1\rho_0$

For comparison, we present results for the model with Pareto distributed match quality and $\rho_1 = 0.1\rho_0$. We re-calibrate the values of $\Delta_p, \Delta_q, \theta, \xi$ to match the variance decomposition presented in Bonhomme, Lamadon and Manresa (2019), and keep the other parameters the same as in the main text. Table 3 shows the parameter values with $\rho_1 = 0$ (no on-the-job search, Section 5), $\rho_1 = 0.1\rho_0$, and $\rho_1 = 0.2\rho_0$ (Section 6.2). As search on the job becomes more effective, the amount of firm heterogeneity required to match the variance of firm effects becomes smaller, with the parameter $\Delta_q$ decreasing significantly. More search while employed also generates more wage dispersion, and so in order to keep the variance of log wages and the variance of the error term equal to the targeted values, the parameter $\theta$ increases, which reduces the variance of the match-specific productivity shock.

Figure 7 shows the results with $\rho_1 = 0.1\rho_0$. In the first row we show the average log wage in matches and the distribution of worker types conditional on firm type. In both case, the log wage is monotone in firm type for each worker type, and submodular. The top right panels show that there is strong sorting, so low productive workers are only rarely employed

by firms with high productivity. The bottom rows of the figures show that the reason behind this pattern is selection. While the bottom left figures show a large increase in the average surplus in matches as one moves from low to high productivity firm, the right panel of the bottom row shows that this is not the case for the average surplus in meetings. In the model without on the job search, $\rho_1 = 0$, the average surplus in meetings is hump-shaped for low types, indicating that these workers benefit most from meeting a slightly better firm, in the case of $\rho_1 = 0.1\rho_0$, the average surplus in meetings is even declining for the lowest type. Hence, the low productivity workers benefits most from meeting the lowest-productivity firms, and benefit least from meeting firms with the highest productivity.

# E Event Study Methodology

For the event study methodology, we look at all workers who move from a type $y_1$ firm to a type $y_2$ firm at a point in time. Note that this means we weight workers by their likelihood of leaving their job at each instant. The average first log wage is

$$\bar{w}^1_{y_1,y_2} \equiv \frac{\sum_x \int_0^\infty \log W_{x,y_1}(\omega_1)\phi_{x,y_1}(\omega_1)\Big(\delta \frac{v_{y_2}S(Z_{x,y_2}(0))}{\sum_y v_y S(Z_{x,y}(0))} + \rho_1 v_{y_2} S(Z_{x,y_2}(\omega_1))\Big)d\omega_1}{\sum_x \int_0^\infty \phi_{x,y_1}(\omega_1)\Big(\delta \frac{v_{y_2}S(Z_{x,y_2}(0))}{\sum_y v_y S(Z_{x,y}(0))} + \rho_1 v_{y_2} S(Z_{x,y_2}(\omega_1))\Big)d\omega_1}.$$

There are $\phi_{x,y_1}(\omega_1)$ type $x$ workers in a match with type $y_1$ firms and surplus $\omega_1$. These matches end at each instant in two ways. They can end exogenously, at rate $\delta$, in which case the worker (eventually) moves to a type $y_2$ firm with probability $\frac{v_{y_2}S(Z_{x,y_2}(0))}{\sum_y v_y S(Z_{x,y}(0))}$. Or they end endogenously at rate $\rho_1 \sum_y v_y S(Z_{x,y}(\omega_1))$, in which case the worker moves to a type $y_2$ firm with probability $\frac{v_{y_2}S(Z_{x,y_2}(\omega_1))}{\sum_y v_y S(Z_{x,y}(\omega_1))}$.

Now look at the average second wage. Here we need to pay attention to the realized value of that job.

$$\bar{w}^2_{y_1,y_2} \equiv \frac{\sum_x \int_0^\infty \phi_{x,y_1}(\omega_1)\Big(\delta \mathbb{E}(\log W_{x,y_2}) + \rho_1 v_{y_2} \int_{\omega_1}^\infty \log W_{x,y_2}(\omega_2) s(Z_{x,y_2}(\omega_2)) Z'_{x,y_2}(\omega_2) d\omega_2\Big)d\omega_1}{\sum_x \int_0^\infty \phi_{x,y_1}(\omega_1)\Big(\delta \frac{v_{y_2}S(Z_{x,y_2}(0))}{\sum_y v_y S(Z_{x,y}(0))} + \rho_1 v_{y_2} S(Z_{x,y_2}(\omega_1))\Big)d\omega_1},$$

where $\mathbb{E}(\log W_{x,y_2})$ is the average wage that a type $x$ worker gets from a type $y$ firm out of unemployment, multiplied by the probability of matching with such a firm:

$$\mathbb{E}(\log W_{x,y_2}) \equiv \frac{v_{y_2} \int_0^\infty \log W_{x,y_2}(\omega_2) s(Z_{x,y_2}(\omega_2)) Z'_{x,y_2}(\omega_2) d\omega_2}{\sum_y v_y S(Z_{x,y}(0))}.$$

This is essentially the same expression, but in each event we record the new wage.

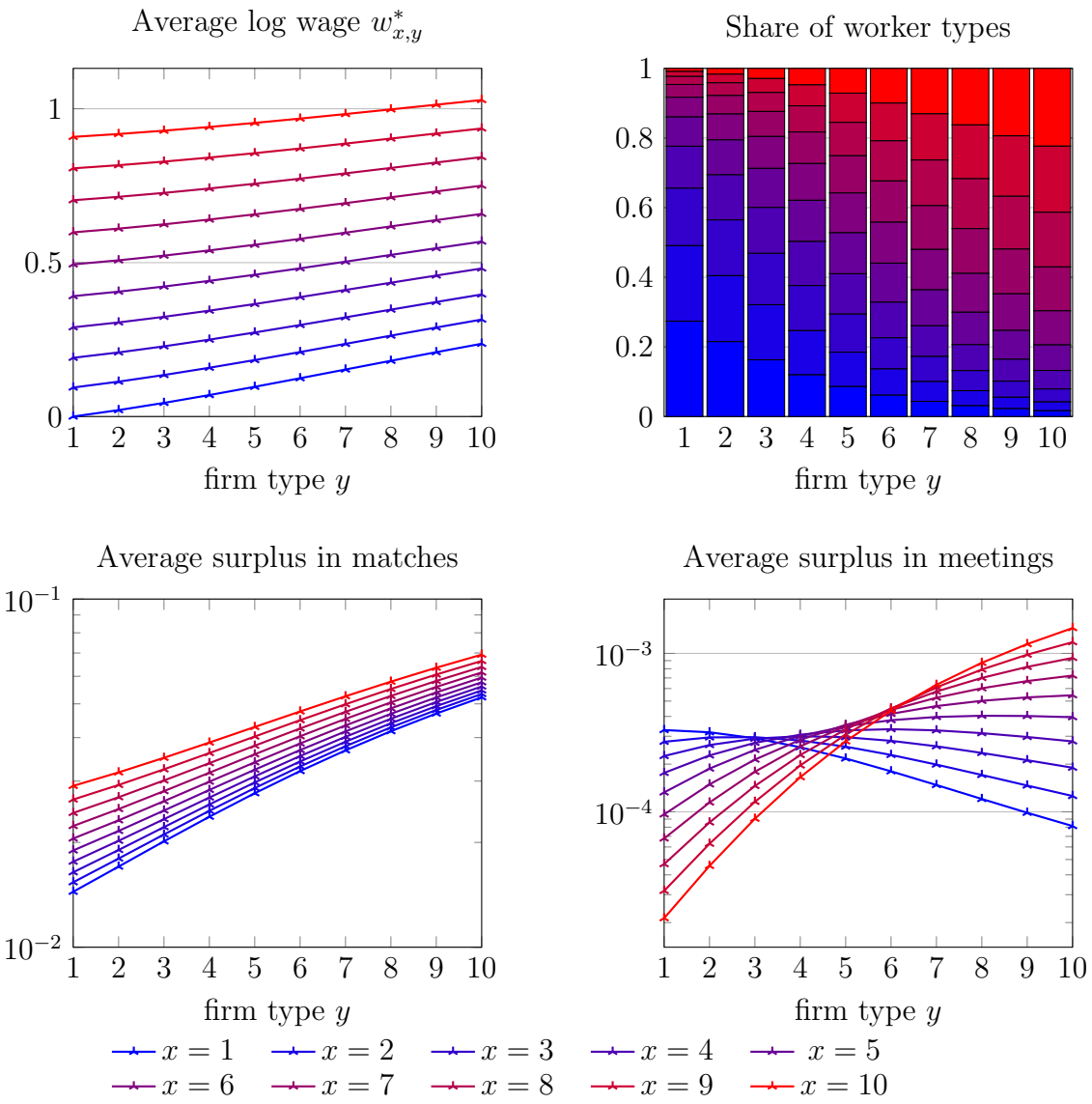



Figure 7: Results for model with Pareto distributed match quality shocks and $\rho_1 = 0.1\rho_0$. The model is re-calibrated to match the variance decomposition in Bonhomme, Lamadon and Manresa (2019).

To measure the average wage of workers who move from quantile $q_1$ to quantile $q_2$, we sum both the numerator and denominator of these two expressions across $y_1$ in quantile $q_1$ and $y_2$ in quantile $q_2$.

The expressions simplify considerably if there is no on-the-job search. Start with the first wage:

$$\bar{w}^1_{y_1,y_2} \equiv \frac{\sum_x \frac{v_{y_2} S(Z_{x,y_2}(0))}{\sum_y v_y S(Z_{x,y}(0))} \int_0^\infty \log W_{x,y_1}(\omega_1)\phi_{x,y_1}(\omega_1)d\omega_1}{\sum_x \frac{v_{y_2} S(Z_{x,y_2}(0))}{\sum_y v_y S(Z_{x,y}(0))} \int_0^\infty \phi_{x,y_1}(\omega_1)d\omega_1}.$$

Using equation (12) and (15), we can rewrite this as

$$\bar{w}^1_{y_1,y_2} \equiv \frac{\sum_x \frac{\phi_{x,y_1}\phi_{x,y_2}}{\sum_y \phi_{x,y}} w^e_{x,y_1}}{\sum_x \frac{\phi_{x,y_1}\phi_{x,y_2}}{\sum_y \phi_{x,y}}}.$$

This is indeed analogous to the expression in the paper for that case. The equation for the second wage without on-the-job search is analogous:

$$\bar{w}^2_{y_1,y_2} \equiv \frac{\sum_x \frac{\phi_{x,y_1}\phi_{x,y_2}}{\sum_y \phi_{x,y}} w^e_{x,y_2}}{\sum_x \frac{\phi_{x,y_1}\phi_{x,y_2}}{\sum_y \phi_{x,y}}}.$$