Economics Working Papers

2014-11

An Empirical Model of Wage Dispersion with Sorting

Jesper Bagger and Rasmus Lentz

AARHUS UNIVERSITY
BUSINESS AND SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS AND BUSINESS

An Empirical Model of Wage Dispersion with Sorting*

Jesper Bagger† Rasmus Lentz‡

March 28, 2014

Abstract

This paper studies wage dispersion in an equilibrium on-the-job-search model with endogenous search intensity. Workers differ in their permanent skill level and firms differ with respect to productivity. Positive (negative) sorting results if the match production function is supermodular (submodular). The model is estimated on Danish matched employer-employee data. We find evidence of positive assortative matching. In the estimated equilibrium match distribution, the correlation between worker skill and firm productivity is 0.12. The assortative matching has a substantial impact on wage dispersion. We decompose wage variation into four sources: Worker heterogeneity, firm heterogeneity, frictions, and sorting. Worker heterogeneity contributes 51% of the variation, firm heterogeneity contributes 11%, frictions 23%, and finally sorting contributes 15%. We measure the output loss due to mismatch by asking how much greater output would be if the estimated population of matches were perfectly positively assorted. In this case, output would increase by 7.7%.

Keywords: Sorting, Worker heterogeneity, Firm heterogeneity, On-the-job search, Wage dispersion, Matched employer-employee data

JEL codes: J24, J33, J62, J63, J64

*We are indebted to a great number of people, too many to mention all, who have commented on past versions of the paper and in seminar and conference presentations. In particular, we have benefited from the numerous discussions we have had with Rafael Lopes de Melo, Jan Eeckhout, Philipp Kircher, Jeremy Lise, Costas Meghir, and Jean-Marc Robin. In addition, John Kennan, Dale T. Mortensen, Giuseppe Moscarini, Richard Rogerson, Robert Shimer, and Chris Taber have given us useful ideas and comments. Henning Bunzel has been invaluable in the development of the data we use. We acknowledge financial support from LMDG and CAP, Aarhus University (LMDG is a Dale T. Mortensen Visiting Niels Bohr professorship project funded by the Danish National Research Foundation; CAP is a research unit funded by the Danish Council for Independent Research).

†Royal Holloway, University of London; LMDG; CAP; E-mail: jesper.bagger@rhul.ac.uk

‡University of Wisconsin-Madison; NBER; LMDG; CAP; E-mail: rlentz@ssc.wisc.edu

1 Introduction

What causes wages to vary across jobs? There is ample empirical evidence that worker skill and firm productivity heterogeneity are both important contributors to observed wage dispersion.[1] Labor market frictions allow firm productivity heterogeneity to manifest itself in wages and also provide a separate source of wage dispersion. Of particular interest in this respect is the observation that labor markets are characterized by a large amount of worker reallocation through job-to-job transitions, and that the transitions tend to be in the direction of higher wages.[2] Wages guide reallocation and are thus also in part determined by it.[3] Therefore, the study of wage dispersion should ideally contain an understanding of the allocation of workers to firms that the labor market is seeking to implement. In particular, this includes the issue of sorting. However, the impact of labor market sorting on wages is not well understood.

In this paper we quantify the sources of wage dispersion in an equilibrium on-the-job search model with firm and worker heterogeneity and the possibility of sorting. Heterogeneity is single dimensional: Workers differ in their skill, firms differ in the productivity with which they employ a given skill level worker. Wage determination is the same as in Dey and Flinn (2005) and Cahuc et al. (2006). A worker's current wage depends on her skill level, her employer's productivity as well as her bargaining position. The latter is shorthand for the competition that arises between two firms when a currently employed worker meets another firm. In a frictional labor market, meetings are chance events, and individual bargaining positions, and therefore wages, evolve stochastically, even among similar workers employed in similarly firms. Hence, labor market frictions contribute to wage variation, over and above the dispersion that arises via worker skill and firm productivity heterogeneity (Postel-Vinay and Robin, 2002).

Workers' search intensity is endogenous as in Christensen et al. (2005). The core sorting mechanism in the model is analyzed in detail in Lentz (2010). Depending on the presence of complementarities in the production function between worker skill and firm

[1]See among others Abowd et al. (1999), Postel-Vinay and Robin (2002), and Cahuc et al. (2006).

[2]See for example Fallick and Fleischman (2001), Christensen et al. (2005), Nagypál (2005) and Jolivet et al. (2006).

[3]See Burdett and Mortensen (1998), van den Berg and Ridder (1999), Bunzel et al. (2001), Postel-Vinay and Robin (2002), Christensen et al. (2005), Cahuc et al. (2006), Hornstein et al. (2011), and Bagger et al. (2014).

productivity, the worker's search intensity choice can vary with the skill level and sorting results. For example, if there are positive complementarities in production more skilled workers will search more intensely to move up the firm hierarchy and will tend to be matched with more productive firms, and positive assortative matching prevails. Sorting naturally adds to wage dispersion. The equilibrium wage distribution is affected by how the labor market combines given distributions of worker skill and firm productivity as well as the reallocation flows it produces in the process.

Using Danish matched employer-employee data, the match production function is estimated to be supermodular, and the equilibrium is characterized by positive assortative matching; more skilled workers tend to be matched with more productive firms. The correlation between worker skill and firm productivity is 0.12. The strength of the complementarities are relatively weak which is illustrated by the modest aggregate productivity gain of 7.7% that would result from a reshuffling of firms and workers in existing matches into a perfectly sorted, "no mismatch", allocation.

For the estimated model, we can decompose wage dispersion into four components: Worker heterogeneity, firm heterogeneity, labor market frictions, and sorting.[4] Worker heterogeneity is found to account for 51% of the wage variance in the data. Firm productivity heterogeneity is responsible for 11%. Frictions account for 23%. Finally, sorting contributes 15% of the wage variation.

As in the partnership models of Becker (1973) and Shimer and Smith (2000), our analysis links sorting in the match distribution primarily to match production function modularity. However, in our setup, due to an assumption of constant returns to scale in production and workers' opportunity to search on-the-job, neither firms nor workers view the match decision to include a substantial search opportunity loss. Hence, acceptance of a match opportunity does not require compensation for the loss of this opportunity. In contrast, the core of the acceptance/rejection decision in Shimer and Smith (2000) relies on a fundamental scarcity: Once matched, the agent gives up the opportunity to search until once again unmatched. This is not an altogether unreasonable assumption in the study of marriage, as in Becker (1973), but the vital role of on-the-job

[4] Abowd et al. (1999) consider the first two components in a reduced form regression framework. Postel-Vinay and Robin (2002) decompose dispersion into the first three components, subject to the assumption of no sorting in the match distribution. The match production function in Postel-Vinay and Robin (2002) is supermodular, but there is no mechanism in the model by which sorting can happen.

search for workers and the fact that a single firm can match with many workers make the scarcity assumption in labor market matching much less obvious.

Identification of sorting, i.e. recovering from observed data the relationship between unobserved worker skill and unobserved firm productivity, is inherently difficult, and is a central question in the paper. In this respect, it is joined by Eeckhout and Kircher (2011), Lise et al. (2013), Lopes de Melo (2008), and more recently by Bartolucci and Devicienti (2013) and Hagedorn et al. (2013). Like our paper, these papers also adopt a maintained identifying assumption regarding match opportunity scarcity, but they go to the other extreme benchmark and assume match opportunity scarcity at the job level as it exists in the partnership model. That is, once a job is filled, the matching opportunity capital that is embodied in the job is lost. Hence, the literature still lacks an actual identification of match opportunity scarcity at the firm level. Our identification strategy is not fully appropriate in the partnership model setting, nor are the identification strategies used for identification in the partnership model setting appropriate in ours. Taken together, all the mentioned papers produce a robustness of sorts to the issue of core identifying assumptions. Like ours, both Lopes de Melo (2008) and Lise et al. (2013) obtain results that suggest positive assortative matching in the labor market but with substantial imperfection. Results in Bartolucci and Devicienti (2013) that add profit data to the analysis, are also supportive of the presence of positive assortative matching.

Our framework thus emphasizes worker reallocation as the first order channel through which sorting happens, and our identification strategy concerning sorting is focused on worker reallocation rate heterogeneity. The sorting mechanism in the model implies that more mismatched workers are more likely to reallocate to another job. In a frictional labor market, in any given firm, there will be heterogeneity among its workers as to the degree of mismatch and consequently separation rate heterogeneity. The same is true among the population of unemployed workers; if some unemployed workers are more mismatched than others, they will leave unemployment faster, which leads to unemployment hazard heterogeneity. The identification of sorting in the paper utilizes the presence of heterogeneity in spell hazards and links it to worker and firm type rank. Firm type rank is identified using the composition of a firm's worker inflow, specifically what fraction of its hires that come directly from other firms relative to unemployment. We only identify worker type rank for a subpopulation of workers where the model predicts that wages can be used to rank individuals. This is the population of workers hired into top rank firms directly out of unemployment.

The paper is structured as follows: The model is presented in section 2, with its key properties discussed in section 3. Sections 4 and 5 present data and estimation, respectively. Section 5 is divided into a discussion of the identification strategy and the estimation which is done by Indirect Inference. Section 6 discusses the implied estimate for efficiency loss due to mismatch and in section 7 we decompose the estimated wage variance into its four distinct sources; worker heterogeneity, firm heterogeneity, friction, and sorting. Section 8 concludes.

2 Model

There is a continuum of firms with measure m, and a continuum of risk neutral workers with measure normalized at unity. Time is continuous and firms and workers discount time at a common rate r. Workers maximize income and firms maximize profits. A worker is characterized by his or her permanent innate skill level $h \in [0,1]$ which is independently and identically distributed across workers according to the cumulative distribution function $\Psi(\cdot)$. Firms differ with respect to their permanent productivity realization $p \in [0,1]$ which is independently and identically distributed across firms according to the cumulative distribution function $\Phi(\cdot)$.

Workers can be either employed or unemployed. Regardless of employment state, a worker generates outside employment opportunities through a choice of search intensity s at increasing and convex cost $c(s)$. The analysis allows that the search technology efficiency can differ across the two employment states. Specifically, a search intensity s results in the arrival rate of new job opportunities of $(\mu + \kappa s)\lambda(\theta)$ or $s\lambda(\theta)$ if unemployed or employed, respectively, where $\kappa > 0$. If $\kappa > 1$ then search is more efficient in the unemployed state. $\mu \geq 0$ represents an arrival of offers that is unrelated to the search decision of the worker. $\lambda(\theta)$ is the equilibrium arrival rate of offers per search unit and θ is the market tightness to be determined in equilibrium. By assumption $\lambda'(\theta) \geq 0$. For notational simplicity, we will often suppres θ in the expression for $\lambda(\theta)$.

A match between a worker of skill level h and a firm of productivity p produces output $f(h,p)$. It is assumed that $f(h,p)$ is twice continuously differentiable with

$f_p(h,p) \geq 0$ and $f_h(h,p) \geq 0$ for all (h,p). Hence, more skilled workers enjoy an absolute advantage relative to less skilled workers regardless of the firm type p they are matched with. Likewise for the ranking of firms.

Match separation occurs as the result of one of three mutually exclusive events. First, the worker in the match may receive an offer from an outside firm with greater productivity than the current firm which will induce a quit. Second, at rate $\delta_0\lambda$ the worker makes a job-to-job transition where the new job is drawn randomly from the vacancy offer distribution $\Gamma(\cdot)$ and the outside option in the new job is unemployment. The process is meant to capture the possibility that some job-to-job transitions are not up the firm hierarchy.[5] One possible explanation is that an (to the econometrician) unobserved shock has reduced the worker's valuation of the current match which induces a job-to-job transition. Nagypál (2005) provides an explicit argument for such a process. Alternatively, the worker may have to reallocate for family reasons. It may also be that the worker has been given notice of a lay-off sufficiently far in advance that the worker got a new job without an actual unemployment spell in between. The model does not take a stand on the nature of the shock. Third, at rate δ_1, the worker is laid off and moves into unemployment. The model allows the layoff rate to be worker type dependent. Specifically, the layoff rate can be high or low, $\delta_H > \delta_L > 0$. The layoff rate is modeled as a worker random effect such that the probability that a worker is a high layoff rate type is given by $\xi_j = \Pr\left(\delta_1 = \delta_j\right), j \in \{L, H\}$.

Employment contracts between workers and employers are set through a Rubinstein (1982) style bargaining game following the same protocol as in Cahuc et al. (2006). The exact protocol is described in detail in the online appendix. It is assumed that the worker can use a contact with one employer as a threat point in a bargaining game with another. The bargaining procedure is observationally equivalent to the one in Dey and Flinn (2005), which is generalized Nash bargaining with worker bargaining power β. Here, competition between firms is such that if two firms compete with each other over a worker, the more productive firm will win and the worker will bargain with full surplus extraction with the losing firm as the outside option. An employment contract can only be re-negotiated by mutual consent. If the worker is unemployed, then the

[5]Christensen et al. (2005), Nagypál (2005) and Bagger et al. (2014) emphasize that this type of separation shock is empirically important.

value of unemployment will be the worker's threat point.

As in Cahuc et al. (2006), denote by $V_j\left(h,q,p\right)$ a (j,h)-type worker's valuation of a job with a productivity p firm given an employment contract that the worker negotiated with an outside option of full surplus extraction with a productivity q firm. Furthermore, define $V_j\left(h,p\right) \equiv V_j\left(h,p,p\right)$. This is the full value of a match between a (j,h) worker and a productivity p firm. Since the firm is competing with an equally productive firm, the competition between the two results in the worker extracting all the rents from the relationship. The generalized Nash bargaining with worker bargaining power β then implies,

$$V_j\left(h,q,p\right) = \beta V_j\left(h,p\right) + \left(1-\beta\right) V_j\left(h,q\right). \tag{2.1}$$

The value of unemployment, $V_j^0\left(h\right)$, solves,

$$rV_j^0(h) = \max_{s\geq 0}\left\{ f(h,0) - c(s) + (\mu + \kappa s)\lambda\beta \int_{R_j(h)}^1 \left[V_j(h,p') - V_j^0(h)\right] d\Gamma(p') \right\}, \tag{2.2}$$

where $R_j(h)$ is the reservation productivity level. It is defined implicitly by,

$$V_j^0\left(h\right) = V_j\left(h, R_j(h)\right). \tag{2.3}$$

At rate $\lambda s\left(h,p\right)$ the worker meets an outside firm. If it is better than the current firm who has productivity p, the worker moves to the new firm and receives a contract with value $V_j\left(h,p\right)$.

An employment contract consists of a worker's wage level and search intensity (w,s). The wage profile is flat until it is renegotiated, which by assumption only happens if both parties agree to do so. The analysis assumes search intensities can be contracted upon, which results in the implementation of the jointly (between employer and employee) efficient search intensity level, a useful benchmark.[6] The outcome of the employment contract bargaining is such that the agreed upon search intensity maximizes the joint surplus of the match and the wage dictates the surplus split.

[6]Lentz (2014) studies the mechanism design problem in the case where search intensity is not contractable. Here, a flat wage profile that does not deliver the entire surplus to the worker results in the worker searching too much relative to the jointly efficient level because part of the incentive to generate outside offers now includes rent extraction from the current match.

The flow value equation for the value of employment can be written as,

$$
\begin{aligned}
rV_j(h,q,p) = w_j(h,q,p) - c(s_j(h,p)) + \lambda s_j(h,p) \int_p^1 \left[V_j(h,p,p') - V_j(h,q,p)\right] d\Gamma(p') \\
+ \lambda s_j(h,p) \int_q^p \left[V_j(h,q',p) - V_j(h,q,p)\right] d\Gamma(q') \\
+ \left[\delta_j + \delta_0 \lambda \Gamma\left(R_j(h)\right)\right] \left[V_j^0(h) - V_j(h,q,p)\right] \\
+ \delta_0 \lambda \int_{R_j(h)}^1 \left[V_j(h, R_j(h), p') - V_j(h,q,p)\right] d\Gamma(p'). \quad (2.4)
\end{aligned}
$$

If a worker meets an outside firm that has productivity greater than her current employer, $p' > p$, then the worker reallocates to the new employer with a contract that has value $V_j(h,p,p')$, reflecting her bargaining position of full surplus extraction with her old employer. If she meets a firm that improves on her bargaining position but is not more productive that her current firm, $q < p' \leq p$, then she stays with her current employer with a renegotiated contract that has value $V_j(h,p',p)$. The worker moves into unemployment either directly through a layoff (at rate δ_1) or by refusing an exogenous reallocation. If she accepts the exogenous reallocation shock with a productivity p' firm, her contract has value $V_j\left(h, R_j(h), p'\right)$ because her bargaining position is now unemployment.

The search intensity for a match between a (j,h)-type worker and productivity p firm, $s_j(h,p)$, is the level of search intensity that maximizes the right hand side of equation (2.4) for $q = p$, that is the total match value $V_j(h,p)$. Thus, it must satisfy the first order condition,

$$
c'(s(h,p)) = \beta\lambda \int_p^1 \left[V_j(h,p') - V_j(h,p)\right] d\Gamma(p'). \quad (2.5)
$$

Notice, that the jointly efficient level of search does not depend on the bargaining position of the worker, q.

The value function is characterized in Lemma 1

Lemma 1. *The worker's valuation of a match $V_j(h,q,p)$ is for any $j \in \{L, H\}$ strictly increasing in all three arguments, (h,q,p).*

Proof. See online appendix. □

2.1 The search choices

Integration by parts of equations (2.2) and (2.4) yields the following expressions for the first order conditions for the unemployed and employed search, respectively,

$$c'(s_j^0(h)) \quad = \quad \kappa\beta\lambda \int_{R_j(h)}^{1} \frac{f_p(h,p')\hat{\Gamma}\,(p')\,dp'}{r+\hat{\delta}_j+\beta\lambda s(h,p')\hat{\Gamma}(p')}, \tag{2.6}$$

and

$$c'\big(s_j(h,p)\big) = \beta\lambda \int_{p}^{1} \frac{f_p(h,p')\hat{\Gamma}\,(p')\,dp'}{r+\hat{\delta}_j+\beta\lambda s_j(h,p')\hat{\Gamma}(p')}, \tag{2.7}$$

where $\hat{\delta}_j = \delta_j + \delta_0\lambda$ and $\hat{\Gamma}\,(p) = 1 - \Gamma\,(p)$. By convexity of $c(\cdot)$, differentiation of equation (2.7) with respect to p immediately yields that $s_j(h,p)$ is monotonically decreasing in $p,\ \forall h$. Lemma 2 characterizes how search intensity varies across different skill workers depending on the kind of complementarity between skill and productivity in the production function. Specifically, if the production function is supermodular (submodular), then a high skill worker will search more (less) intensely for outside job opportunities than a less skilled colleague within a given firm. In the supermodular case, the relative wage gains from upward mobility are greater for high skill workers and consequently they invest more heavily in offer creation. In the submodular case, low skill workers have the greater gains and so search more intensely. If the production function is modular, then search intensity does not vary across worker skill.

Lemma 2. *For either $j = L, H$ and for any pair $(h_0, h_1) \in [0,1]^2$ such that $h_0 < h_1$, and for all $p \in [0,1)$, if $f_{hp} > 0$ (supermodular) then $s_j(h_0,p) < s_j(h_1,p)$. If $f_{hp} < 0$ (submodular) then $s_j(h_0,p) > s_j(h_1,p)$. If $f_{hp} = 0$ (modular) then $s_j(h_0,p) = s_j(h_1,p)$.*

The proof of Lemma 2 is a straightforward application of the employed search intensity first order conditions in equation (2.7). The reservation productivity $R_j(h)$ is characterized in Lemma 3

Lemma 3. *For any $h \in [0,1]$, if $\kappa \leq 1$ and $\mu = \delta_0$ then $R_j(h) = 0$, and if $\kappa > 1$ and $\mu > \delta_0$ then $1 > R_j(h) > 0$. Furthermore, if the production function is modular, then $R_j\,(h_0) = R_j\,(h_1)$ for any $(h_0, h_1) \in [0,1]^2$.*

In the case where $\kappa \leq 1$ and $\mu = \delta_0$, we obtain the trivial case where accepting a job does not involve a loss in search efficiency and consequently any job that is at least as productive as unemployment will be accepted by the unemployed worker. If employed search is less efficient than unemployed search, a sufficient condition for this is $\kappa > 1$ and $\mu > \delta_0$, a job that is exactly as productive as unemployment ($p = 0$) will not be acceptable to the worker since it can at most pay a wage stream equal to that of unemployment and the value of the search option is strictly below that of unemployment. The value of a job is monotonically increasing in the productivity of the job. Consequently, the problem is characterized by a threshold decision rule as to whether or not to accept an employment offer.

In the case where unemployed search is more efficient than employed search, $\kappa > 1$, an obvious question of interest is how $R_j(h)$ varies with h. Lemma 3 states that in the absence of production function complementarities, $R_j(h)$ is identical across worker skill levels. In this case, the gains to search are independent of the worker's own type, and the result follows from this insight. With complementarities, and $\kappa > 1$, the model includes many of the complications associated with the classic stopping problem as analyzed in Shimer and Smith (2000). In particular, it is possible that $R_j(h)$ is not monotone in h.

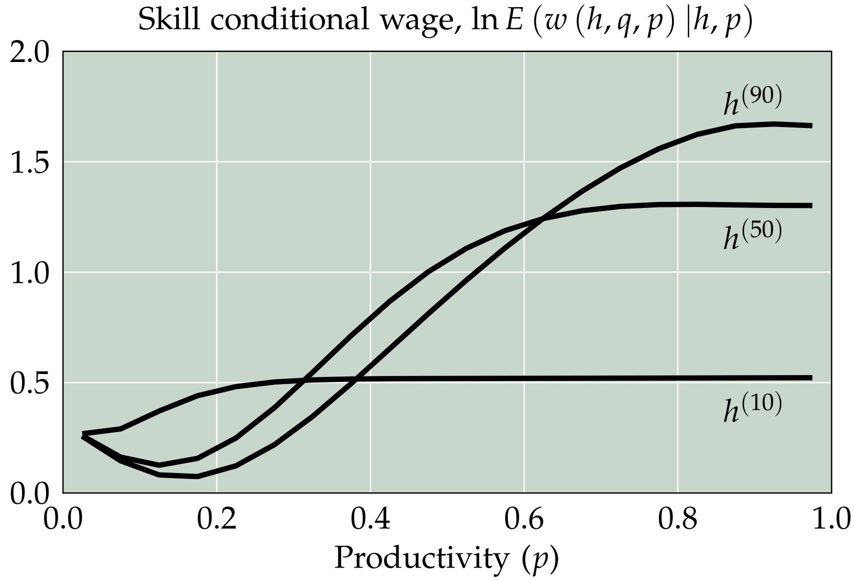
2.2 The wage equation

By equations (2.1) and (2.4) and integration by parts one obtains the following wage equation,

$$\begin{aligned} w_j(h,q,p) = {} & (r+\hat{\delta}_j)V_j(h,q,p) + c\big(s_j(h,p)\big) - \delta_j V_j^0(h) - \\ & \delta_0\lambda \int_{R_j(h)}^1 \frac{\beta f_p(h,p')\hat{\Gamma}(p')\,dp'}{r+\hat{\delta}_j+\beta\lambda s_j(h,p')\hat{\Gamma}(p')} - \lambda s_j(h,p)\left[\int_p^1 \frac{\beta f_p(h,p')\hat{\Gamma}(p')dp'}{r+\hat{\delta}_j+\beta\lambda s_j(h,p')\hat{\Gamma}(p')}\right. \\ & \left. + \int_q^p \frac{(1-\beta) f_p(h,p')\hat{\Gamma}(p')dp'}{r+\hat{\delta}_j+\beta\lambda s_j(h,p')\hat{\Gamma}(p')}\right]. \end{aligned} \tag{2.8}$$

It is a well-known feature in a model with sequential bargaining as in Postel-Vinay and Robin (2002) that wages are not necessarily monotone in the productivity of the firm. As it turns out, wages are also not necessarily monotone in the worker skill index, either. Figure 2.1 illustrates the average wage steady state wage for an (h,p) match, $E\left(w\left(h,q,p\right)|h,p\right)$, where model specification is given in detail in Section 5.1 and parameterization is given in the figure footnote.

Figure 2.1: Non-monotone wages



Note: The wage function is obtained for the parameterization, $(c_0, c_1, \rho, f_0, \alpha, \beta, \kappa, \delta, \delta_0, \underline{h}, \underline{p}) = (0.01, 1, -7, 5, 0.5, 0.1, 1, 0.1, 0.05, 0.25, 0.25)$. In addition, the vacancy distribution is assumed uniform and so is the worker skill distribution. $h^{(x)}$ indicates the xth percentile in the worker skill distribution $\Psi(h)$.

The figure draws the wage as a function of p for the 10th, 50th, and 90th worker skill percentile, denoted $h^{(10)}$, $h^{(50)}$ and $h^{(90)}$. In the example, the production function is supermodular and the worker's bargaining power, β, is relatively low. A worker of given skill level h may, on average, receive lower wages in more productive firms. For a given outside option, a more productive firm is more valuable to a worker because of it offers the possibility of more rent extraction in the event of future outside offers. Hence, bargaining with a more productive firm results in an initially lower wage. However, the realization of future higher wages may tend to take place with an even more productive firm, making it possible that some firms pay lower wages than their less productive peers. This is a feature of the wage determination mechanism that does not rely on the modularity of the production function.

The example in Figure 2.1 also illustrates that, in the less productive firms, the lowest skill worker may have the highest wage, and the highest skill worker the lowest wage. This complete reversal of the ranking of workers by wage does stem from supermodularity. The more skilled worker is expecting greater future wage gains relative to a less skilled worker, an effect that is amplified by the greater search effort among more skilled workers when production is supermodular. Consequently, for identical outside options the current wage is lower for the more skilled workers. Using wages across workers within a given firm to identify worker types is further complicated by the possibility that workers outside option q may vary systematically with worker skill type. In particular, in the case of a submodular production technology, low type workers search more intensely and accumulate a better bargaining position. Low skill workers may thus end up with higher wages within a firm.[7]

2.3 Vacancy creation

Permanent firm types p are distributed according to the cumulative distribution function $\Phi(p)$. A firm's total output Y is the sum of the output of all its matches. Hence, a firm with n workers, recorded in an n-vector $\mathbf{h}^n = (h_1, h_2, ..., h_n)'$, produces,

$$Y(\mathbf{h}^n, p) = \sum_{i=1}^{n} f(h_i, p).$$

The total wage bill of the firm depends not only on the vector of worker types, but also on the next best offer of each worker.

At any given time, each firm chooses a vacancy intensity ν at cost $c_\nu(\nu)$, where $c_\nu(\cdot)$ is strictly increasing and convex. Given the choice of vacancy intensity, the firm meets a new worker at rate $\eta\nu$. If a productivity p firm meets a skill h worker currently matched with a productivity $p' < p$ firm, the worker will accept to match with the productivity p firm. The bargaining will award value $V_j(h, p', p)$ to the worker and the firm will receive value $V_j(h, p, p) - V_j(h, p', p)$, which is the full match surplus minus the worker's share. The vacancy intensity choice is made so as to maximize the value of the firm's hiring

[7] Ranking workers by their wages within a given firm is not an option here. In the partnership model with Nash bargaining, wages are monotone in worker skill for a given firm type which both Hagedorn et al. (2013) and Bartolucci and Devicienti (2013) use explicitly to rank workers.

operation,

$$J_0(p) = \max_{\nu \geq 0}\left[-c_\nu\left(\nu\right) + \eta\nu \sum_{j\in\{L,H\}} \int_0^1 \left\{ \left[V_j\left(h,p\right) - V_j\left(h, R_j(h), p\right)\right] \Lambda_j^0\left(h\right) + \int_{R_j(h)}^p \left[V_j\left(h,p\right) - V_j\left(h,p',p\right)\right] \Lambda_j\left(h,p'\right) dp' \right\} dh\right], \quad (2.9)$$

Conditional on a meeting, $\Lambda_j\left(h,p\right)$ is the likelihood of meeting an employed skill level h, layoff rate δ_j worker who is currently employed with a productivity p firm. $\Lambda_j^0\left(h\right)$ is the likelihood that conditional on meeting a worker, the meeting is with a skill level h, layoff rate δ_j worker who is either currently unemployed or making a job-to-job reallocation, which in either case means that the worker's bargaining position is that of unemployment. They are complicated objects and we provide expressions for them in the online appendix. The density of matches between type-(j,h) workers and productivity p firms is given by $g_j(h,p) = \int_0^p g_j(h,q,p)dq$, where $g_j(h,q,p)$ is the joint PDF of matches. u_j is the layoff rate conditional unemployment rate and $\Upsilon_j(h)$ is the CDF of worker skill in the layoff rate conditional unemployment pool.

The firm's hiring intensity $\nu\left(p\right)$ is the maximand of the right hand side of equation (2.9). A firm's hiring rate is the product of the meeting rate and the probability that the worker in question accepts the firm's offer,

$$\eta(p) = \eta\nu(p) \sum_{j\in\{L,H\}} \int_0^1 I(R_j(h) \leq p) \left[\Lambda_j^0\left(h\right) + \int_{R_j(h)}^p \Lambda_j\left(h,p'\right) dp'\right] dh. \quad (2.10)$$

The expected match separation rate for a type p firm is given by,

$$d\left(p\right) = \sum_{j\in\{L,H\}} \xi_j\delta_j + \lambda\left(\theta\right)\hat{\Gamma}\left(p\right) \frac{\sum_{j\in\{L,H\}} \Delta_j \int_0^1 s_j\left(h,p\right) g_j\left(h,p\right) dh}{\sum_{j\in\{L,H\}} \Delta_j \int_0^1 g_j\left(h,p\right) dh}. \quad (2.11)$$

This expression comes from the result that a firm's expected labor force composition depends only on its productivity type, which is proven in the online appendix.

2.4 Steady state

Denote by $G_j(h,q,p)$, the fraction of employed workers with layoff rate δ_j and skill level no greater than h who are employed with firms of productivity no greater than p at bargaining position no greater than $q \leq p$. By definition, $G_j\left(1,1,1\right) = 1$. In steady state, the flow into this pool must equal the flow out, which leads to the following steady state condition on the match distribution,

$$\hat{\delta}_j G_j(h,q,p) = \int_0^h \mathbf{1}(R_j(h') \leq q)\lambda\left[\Gamma(p) - \Gamma(R_j(h'))\right]\Bigg[\frac{u_j}{1-u_j}[\mu + \kappa s_j^0(h')]v_j(h') + \delta_0\lambda \int_{R_j(h')}^1 \int_{q'}^1 g_j(h',q',p')dp'dq'\Bigg]dh' - \int_0^h \int_{R_j(h')}^q \lambda\Big\{\hat{\Gamma}(p)\int_{q'}^q s_j(h',p')dG_j(h',q',p') + \hat{\Gamma}(q)\int_q^p s_j(h',p')dG_j(h',q',p')\Big\} \quad (2.12)$$

where $\mathbf{1}(\cdot)$ is an indicator function that equals one if its expression is true, zero if false. A worker can leave the pool by moving into unemployment which happens in the case of a layoff at rate δ_j and in the case where exogenous reallocation is rejected by the worker which happens at rate $\delta_0\Gamma(R_j(h)$ for a type-(h,δ_j) worker. The worker can also leave the pool by receiving an outside offer: If a type-(h,δ_j) worker is currently employed at a productivity p' firm with bargaining position q' such that $q' \leq p' \leq q$, then an outside offer makes the worker leave the pool only if the productivity of the outside offer is greater then p. If the outside offer is in the $[q',p]$ range, then the offer changes the worker's employment terms (and the worker possibly changes firms), but the worker stays in the pool. If the worker is currently employed with a productivity $p' \in [q,p]$ firm, then an outside offer that is better than q makes the worker leave the pool, because it pushes the worker's new bargaining position above than q. The worker enters the pool through unemployment and exogenous reallocation if she were previously employed outside the pool by receiving an acceptable offer no greater than p and if her reservation level is no greater than q.

Equation (2.12) implies that the steady state unemployment rate for the population of layoff rate δ_j workers satisfies,

$$u_j = \left[\int_0^1 \left(1 + \frac{\hat{\Gamma}(R_j(h'))[\mu + \kappa s_j^0(h')]\lambda}{\delta_0\lambda\Gamma(R_j(h')) + \delta_j}\right) d\Upsilon_j(h')\right]^{-1}. \quad (2.13)$$

In steady state, the mass of productivity p firms with n workers $m_n(p)$ must be constant. Hence, the steady state firm size distribution satisfies,

$$0 = \eta\,(p)\,m_{n-1}\,(p) + d\,(p)\,(n+1)m_{n+1}\,(p) - (\eta\,(p) + d\,(p)\,n)\,m_n\,(p)\,, \quad (2.14)$$

for all $n \geq 1$ and p. In the online appendix we show that the firm's expected labor force composition is independent of its size. Hence, the expected destruction rate of matches is $d(p)$ for any firm size. Also, in steady state the number of firm births (firms enter with one worker) must equal the number of deaths,

$$\eta(p)m_0(p) = d(p)m_1(p). \tag{2.15}$$

An alternative interpretation of equation (2.15) is that firms do not exit, but they just have no economic activity during periods where they have no workers. During such periods they act like potential entrants. In the estimation we do not use entry and exit information from the data, and so, we do not have to take a stand on the issue. Furthermore, it is given that

$$\sum_{n=0}^{\infty} m_n(p) = m\phi(p), \tag{2.16}$$

where $\phi(p)$ is the firm productivity distribution PDF. Equations (2.14)-(2.16) imply that the type conditional firm size distribution $m_n(p)/(m\phi(p))$ is Poisson with arrival rate $\eta(p)/d(p)$,

$$m_n(p) = \left(\frac{\eta(p)}{d(p)}\right)^n \frac{1}{n!} \exp\left(-\frac{\eta(p)}{d(p)}\right) m\phi(p), \tag{2.17}$$

for all $n \geq 0$.

2.5 Steady state equilibrium

The equilibrium vacancy offer distribution is given by,

$$\Gamma(p) = \frac{\int_0^p \nu(p')\, d\Phi(p')}{\int_0^1 \nu(p')\, d\Phi(p')}. \tag{2.18}$$

In equilibrium, the meeting rates of both workers and firms must balance which implies,

$$\lambda(\theta) = \theta\eta(\theta), \tag{2.19}$$

where by proportional matching,

$$\theta = \frac{m \int_0^1 \nu(p')\, d\Phi(p')}{\sum_{j\in\{L,H\}} \xi_j \left[u_j \int_0^1 [\mu + \kappa s_j^0(h)] d\Upsilon_j(h) + (1-u_j) \int_0^1 \int_0^1 [\delta_0 + s_j(h,p)] dG_j(h,p)\right]}, \tag{2.20}$$

and

$$G_j(h,p) = G_j(h,p,p).$$

The worker skill distribution is related to the employment state conditional worker skill distributions by, $\Psi(h) = (1-u_j)G_j(h,\bar{p}) + u_j \Upsilon_j(h)$ which together with the steady state conditions on $G_j(h,q,p)$ and u_j produce (see detailed derivations in the online appendix),

$$\Upsilon_j(h) = \frac{\int_0^h \frac{\delta_0 \Gamma(R_j(h')) + \delta_j/\lambda(\theta)}{\delta_0 \Gamma(R_j(h')) + \delta_j/\lambda(\theta) + \hat{\Gamma}(R_j(h'))[\mu + \kappa s_j^0(h')]} d\Psi(h')}{\int_0^1 \frac{\delta_0 \Gamma(R_j(h')) + \delta_j/\lambda(\theta)}{\delta_0 \Gamma(R_j(h')) + \delta_j/\lambda(\theta) + \hat{\Gamma}(R_j(h'))[\mu + \kappa s_j^0(h')]} d\Psi(h')}, \; j \in \{L, H\}. \tag{2.21}$$

With these conditions, steady state equilibrium can be defined.

Definition 1. A steady state equilibrium is a collection $\{G_j(h,q,p), \Upsilon_j(h), \Gamma(p), u_j, s_j(h,p), s_j^0(h), R_j(h), \eta, w_j(h,q,p)\}_{j\in\{L,H\}}$ that satisfies equations (2.3), (2.6), (2.7), (2.8), (2.13), (2.12), (2.18), (2.20), and (2.21).

It is a convenient feature of the model that the composition of the unemployment pool can be expressed analytically as a function of the worker type distribution, a model fundamental, and the worker type conditional flow rates in and out of the unemployment pool. The search behavior that dictates the flow out of unemployment is in turn dictated by the offer distribution. The offer distribution is an equilibrium object defined in equation (2.18), which in turn is a simple transformation of firm type distribution $\Phi(\cdot)$, also a model fundamental. Existence and uniqueness of steady state equilibrium reduces to an examination of the existence of a fixed point of a mapping from the offer distribution into itself.

3 Properties of steady state equilibrium

The steady state equilibrium may or may not display sorting depending on the characteristics of the production function. In this section, we make the simplifying assumption that $\mu = \delta_0$. This assumption implies that the unemployed and employed states do not differ in terms of exogenous reallocation, which is helpful for the characterization of the worker reservation level. Our notion of sorting is focused on how worker skill and firm productivity are allocated in steady state equilibrium. Our worker types are two-dimensional so we discuss allocation patterns conditional on the worker's layoff rate type. Proposition 1 states sufficient conditions for positive sorting to occur. First, define the worker type conditional CDF of firm types by,

$$\Omega_j(p|h) = \frac{\int_0^p g_j(h,p')dp'}{\int_0^1 g_j(h,p')dp'}, \; j \in \{L,H\}. \tag{3.1}$$

One can then state the central characterization of sorting in steady state equilibrium.[8]

Proposition 1. *For any $h \in [0,1]$ and $j \in \{L,H\}$, $\Omega_j(0|h) = 0$ and $\Omega_j(1|h) = 1$. Consider any $j \in \{L,H\}$ and pair $(h_0,h_1) \in [0,1]^2$ such that $h_0 < h_1$. If $\kappa = 1$ then for all $p \in (0,1)$,*

- $f_{hp}(h,p) > 0 \,\forall (h,p) \Rightarrow \Omega_j(p|h_0) > \Omega_j(p|h_1)$ (supermodular).
- $f_{hp}(h,p) < 0 \,\forall (h,p) \Rightarrow \Omega_j(p|h_0) < \Omega_j(p|h_1)$ (submodular).
- $f_{hp}(h,p) = 0 \,\forall (h,p) \Rightarrow \Omega_j(p|h_0) = \Omega_j(p|h_1)$ (modular).

The result generalizes to any $\kappa > 0$ as long as $R_j(h)$ is weakly increasing (decreasing) in h when the production function is supermodular (submodular).

Proof. See Lentz (2010). □

If the production function is supermodular, it is for given populations of jobs and workers efficient to match high skill workers together with high productivity firms and low skill workers with low productivity firms. This is for example the case in Postel-Vinay and Robin (2002) and Cahuc et al. (2006). However there is no mechanism in these models to make the agents act on the gains to efficient matching and so these papers have no sorting.

The firms in our model are multi-worker constant returns to scale firms. They are not discriminating between which worker types they match with because hiring a worker at a given point in time does not preclude the firm from engaging in job recruitment in the future. The worker's job acceptance strategy is similarly trivial in a model where employed search is no less efficient than unemployed search. In this case, workers accept any match regardless of firm productivity as long as it is better than the current match.

[8] This proposition is given in Lentz (2010). We state it here for completeness.

The worker then continues to search for better opportunities from the new match. A mismatched worker has relatively larger gains to upward movement than a better matched worker for a given position in the firm hierarchy. The search intensity choice allows mismatched workers to act on the economic incentives in the model, and the more mismatched they are, the more intensely they will search for better opportunities. So, if the production function is supermodular more skilled workers will at any rung on the firm ladder search more intensely for outside options than a less skilled worker. Hence, the more skilled worker will in a stochastic dominance sense end up higher on the firm ladder than a less skilled worker. In the submodular case, it is the low skill workers that have larger relative gains to upward movement and so they search more intensely and end up higher on the firm ladder.[9]

As noted above the statements on sorting are within layoff rate types. As we go across layoff rate types it is possible that the elasticity of the search intensity with respect to a change in the layoff rate is not constant across worker skill levels, in which case layoff rate heterogeneity can drive a particular allocation pattern of worker skill to firm productivity in steady state. In our estimation, this point turns out to be moot because the high layoff rate type accounts for only about 1 percent of the employed workers.

4 Data

Our empirical analysis is conducted using Danish register-based matched employer-employee (MEE) panel dataset.

4.1 Data sources

The backbone of our data is individual level labor market spells recorded at a weekly frequency during 1985-2003 for the entire Danish population aged 15-70. Workers and firms are identified via unique IDs. Spells are constructed from administrative registers with information on public transfers, hourly wages, and start and end dates for all jobs reported by employers to tax authorities, and mandatory employer pension contributions.

[9]It is worth noting that the stochastic dominance results in Proposition 1 do not cleanly extend to the firm productivity conditional worker skill distribution.

The raw data identify five labor market states: employment (jobs), unemployment, retirement, self-employment and non-participation. By construction, non-participation is a residual state reflecting that an individual is neither employed nor self-employed nor receiving any kind of public transfer that would categorize him/her as unemployed or retired. Hence, in addition to genuine out-of-the-labor-force spells, non-participation captures imperfect take-up rates of public transfers, reception of transfers not used in the construction of the spell data and misreported start and end dates of spells.

Using person and firm IDs we merge the spells data with information on individual education and wages, and firm's sector of operation from IDA (Integreret Database for Arbejdsmarkedsforskning), an annual population-wide (age 15-70) Danish MEE panel constructed and maintained by Statistics Denmark from several administrative registers. Our wage measure is an estimate of the average hourly wage for jobs that are active in the last week of November. No wage information is available for job spells that do not overlap with a last week of November.

4.2 Analysis panel

A number of selection criteria and data manipulations are imposed in order to rid the data of invalid observations and to reduce un-modeled heterogeneity as well as other features of the data that our model is not designed to deal with.

First, we truncate individual labor market histories at age 55 and discard any labor market history that predates labor market entry as measured by date of graduation from highest completed education. Second, we discard all workers ever observed in employment in the public sector, in self-employment, in retirement or in Agriculture. Third, we recode non-participation spells as unemployment spells.[10] We recode unemployment spells with duration no greater than 13 weeks followed by recall of the worker back to the same employer as part of the original employment spell.[11] In addition, we recode unemployment spells of duration 2 weeks or less in between two employment spells with

[10]Our model features two labor market states and we must decide how to treat the empirical observation of nonparticipation in relation to the model's notion of unemployment. Coding nonparticipation as unemployment implies a broad definition of unemployment. This is appropriate as our model allows for layoff rate heterogeneity across workers. High layoff type workers will have weak labor force attachment, akin to non-particiaption.

[11]We effectively treat workers on recall unemployment as being employed. This is in line with recent evidence presented in Fujita and Moscarini (2014).

Table 1: Analysis data—summary statistics

	All years	1994	2003
Number of observations	6,815,884	658,465	703,707
Number of individuals	782,951	552,869	588,643
Number of job spells	1,698,990	490,309	511,604
Number of unemployment spells	608,065	168,155	192,102
Number of firms	117,847	53,537	58,210
Number of firm-years	559,920	53,537	58,249

different employers as a transition between two job spells within a single employment spell. Fourth, We select the period 1994-2003 for our analysis. Our structural model assumes permanent worker and firm types. We want to have a long enough panel to be able to effectively measure worker flows, but do not want to push the type permanence assumption too much. Fifth, we trim the annual individual hourly wage at the 1st and 99th percentiles, and trend them to 2003 levels using their implicit deflators.

Table 1 provides basic summary statistics on the final analysis panel and also shows statistics for the first (1994) and last (2003) annual cross section in the data.

5 Model estimation

5.1 Parameterization and estimation

We adopt a cost function specification where $c(\underline{s}) = c'(\underline{s}) = 0$ for some $\underline{s} \geq 0$. The worker's choice of offer arrival rate is in the range $s \in [\underline{s}, \infty[$. This is done to allow the possibility that worker search intensity is not an essential good in the creation of matches. The firm's recruitment cost function is $c_\nu(\nu)$ for recruitment intensity $\nu \in [0, \infty[$. The cost functions are given by increasing and convex functions,

$$c(s) = \frac{\left(c_0\left(s - \underline{s}\right)\right)^{1+1/c_1}}{1 + 1/c_1} \quad \text{and} \quad c_\nu(\nu) = \frac{\left(c_{\nu 0}\nu\right)^{1+1/c_{\nu 1}}}{1 + 1/c_{\nu 1}}. \tag{5.1}$$

where $c_0 > 0$ and $c_{\nu 0} > 0$ are scale parameters and $c_1 > 0$ and $c_{\nu 1} > 0$ set curvatures.

The match production function specified as,

$$f(h, p) = f_0\Big(\alpha\left(h + \underline{h}\right)^\rho + (1 - \alpha)\left(p + \underline{p}\right)^\rho\Big)^{\frac{1}{\rho}}, \tag{5.2}$$

where f_0 is a scale parameter, and $\alpha \in [0, 1]$ sets the weight that is put on the skill index relative to the firm productivity index. $\underline{h}$ and $\underline{p}$ are lower support bounds that are relevant only because of our parametric specification of the h and p distributions. The modularity of the CES function is governed by ρ. If $\rho < 1$, then the production function is supermodular. It is submodular for $\rho > 1$, and it is modular for $\rho = 1$.

We parameterize the firm productivity distribution, $\Phi(p)$, as a Beta distribution with parameters $\left(\beta_0^\Phi, \beta_1^\Phi\right)$ and the worker skill distribution is assumed to be a Beta distribution with parameters $\left(\beta_0^\Psi, \beta_1^\Psi\right)$.[12] We allow for classical measurement errors ε^w in annual individual wage observations, with $\varepsilon^w \backsim \mathcal{N}\left(0, \sigma_w^2\right)$.

The discount rate is fixed at $r = 0.05$ and equilibrium market tightness is normalized at $\lambda(\theta) = 1$. Furthermore, c_0 and c_{v0} are not separately identified so we normalize $c_{v0} = 1$. Finally, we set $\underline{h} = \underline{p} = 0.1$ to avoid extreme behavior in cases where the production function treats h and p as essential goods. This leaves us with 18 free structural parameters $\boldsymbol{\omega} = \left(c_0, c_1, \kappa, c_{v1}, \delta_0, \delta_L, \delta_H, \xi_L, \beta_0^\Psi, \beta_1^\Psi, \beta_0^\Phi, \beta_1^\Phi, \beta, f_0, \alpha, \rho, \sigma_w, \underline{s}\right)'$, which we estimate by Indirect Inference (Gourieroux et al., 1993).

The Indirect Inference estimator is

$$\widehat{\boldsymbol{\omega}} = \arg\min_{\omega} [\mathbf{a}(\boldsymbol{\omega}_0) - \mathbf{a}^S(\boldsymbol{\omega})]' \widehat{\mathbf{W}}^{-1} [\mathbf{a}(\boldsymbol{\omega}_0) - \mathbf{a}^S(\boldsymbol{\omega})],$$

where $\mathbf{a}(\boldsymbol{\omega}_0)$ is a vector of specific auxiliary statistics and data moments computed on real data, a function the true parameter value $\boldsymbol{\omega}_0$, $\mathbf{a}^S(\boldsymbol{\omega}) = \frac{1}{S}\sum_{s=1}^S \mathbf{a}_s(\boldsymbol{\omega})$ is the same vector, but computed on S simulated datasets from the structural model at some parameter value $\boldsymbol{\omega}$, and $\widehat{\mathbf{W}}$ is an estimate of the variance-covariance matrix of $\mathbf{a}(\boldsymbol{\omega}_0)$. For suitable choices of $\mathbf{a}$, and under some regularity conditions, see Gourieroux et al. (1993), $\sqrt{N}(\widehat{\boldsymbol{\omega}} - \boldsymbol{\omega}_0) \to^d \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega})$ where N is the number of observations in the data, $\boldsymbol{\Omega} = (1 + S^{-1})(\mathbf{J}'\widehat{\mathbf{W}}^{-1}\mathbf{J})^{-1}$, and $\mathbf{J} = \partial \mathbf{a}(\boldsymbol{\omega})/\partial \boldsymbol{\omega}$. We report standard errors from an estimate of $\boldsymbol{\Omega}$.[13]

[12] This means that we are solving for the equilibrium fixed point vacancy offer distribution $\Gamma(p)$ in each simulation iteration.

[13] We compute $\widehat{\mathbf{W}}$ by block-bootstrapping worker careers in the real data (200 bootstrap repetitions). $\mathbf{J}$ is estimated by numerically differentiating the simulated vector of moments $\mathbf{a}^S(\boldsymbol{\omega})$ with respect to $\boldsymbol{\omega}$ (at $\boldsymbol{\omega} = \widehat{\boldsymbol{\omega}}$). When the model is simulated, we replicate that there are 8.84 workers per firm in the data, which directly determines the size of the firm population for a given number of simulated workers. The estimate is obtained by simulating economies with a worker population of 100,000 over 10 years, $S = 288$ times.

5.2 Identification

Identification of sorting requires auxiliary models or data moments that rank workers and firms in terms of their unobserved skills and productivities h and p. Such statistics are difficult to find. Given a matched employer-employee dataset, one might think that worker and firm fixed effects from a log wage regression would provide such rankings. However, in general, wage data alone does not suffice for identification of sorting. Indeed, as pointed out in section 2.2, as well as in Eeckhout and Kircher (2011), Lopes de Melo (2008), and Lise et al. (2013), because wages are not monotone in the productivity indices, worker and firm fixed effects obtained from wage regressions do not correctly rank workers and firms in terms skill levels and productivities, h and p.[14] The implication is that the correlation between estimated worker and firm fixed effects does not identify the correlation between worker skill and firm productivity indices in the steady state match distribution.

In Appendix A we document how the wage non-monotonicity bias the correlation between worker and firm fixed effects downward relative to the correlation between h and p. This suggests that the fixed effects correlation provides a lower bound on the true degree of assortative matching between productivity types in the equilibrium.[15]

At this point it is worth emphasizing that non-monotonicity extends to a measure like the firm's measured labor productivity, $Y\left(\mathbf{h}^n, p\right)/n$, because of the way that the expected labor force composition changes across firms with different productivity levels. It is a feature of the model that the equilibrium firm productivity conditional worker skill distribution

$$\Omega_j(h|p) = \frac{\int_0^h g_j(h', p)dh'}{\int_0^1 g_j(h', p)dh'}, \quad j \in \{L, H\}$$

is not necessarily stochastically increasing (decreasing) in p when the production function is supermodular (submodular); see Lentz (2010). Hence, firm level output data does not solve the firm rank identification problem.

Instead of a relying on wage regressions or output data for identification, we utilize that in our model sorting is fundamentally driven by reallocation rate heterogeneity

[14] Of course, the mechanisms by which non-monotonicity arise are different in the partnership model relative to the one in this paper.

[15] For $\beta = 1$ one obtains trivially that $w\left(h, q, p\right) = f\left(h, p\right)$, which by construction is monotone, and indeed wages are fully reflective of the match production function.

across different types of workers. Our identification strategy consequently puts considerable emphasis on the reallocation aspect of the data.

In what follows, we detail the key parts of the identification strategy. First, we discuss how we rank firms based on worker reallocation, and subsequently proceed to discuss the various auxiliary statistics and data moments we use in our Indirect Inference procedure.

5.2.1 Firm ranking

We can identify a firm's position within the hierarchy by measuring the origin composition of its worker inflow. Specifically, the fraction of the worker inflow that comes directly from other firms as opposed to from the unemployment pool. Assuming $\kappa \leq 1$, by the proportional matching assumption, the probability that a hire comes directly from another firm is given by

$$\iota(p) = \frac{\sum\limits_{j \in \{L,H\}} \xi_j (1-u_j) \int_0^1 \left[\delta_0 g_j(h) + \int_0^p s(h,p') g_j(h,p') dp'\right] dh}{\sum\limits_{j \in \{L,H\}} \xi_j (1-u_j) \int_0^1 \left[\frac{u_j}{1-u_j}[\mu + s_j^0(h)] v_j(h) + \delta_0 g_j(h) + \int_0^p s_j(h,p') g_j(h,p') dp'\right] dh},$$

where u_j is the layoff rate conditional unemployment rate (see (2.13)), $g_j(h) = \int_0^1 g_j(h,p)dp$, and from which it trivially follows that $\iota_p(p) \geq 0$ with strict inequality if, for any $j \in \{L,H\}$, $\int_0^1 s_j(h,p) g_j(h,p) dh > 0$. Hence, the measure of the fraction of the inflow that comes directly from other firms is monotonically increasing in the productivity index of the firm.[16]

In terms of empirical implementation, for all firms in our analysis data, we collect all hires they make during the 10 year window and calculate the fraction that come directly from other firms.[17] The inflow rank is measured only for firms with a total inflow of more than 15 hires, and at least one hire from unemployment; effectively truncating observations on very small firms for statistics involving firm rank measures. Given the prevalence of exogenous job-to-job reallocation in the data, the noise in the inflow

[16] In the case where $\kappa > 1$, the match acceptance decision, as represented by the productivity threshold $R_j(h)$, must be taken into account. Monotonicity does not necessarily hold in cases where $R_j(h)$ varies across h and j. The case $\kappa < 1$ turns out to be the empirically relevant one.

[17] We replicate the same selection in our model simulation, as well as the average firm size in the data so as to emulate the noise in the firm rank measure that is related to small numbers.

measure is considerable for very small firms, and consequently, the conditioning on extreme realizations (either low or high) of the inflow measure over selects the very small firms. The inflow measure of large firms is a less noisy reflection of their underlying propensity to hire from other firms. In the following we refer to a firm's measured inflow rank as $\hat{\iota} \in [0, 1]$. It is the firm's percentile position in the inflow measure distribution.

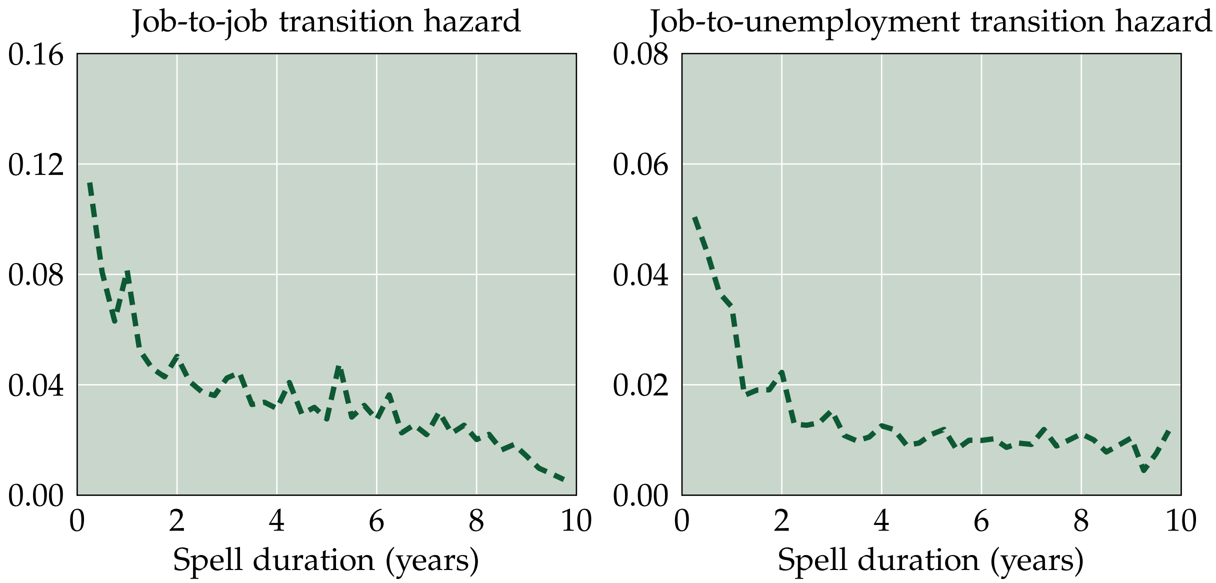
We next turn to a description of the auxiliary statistics and data moments we base our estimation on. These fall in three broad categories, namely statistics related to worker reallocation, cross sectional heterogeneity, and unemployment duration and starting wages.

5.2.2 Worker reallocation

From the analysis panel we extract a flow and a stock dataset. For the flow data, we select all employment spells not initiated in the final year of our data period with non-missing inflow rank measure $\hat{\iota}$. The unit of observation in the flow data is a spell, and the data is $\{t_i, ee_i, eu_i, \hat{\iota}_i\}$ for $i = 1, 2, ..., I$, where t_i is spell duration, ee_i is a job-to-job transition indicator, and eu_i is a job-to-unemployment indicator. The stock data contains a sequence of annually stock sampled employment spells. The unit of observation is a given spell in a given year. Let $\mathcal{I}(n)$ be the index set of employment spells active in the last week of November in year n. The stock data is $\{ee_{it}, eu_{it}, m_{it}, n\}$ for $n = 1, 2..., 10$ and $i \in \mathcal{I}(n)$, where m_{in} is the number of jobs in the employment cycle that spell i is part of, a statistic that is easy to compute from the analysis panel.

Kaplan-Meier job hazard functions: We include Kaplan-Meier estimates of the unconditional job-to-job and job-to-unemployment transition hazard functions from the flow data. Heterogeneity in search intensities exhibits itself through negative duration dependence in the Kaplan-Meier job-to-job hazard function. A similar dynamic selection on layoff rate types produce negative duration dependence in the estimated job-to-unemployment hazard function. We measure the hazard functions at a quarterly frequency over a 10 year period. Let $\mathcal{R}(t^*)$ be the set of spells at risk of ending within quarter $t^* = 1, 2, ..., 40$, and $\mathcal{E}(t^*)$ and $\mathcal{U}(t^*)$ the sets of spells that end in a job-to-job, respectively job-to-unemployment, transition within quarter t^*. The Kaplan-Meier esti-

Figure 5.1: Quarterly Kaplan-Meier job hazards.



mates of the two quarterly hazard functions are

$$h^{ee}(t^*) = \frac{|\mathcal{E}(t^*)|}{|\mathcal{R}(t^*)|} \quad \text{and} \quad h^{eu}(t^*) = \frac{|\mathcal{U}(t^*)|}{|\mathcal{R}(t^*)|},$$

for $t^* = 1, 2, ..., 40$.

Figure 5.1 plots $h^{ee}(t^*)$ and $h^{eu}(t^*)$; both empirical Kaplan-Meier hazards exhibit clear negative duration dependence.

Inflow rank conditional job-to-job transition hazard rates: To directly discipline the model estimate to fit the empirical firm ladder, we include as an auxiliary statistic to be matched the job-to-job transition hazard rate as function of inflow rank $\hat{\imath}$, computed on the flow data. Let d_i^{ee} and d_i^{eu} be binary indicators for spell i ending in a job-to-job, respectively job-to-unemployment, transition within a quarter of its initiation. We estimate $\widehat{P}^{ee}(\hat{\imath}) = \Pr(\widehat{d_i^{ee} = 1}|\hat{\imath})$ and $\widehat{P}^{eu}(\hat{\imath}) = \Pr(\widehat{d_i^{eu} = 1}|\hat{\imath})$ by non-parametrically regressing d_i^{ee} and d_i^{eu} on $\hat{\imath}_i$. We transform the estimated job-to-job transition probabilities into hazard rates, thus including in the estimation

$$h^{ee}(\hat{\imath}) = -\frac{\widehat{P}^{ee}(\hat{\imath}) \ln\left\{1 - \widehat{P}^{ee}(\hat{\imath}) - \widehat{P}^{eu}(\hat{\imath})\right\}}{\widehat{P}^{ee}(\hat{\imath}) + \widehat{P}^{eu}(\hat{\imath})}$$

evaluated at ten equidistant values in $\hat{\imath} \in [0, 1]$.

Figure 5.2: Quarterly inflow rank conditional job-to-job hazard rates

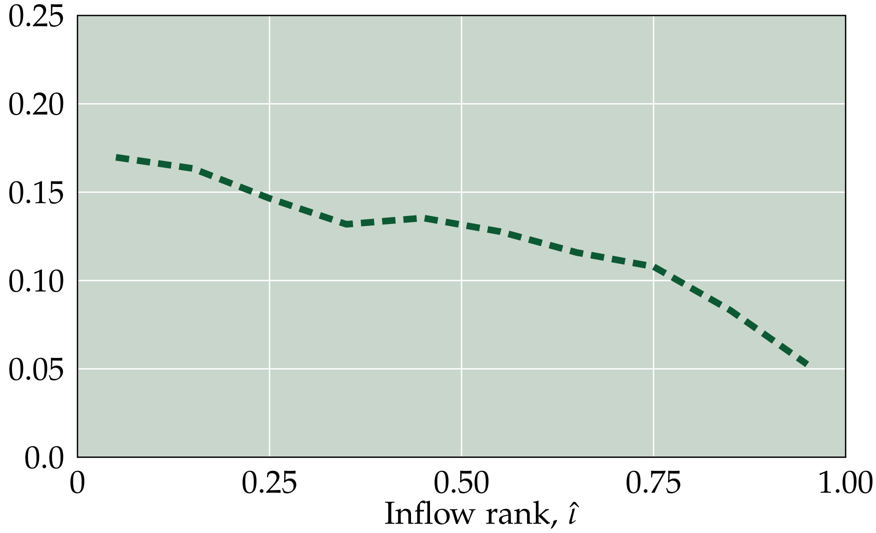


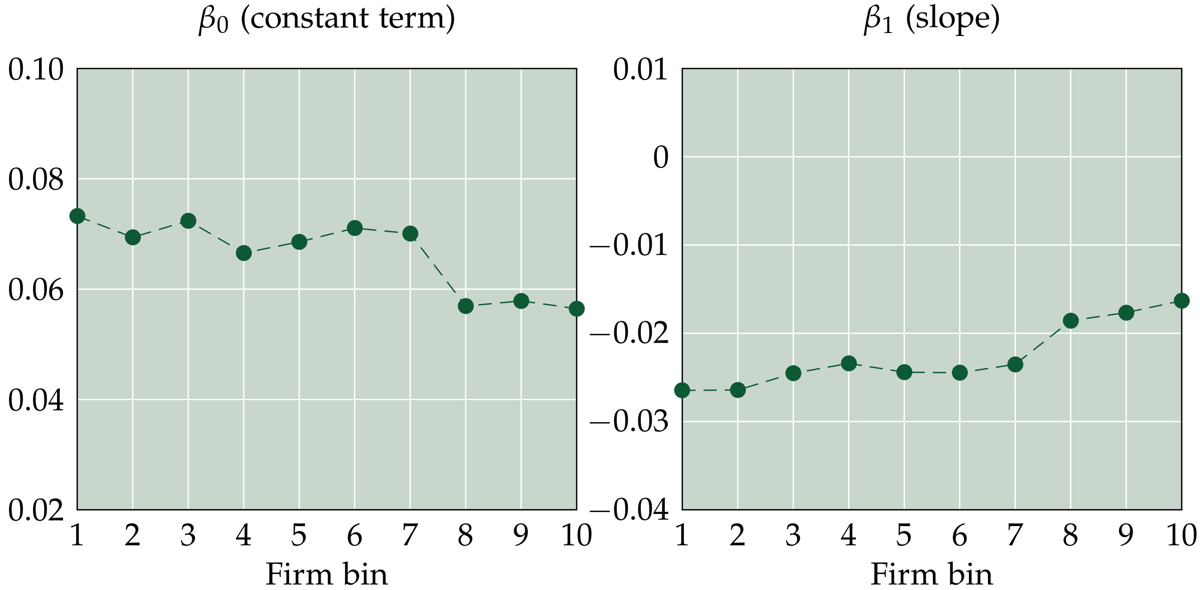
Figure 5.2 plots $h^{ee}(\hat{\imath})$ and shows a clear negative relationship between the EE-transition rate and firm rank with workers in less productive firms transitioning to other firms at greater rates than workers that are placed with more productive firms.[18] Furthermore, at the top rank firms, search intensity is near zero, and the observed job-to-job transition rate at these firms provide information on δ_0, the exogenous job-to-job reallocation rate.

Duration dependence and the firm ladder: If search intensity heterogeneity varies over the firm ladder, as would be the case if there is sorting, we should observe differences in the duration dependence of job-to-job transition hazards across the firm ladder. To utilize any such variation in the estimation, we split the flow data job spells into ten bins according to the deciles of the firm-level distribution of $\hat{\imath}$. Within each bin we compute the Kaplan-Meier estimate of the quarterly job-to-job transition hazard function, denoted $h^{ee}_k(t^*)$ for bin $k = 1, 2, ..., 10$, where $t^* = 1, 2, ..., 40$ index quarters over our 10 year window of observation. We now consider $\widehat{\beta}_{0k}$ and $\widehat{\beta}_{1k}$ estimated from the regression

$$\ln h^{ee}_k(t^*) = \beta_{0k} + \beta_{1k} \ln t^* + \epsilon_{t^*k}, \tag{5.3}$$

[18] It is important to note that because of noise in the firm rank measure one would expect substantial attenuation bias if this is to be interpreted as actual firm ladder conditional job-to-job hazards.

Figure 5.3: Duration dependence and the firm ladder



which are rendered graphically in Figure 5.3 for bins $k = 1, 2, ..., 10$. The left panel, presenting $\widehat{\beta}_{0k}$, shows a declining pattern across the deciles of the inflow rank distribution. The right panel, showing $\widehat{\beta}_{k1}$, trace out an increasing-towards-zero profile. That is, the duration dependence of the job-to-job transition hazard function is weakens towards the top of the firm ladder, suggesting that search intensity variation declines towards the top of the firm ladder.

The observed negative duration dependence of the Kaplan-Meier EE hazards within firm type will require the model to produce variation in the search intensity choices across worker within these bins of roughly identical types of firms. Part of this variation can come from measurement noise in the firm's rank and our simulation procedure will replicate this mechanism. But, of course, by introducing complementarities in production, the model can induce search choice variation in worker skill for worker within identical type firms and thereby seek to match the negative duration dependence this way.

Employment cycles: In the stock data, within a given cross section n, each ongoing job $i \in \mathcal{I}(n)$ is part of an employment cycle, a sequence of employment spells with no intervening unemployment spells. The number of jobs in each of these employment cycles is m_{in}. Let $\overline{m}_n = \frac{1}{|\mathcal{I}(n)|}\sum_{i\in\mathcal{I}(n)} m_{in}$ be the average number of employment jobs in employment cycles ongoing in cross section n, and $\text{se}(m)_n = \left[\frac{1}{|\mathcal{I}(n)|}\sum_{i\in\mathcal{I}(n)}(m_{in} - \overline{m}_n)^2\right]^{1/2}$ the standard error. We include $\overline{m} = \frac{1}{10}\sum_{n=1}^{10}\overline{m}_n$ and $\text{se}(m) = \frac{1}{10}\sum_{n=1}^{10}\text{se}(m)_n$ in the set of moments to be matched. In the data, $\overline{m} = 2.182$ and $\text{se}(m) = 1.541$. These statistics are a particular aggregation of the job-to-job hazards relative to the layoff hazards, but measured in the stock data, and thu provide additional identification of these rates.

Record statistics: To provide additional information on the exogenous reallocation rate δ_0 we consider the probability that a randomly stock sampled job spell ends in a layoff. In a constant offer arrival rate, on-the-job search model without exogenous job-to-job reallocation, Barlevy and Nagaraja (2013) show that the statistic is bounded below by 1/2.[19] From the stock data, we know the exit transition type (job-to-job or job-to-unemployment) of all jobs ongoing at cross section n. $\mathcal{I}(n)$ is the index set of employment spells in cross section n, and now let $\mathcal{U}(n)$ be the set of spells that end in a job-to-unemployment transition. We include in the vector of moments to be fitted in the estimation $\overline{S}^{eu} = \frac{1}{10}\sum_{n=1}^{10}|\mathcal{U}(n)| \,/\, |\mathcal{I}(n)|$, the average share of matches in a cross section that ends in a job-to-unemployment transition. Empirically, $\overline{S}^{eu} = 0.338$, inconsistent with a pure on-the-job search model where all job-to-job transitions are from lower to higher ranked firms.

5.2.3 Cross section heterogeneity

The analysis data provides annual measurements on workers' wages. Wage measurements are available for all jobs ongoing in the last week of November. Hence, slightly changing the notation, $\mathcal{I}(n)$ is now the index set of workers with employment spells ongoing in the last week of November in year n. For the cross section heterogeneity moments we extract $\{w_{in}, ueinit_{in}, i, \text{K}(i,n)\}$ for $n = 1, 2, ..., 10$ and $i \in \mathcal{I}(n)$. Here, $ueinit_{in}$ indicates whether or not w_{in} is the first wage observation in a job initiated via a

[19]Hence, we take information about δ_0 both from this record type statistics as well as from the job-to-job transition hazard at top rank firms as described above. The use of the latter statistic to identify δ_0 is possibly sensitive to the assumption that all workers agree on the ranking of firms. This is for example not the case in the partnership sorting model in Shimer and Smith (2000) if there are complementarities in the production function. However, the use of the Barlevy and Nagaraja (2013) statistic to inform δ_0 is robust to this issue since it is primarily a reflection of the relative rates by which workers move up their respective ladders.

unemployment-to-job transition, and $\mathrm{K}(i,n)$ indicates the employer of worker i in cross section n. That is, $\mathrm{K}(i,n) = k$ if worker i is employed by firm k in cross section n.

Raw moments: We include the first and second moment of the empirical distribution of log wages. Let $\overline{\ln w}_n = \frac{1}{|\mathcal{I}(n)|}\sum_{i\in\mathcal{I}(n)} \ln w_{in}$ and $\mathrm{se}(\ln w)_n = \frac{1}{|\mathcal{I}(n)|}\sum_{i\in\mathcal{I}(n)}(\ln w_{in} - \overline{\ln w}_n)^2$. We include $\overline{\ln w} = \frac{1}{10}\sum_{n=1}^{10} \overline{\ln w}_n$ and $\mathrm{se}(\ln w) = \frac{1}{10}\sum_{n=1}^{10} \mathrm{se}(\ln w)_n$ as moments to be matched in the estimation. To capture the firm ladder effect on wages, we also include the first and second moments of the empirical distribution of starting wages in jobs initiated from unemployment, denoted $\overline{\ln w^{ueinit}}$ and $\mathrm{se}(\ln w^{ueinit})$. Empirically, $\overline{\ln w} = 5.254$ and $\overline{\ln w^{ueinit}} = 5.167$, and $\mathrm{se}(\ln w) = 0.168$ and $\mathrm{se}(\ln w^{ueinit}) = 0.222$. These measurements are consistent with a job ladder model where on-the-job search implies $\overline{\ln w} > \overline{\ln w^{ueinit}}$ and $\mathrm{se}(\ln w) < \mathrm{se}(\ln w^{ueinit})$.

To discipline the estimated model in relation to firm sizes, we include average firm size, as well as the ratio of firms to workers. Let K_n be the number of firms with employees in cross section n. For each firm $k = 1, 2, ..., K_n$ in November cross section n, let N_{jn} be a simple count of the number of employed workers. Average firm size is computed as $\overline{N} = \frac{1}{10}\sum_{n=1}^{10} \frac{1}{K_n}\sum_{k=1}^{K_n} N_{kn}$. We also include the ratio of firms to the workforce as a moment to be matched. Let N_n^* be the size of the workforce in cross section n, i.e. $N_n^* = \sum_{k=1}^{K_n} N_{kn} + U_n$ where U_n is the number of unemployed workers. Then we include $\frac{1}{10}\sum_{n=1}^{10} K_n/N_n^*$ as a target in the estimation. We find $\overline{N} = 8.646$ and $\frac{1}{10}\sum_{n=1}^{10} K_n/N_n^* = 0.091$.

The model links search behavior and within-job wage growth. The quantitative effect of on-the-job search on within-job wage growth depends on workers bargaining power parameter. We add average annual within-job wage growth to the vector of auxiliary statistics. Let $\Delta \ln w_{in} = \ln w_{in} - \ln w_{i,n-1}$ be the year n year-on-year wage growth for individual i. Our vector of moments to be fitted includes

$$\overline{\Delta_{\text{Within}} \ln w} = \frac{\sum_{n=2}^{N}\sum_{i\in\mathcal{I}(n)} \Delta \ln w_{in} 1_{\{\mathrm{K}(i,n)=\mathrm{K}(i,n-1)\}}}{\sum_{n=2}^{N}\sum_{i\in\mathcal{I}(n)} 1_{\{\mathrm{K}(i,n)=\mathrm{K}(i,n-1)\}}}.$$

The data reveals $\overline{\Delta_{\text{Within}} \ln w} = 0.009$.

Mean-min ratio: Hornstein et al. (2011) propose the mean-min ratio as a useful and parsimonious measure of wage dispersion, and argue that a basic wage search model without worker heterogeneity cannot generate enough wage dispersion as reflected in the mean-min ratio. However, the wage process we employ may in fact produce too much dispersion as measured by the mean-min ratio because initial wages can be very low. The extent to which this occurs is driven by workers' bargaining power parameter β, and the mean-min ratio thus serves to disciplineβ in the estimation.[20] The data used for computing the mean-min ratio is the same as that used for estimation of the auxiliary log wage regression. We estimate the minimum wage as the average wage among the lower 5 percentiles in the wage distribution. Denote the estimate minimum wage by $\underline{w}$, and the mean wage by $\overline{w}$. Then we include $Mm = \overline{w}/\underline{w}$ in the vector of auxiliary statistics. The empirical mean-min ratio is 1.854.

Log wage regression: We include a restricted version of the Abowd et al. (1999) log wage regression in our set of auxiliary models. Specifically, consider the following log wage regression

$$\ln w_{in} = \varphi_0 + \chi_i + \varphi_{\mathrm{K}(i,n)} + \epsilon_{in} \tag{5.4}$$

where χ_i is a worker effect, $\varphi_{\mathrm{K}(i,n)}$ is a firm effect, and ε_{in} is a residual. When estimating (5.4) we impose the following restrictions:

$$E\left[\chi_i \epsilon_{in}\right] = 0, \qquad E\left[\varphi_{\mathrm{K}(i,n)} \epsilon_{in}\right] = 0, \tag{5.5}$$

and

$$\mathrm{E}\left[\chi_i \varphi_{\mathrm{K}(i,n)}\right] = 0 \tag{5.6}$$

The first two assumptions in (5.5) impose "exogenous mobility" in the terminology of Abowd et al. (1999), allowing for estimation of the parameters in (5.4), including the fixed effects, by OLS. The third restriction (5.6), not imposed in Abowd et al. (1999), implies uncorrelated firm and worker effects, as estimated from (5.4). Let K, I and N be the total number of firms, workers and observations, respectively. We estimate (5.4) by OLS and include the average firm effect $\overline{\varphi} = \widehat{\varphi}_0 + \frac{1}{N}\sum_{i=1}^{I}\sum_{n=1}^{10}\widehat{\varphi}_{\mathrm{K}(i,n)}$ and their standard deviation $\mathrm{se}(\varphi) = \left[\frac{1}{N}\sum_{i=1}^{I}\sum_{n=1}^{10}(\widehat{\varphi}_{\mathrm{K}(i,n)} - \overline{\varphi})^2\right]^{1/2}$, as well as the standard deviations of the estimated worker effects and the residuals, $se(\chi) = \left[\frac{1}{N}\sum_{i=1}^{I}\sum_{n=1}^{10}(\widehat{\chi}_i - \overline{\chi})^2\right]^{1/2}$

[20] In the extreme, if $\beta = 1$ the worker simply gets $w\,(h,p) = f\,(h,p)$. For lower bargaining power, the initial wage in an employment relationship will be reduced by the expectation of future wage gains.

and $\text{se}(\epsilon) = \left[\frac{1}{N}\sum_{i=1}^{I}\sum_{n=1}^{10}(\widehat{\epsilon}_{in} - \overline{\epsilon})^2\right]^{1/2}$. Besides $\overline{\varphi} = 5.238$, we find $\text{se}(\varphi) = 0.179$, $\text{se}(\chi) = 0.218$ and $\text{se}(\epsilon) = 0.134$. The auxiliary log wage regression implies a variance decompositions of $Var(\ln w_{in})$ where firm effects account for 33%, worker effects for 49%, and residual variation for 18%.

5.2.4 Unemployment duration and starting wages

Consider the initial wages of workers that are hired out of unemployment into top rank firms, i.e. firms of type $p = 1$. For $\kappa \leq 1$ and with $\underline{s} = 0$ such wages are direct reflections of the model object,

$$w(h, 0, 1) = (1 - \beta) r V_0(h) + \beta f(h, 1). \tag{5.7}$$

These particular wages are monotonically increasing in h. Observed initial wages within this group of workers therefore provide a ranking of these workers by skill level. With a supermodular production technology, $\partial s_j^0(h) / \partial h > 0$; that is, unemployed search intensity is increasing in h. Absent any other sorting forces, supermodular production functions induce a negative correlation between unemployment duration and subsequent initial wages among workers hired into top rank firms.[21] If sorting is negative the correlation has the opposite sign.

Layoff rate heterogeneity affects this statistic. A high layoff rate implies reduced incentives to search, a low value of unemployed search, and thus, lower wages due to a reduced bargaining position. Layoff rate heterogeneity manifest itself as a spurious negative correlation between unemployment duration and initial wages out of unemployment into top ranked firms. Since the identification of layoff rate heterogeneity comes from other sources, in particular duration dependence in the job-to-unemployment hazard function, this does not present an identification problem. Moreover, to anticipate some of our empirical results, the high layoff type workers are hardly ever employed, and the confounding effect of layoff heterogeneity on labor market sorting is minimal.

Exogenous search $\underline{s} > 0$ introduces an option value to employment in the most productive firms where the efficient contract otherwise is designed to eliminate on-the-job search. This implies that initial wages in the top ranked firms may not be monotone in

[21]If $\kappa > 1$, unemployed search is more efficient than employed search. With a supermodular production function, high skill workers may now reject offers from the bottom of the firm ladder and unemployment durations may be increasing in the worker skill type. Empirically we find $\kappa \leq 1$.

worker skill level h. As a result, the correlation between unemployment duration and initial wages in top ranked firms may formally no longer correctly reflect modularity of the production function. In practice, we found that this bias is minimal, and are not concerned that exogenous search $\underline{s}$ is impacting our conclusions.

For each worker in the analysis panel ever observed in unemployment we compute the individual average unemployment duration. We can use our inflow rank measure $\hat{\iota}$ to identify top rank firms. Specifically, we take firms in the 95th percentile or above in the distribution of $\hat{\iota}$ as "top ranked". We then extract all unemployment-to-job transitions into these firms (index these observation by $i = 1, 2, ..., I$) and, for each transition i, record the individual average unemployment duration, denoted $\overline{t^u}_i$, and the starting wage, denoted w_i^0.[22] The resulting dataset is $\{\overline{t^u}_i, w_i^0\}$ for $i = 1, 2, ..., I$, and the moment of interest is

$$Corr(\bar{t}^u, w^0 | \hat{\iota} \geq 0.95) = \frac{\frac{1}{I}\sum_{i=1}^{I}(\overline{t^u}_i - \overline{t^u}) \cdot (w_i^0 - \overline{w^0})}{\text{se}(\overline{t^u}) \cdot \text{se}(w^0)},$$

where $\overline{t^u} = \frac{1}{I}\sum_{i=1}^{I}\overline{t^u}_i$, $\text{se}(\overline{t^u}) = \left[\frac{1}{I}\sum_{i=1}^{I}(\overline{t^u}_i - \overline{t^u})^2\right]^{1/2}$, $\overline{w^0} = \frac{1}{I}\sum_{i=1}^{I} w_i^0$ and $\text{se}(w^0) = \left[\frac{1}{I}\sum_{i=1}^{I}(w_i^0 - \overline{w^0})^2\right]^{1/2}$ are the mean and standard errors of unemployment duration and starting wage, respectively. We include also these statistics in the estimation. Empirically, we find $\overline{t^u} = 57$ weeks with $\text{se}(\overline{t^u}) = 70$ weeks, $\overline{w^0} = 186$DKK with $\text{se}(\overline{w^0}) = 62$DKK, and finally $Corr(\bar{t}^u, w^0 | \hat{\iota} \geq 0.95) = -0.168$. This moment points towards positive sorting.

5.3 Model estimate

The estimated structural parameters are presented in Table 2. The structural parameters are all precisely estimated. With an annual layoff rate of $\delta_L = 0.063$, the expected duration between unemployment spells for the low layoff rate type worker is 16 years. The estimated rate for the high layoff type is $\delta_H = 1.905$, leading to an expected duration between unemployment spells of about 5 months. Workers improve their bargaining position by accumulating job offers, but this search capital depreciates completely at entry into unemployment. Hence, the returns to search are substantially lower for the high layoff type workers. In the population, high layoff types account for 14 percent, but are

[22] We measure the starting wage as the wage on record at the first post-transition November cross section date. For the estimation, we reproduce this observation scheme in the simulations.

almost completely absent from the steady state population of employed workers, where they account for only 1%.

The estimated search cost function $c(s)$ implies an elasticity of search cost to search effort of 1.077. The recruitment cost function $c_\nu(\nu)$ is highly convex, with an elasticity with respect to recruitment intensity at 84.333. The elasticity is driven up to ensure that the more productive firms do not take over the market, a result that would be at odds with the data. The endogenous arrival rate of job offers is supplemented in two ways. First, job offers arrive at an annual exogenous rate $\underline{s} = 0.034$, independent of search effort. Second, workers are hit by reallocation shocks at an annual rate $\delta_0 = 0.106$.

Our estimate of $\kappa = 0.845$, below unity, implies employed job search is slightly more efficient than unemployed job search. With $\kappa < 1$, unemployed workers accept any job offer they receive, independent of their skill level h, and sorting through differential reservation productivities does not arise.

The estimated parameters of the firm productivity and worker skill CDFs, $\Phi(p)$ and $\Psi(h)$, are difficult to interpret. The estimated $\Phi(p)$ has mean 0.274, standard deviation 0.193, and is right skewed and excess kurtosis with skewness 0.740 and kurtosis 3.598. The estimated $\Psi(h)$ has mean 0.141 and standard deviation 0.079. $\Psi(h)$ is right skewed and excess kurtosis with skewness 0.884 and kurtosis 4.077. The left panel of Figure 5.4 plots the estimated heterogeneity distributions $\Phi(p)$ and $\Psi(h)$.

With respect to the CES match production function, we focus attention on ρ, which governs modularity. At an estimate of $\rho = -2.045$ the estimated production function is supermodular. Hence, conditional on layoff type and employer productivity p, more skilled workers search more intensely than less skilled workers. The resulting equilibrium match distribution is such that skilled workers tend to match with more productive firms. As can be seen from the right panel of Figure 5.4, among the low layoff type workers, making up 99% of the employed workers, the equilibrium distribution of firm productivity p for high skilled workers stochastically dominates that of low skilled workers. For low layoff type workers, the estimated correlation between worker skill and firm productivity in the steady state equilibrium is 0.12. Given the absence of high layoff type workers among employed workers, the unconditional (on layoff type) steady state correlation also comes out at 0.12.

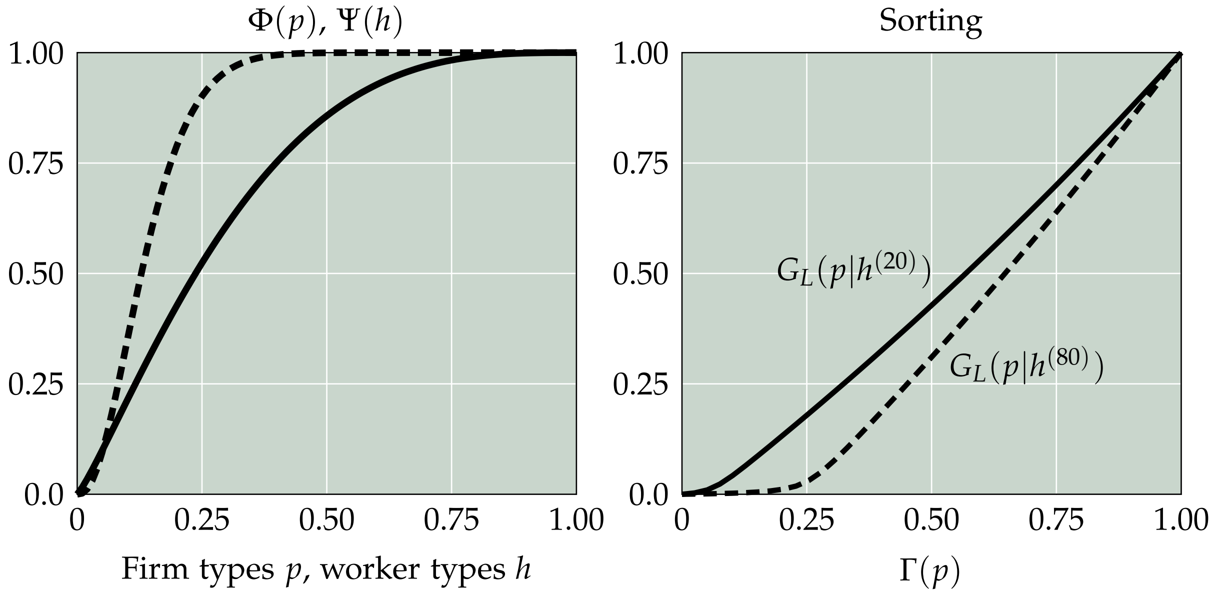
The bargaining parameter is estimated at $\beta = 0.177$. The parameter takes identifica-

Table 2: Structural parameter estimates

Annual job destruction rate, low type, δ_L	0.063 (0.0004)
Annual job destruction rate, high type, δ_H	1.905 (0.0007)
Job destruction type distribution, $\xi_L = \Pr(\delta = \delta_L)$	0.858 (0.0001)
Search cost function $c(s) = \frac{(c_0 s)^{1+1/c_1}}{1+1/c_1}$	
c_0	54.420 (0.0026)
c_1	12.911 (0.0007)
Recruitment cost function $c_v(v) = \frac{v^{1+1/c_{v1}}}{1+1/c_{v1}}$	
c_{v1}	0.012 (0.0004)
Exogenous search, $\underline{s}$	0.034 (0.0004)
Annual reallocation rate, δ_0	0.106 (0.0004)
Off-the-job to on-the-job relative search efficiency, κ	0.845 (0.0004)
Firm productivity CDF on $p \in [0,1]$, $\Phi(p) = Beta(\beta_0^\Phi, \beta_1^\Phi)$	
β_0^Φ (scale)	1.188 (0.0006)
β_1^Φ (shape)	3.151 (0.0007)
Worker skill CDF on $h \in [0,1]$, $\Psi(h) = Beta(\beta_0^\Psi, \beta_1^\Psi)$	
β_0^Ψ (scale)	2.638 (0.0050)
β_1^Ψ (shape)	16.022 (0.0053)
Match production function, $f(h,p) = f_0\Big(\alpha\big(h+\underline{h}\big)^\rho + (1-\alpha)\big(p+\underline{p}\big)^\rho\Big)^{\frac{1}{\rho}}$	
ρ	-2.045 (0.0004)
α	0.311 (0.0005)
f_0	931.169 (0.0020)
Workers' bargaining power, β	0.177 (0.0291)
Std. deviation, wage measurement error, σ_w	0.094 (0.0019)

Note: Standard errors in parentheses.

Figure 5.4: Firm productivity- and worker skill CDFs, and sorting



Note: Right panel: The black solid line shows the estimated population firm productivity CDF $\Phi(p)$. The dashed line shows the estimated population worker skill CDF $\Psi(h)$. Left panel: $h^{(x)}$ is the xth percentile in $\Psi(h)$. The match distribution CDFs of the 20th skill percentile worker, $G_L(p|h^{(20)})$, is drawn in solid and those of the 80th percentile worker, $G_L(p|h^{(80)})$, is drawn with dashed line. The match distributions refer to low layoff type workers.

tion from a number of dimensions in the auxiliary model such as the mean-min wage ratio, and the difference between wages of newly hired workers out of unemployment relative to the steady state wage distribution. As we shall see shortly, the estimate is consistent with the data along all of these dimensions, and is also broadly consistent with other estimates reported in Bagger et al. (2014) and Bagger et al. (2014).

As expected, wages are measured with error although with an estimate $\sigma_w = 0.094$ measurement errors are of modest importance.

In the online appendix, we present the estimation results from a stratification of the data on worker education along with details comments. We stratify the data into two groups: One with less than 15 years of education and one with more than 15 years. Qualitatively, the auxiliary data moments do not vary substantially across the two groups and so it is not surprising that the main conclusions of the pooled sample replicate in the stratified estimations. To the extent that they differ, the high education estimation shows slightly less sorting with a correlation between worker skill and firm productivity of

0.11 whereas the low education group is a bit more sorted at a correlation of 0.15. Less educated workers are estimated to have a somewhat higher bargaining power parameter than high educated workers. The estimated productivity distributions are such that the high education group of workers are on average paid more. Frictions are estimated to play a greater role for the high education group. We quantify the impact of frictions on wages in Section 7 and in that decomposition, frictions turn out to contribute substantially more to high education wage dispersion than to that of the low education workers.

5.4 Model fit

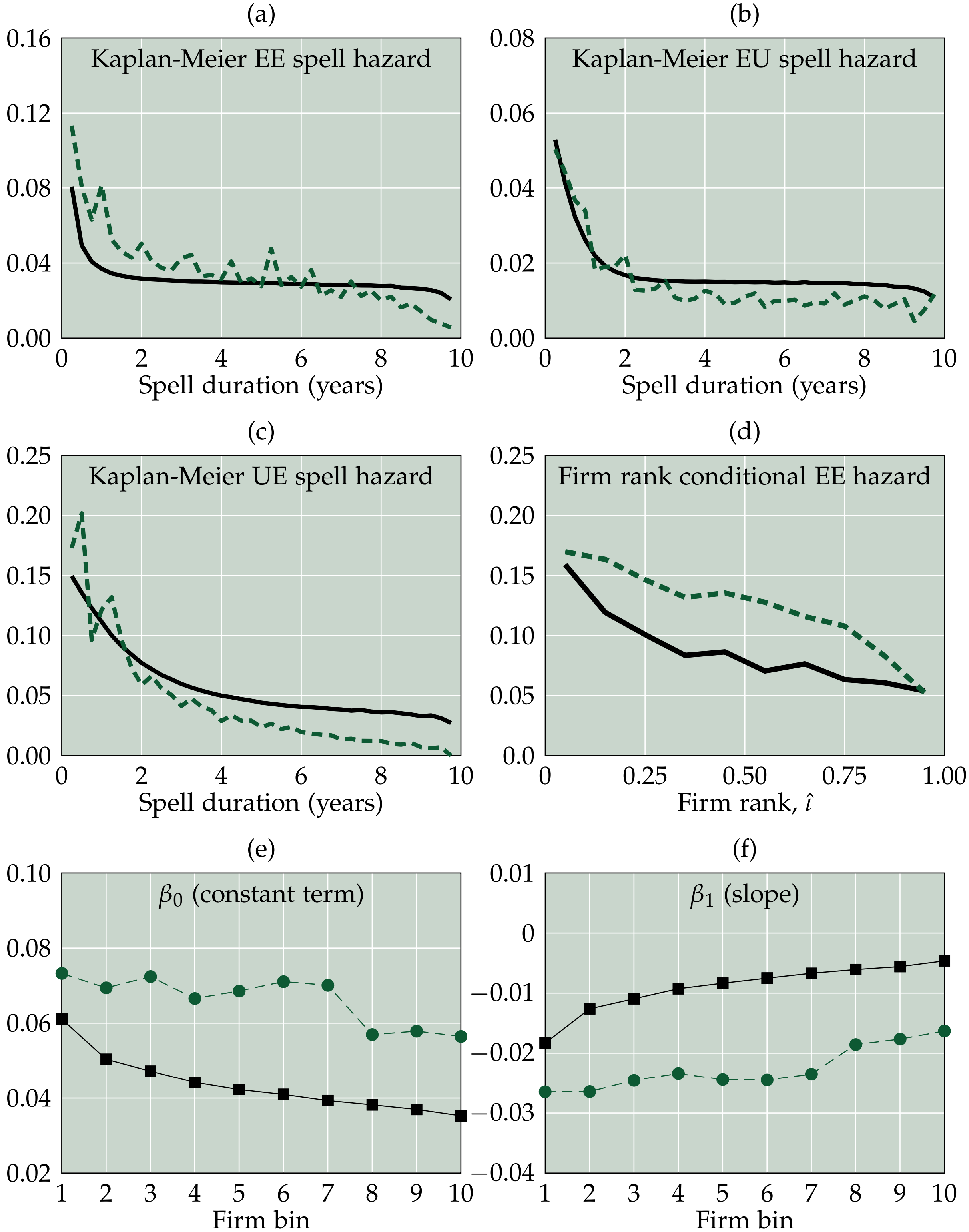
The estimated model's fit to the moment of auxiliary statistics is reported in Table 3 and Figures 5.5. Overall the model fits the data well.

5.4.1 Worker reallocation

Kaplan-Meier hazard functions: Figure 5.5, panels (a) and (b), plot Kaplan-Meier job-to-job and job-to-unemployment hazard rates for real and simulated data. The estimated model reproduces the empirical negative duration dependence in the Kaplan-Meier job-to-job transition hazard function through search intensity heterogeneity induced by layoff, skill and firm productivity heterogeneity, and in the Kaplan-Meier job-to-unemployment transition hazard function, through layoff heterogeneity. Figure 5.5 reveals a good fit to the empirical job-to-job transition hazard, although we underestimate the hazard rate for durations less than 5 years, and overestimate it at longer durations. The model delivers an impressive fits to the empirical job-to-unemployment transitions, especially at shorter job duration of less than 3 years.

We did not use the unemployment-to-job transition hazard rate as a target in the estimation; in fact, our estimation is minimally disciplined with respect to unemployment data. It is therefore reassuring to see from Figure 5.5, panel (c), that the estimated model does a good job in reproducing the observed unemployment-to-job transition hazard rate. Both the empirical and the simulated unemployment-to-job transition hazard function exhibits negative duration dependence, generated by the supermodular production function and layoff rate heterogeneity. If the focus of the model were to fully capture the

Figure 5.5: Model fit



Note: Data in dashed line. Model estimate in solid line.

unemployment-to-job transition rate heterogeneity in the data, Figure 5.5, panel (c) does suggest a role for heterogeneity in the exogenous arrival of job offers across workers so that some workers have close to a zero job finding rate.

Inflow rank conditional job-to-job transition hazard rates: The model's fit to the inflow rank conditional job-to-job transition hazard is generally good. As seen in Figure 5.5, panel (d), the model captures well the declining relationship between firm rank and job-to-job separations. The estimated job separation rate falls a bit too fast with firm rank and underestimates the job separations for the middle firm ranks.

Within-firm job-to-job transition hazard rate duration dependence: Consider coefficients β_{0k} and β_{1k} from (5.3) across inflow rank bins $k = 1, 2, ..., 10$ to assess how the Kaplan-Meier job-to-job hazard rate changes over the firm ladder. In the model, these changes are primarily driven by changes in worker skill heterogeneity over the firm ladder, that is, by sorting. Figure 5.5, panels (e) and (f), plot β_{0k}, in the left panel, and β_{1k}, in the right panel for real and simulated data. The estimated model reproduces the slope of the β_{1k}-profile well, but underestimates β_{1k} at each firm productivity bin k. With respect to the β_{0k}-profiles, the model is again able to reproduce the declining pattern observed in the data, but overestimates the β_{0k} for all k.

Employment cycles: The estimated model also delivers solid fit to the distribution of the number of jobs in employment cycles as shown in Table 3. The model gets the average number of jobs per employment cycle almost exactly right, while it underestimates the variance slightly.

Record statistics: From Table 3, we also note that our model explains well the fraction of existing matches that end in a layoff to unemployment, almost exactly matching the 0.338 fraction of existing jobs that end in a layoff into unemployment.

5.4.2 Cross section heterogeneity

Raw moments: Our model replicates well both the first two moments of the employment weighted distribution of log average firm wages, and the first two moments of

Table 3: Model fit

	Data	Sim.
Worker reallocation		
Number of jobs in employment cycle, average	2.182	2.178
Number of jobs in employment cycle, std. dev.	1.541	1.260
Share of matches ending in EU-transition	0.338	0.348
Cross section heterogeneity		
Log firm wage, employment weighted average	5.254	5.252
Log firm wage, employment weighted std. dev.	0.168	0.160
Log firm wage, newly hired workers average	5.167	5.166
Log firm wage, newly hired workers std. dev.	0.222	0.201
Firm size, average	8.646	8.874
Fraction of active firms to worker population	0.091	0.089
Within-job annual log wage growth, average	0.009	0.005
Mean-min wage ratio	1.854	1.799
Firm effects from (5.4), average	5.238	5.217
Firm effects from (5.4), std. dev.	0.179	0.149
Worker effects from (5.4), std. dev.	0.218	0.220
Residuals from (5.4), std. dev.	0.134	0.129
Unemployment duration and starting wages*		
Unemployment duration (weeks), average	57.395	78.466
Unemployment duration (weeks), std. dev.	69.853	80.484
Starting wage (DKK), average	186.0	172.8
Starting wage (DKK), std. dev.	62.3	44.8
Correlation(unemployment duration, starting wage)	−0.168	−0.381

* All moments in the sorting panel computed on workers hired from unemployment into top ranked firms.

the distribution of log average firm wages among newly hired workers. The difference between the two distributions results from the accumulation of job offers, i.e. the job-to-job transition process as well as the wage bargaining mechanism. The estimated model reproduces shifts in both location and scale between the two distributions. Firm size, in the data measured as the number of workers in a firm at a given cross section date (last week of November), as well as the ratio of firms to workers are also fitted well.[23] The

[23]The design of the simulation replicates the number of firms relative to population in the data and the auxiliary model includes a moment of how many active firms there are in the data as defined by an average labor force size of no less than half a worker at an annual level, which the model fits by largely making the vacancy intensity choice constant across firms. Firm size heterogeneity across firm types is as result primarily determined by differential worker separation rates. One can expand on the firm size fit in the model, but our concern is primarily to make sure that firms have the right labor force size in a first

estimated model, where on-the-job search is the only source of within-job wage dynamics, delivers a log wage growth of 0.005, thus accounting for 56% of the empirical wage growth.

Mean-min ratio: The empirical mean-min ratio is 1.854 while the simulated ratio is 1.799, an almost perfect fit. Unlike the second moment of the wage distribution, the mean-min ratio is not trivially captured by σ_w. The bargaining parameter β plays an important role here as well. The mean-min wage ratio can get very large for low β values because initial wages in employment relationships can in some cases be very low.

Log wage regression: From Table 3 we see that the estimated model is able to reproduce the first two moments of the distributions of firm effects, worker effects and residuals from the auxiliary log wage regression (5.4), even if the model slightly underestimates the overall log wage variance. The empirical log wage variance is 0.097 while its simulated counterpart is 0.087. The underestimation is due to a smaller firm wage effect variance in the model estimate than in the data.

The auxiliary log wage regression is a restricted version of the two-way error component model applied in Abowd et al. (1999). The restriction is (5.6), that worker and firm fixed effects are uncorrelated. We imposed assumption (5.6) to ease computations of the auxiliary firm and worker fixed effects in the Indirect Inference estimation procedure, and because the correlations between worker and firm fixed effects need not be particularly informative about sorting, as shown in section 2.2. However, subsequent to estimation, it is straightforward to assess whether the estimated structural model is in fact able reproduce the Abowd et al. (1999) specification.

Table 4 presents the log wage variance decomposition obtained from a Abowd et al. (1999) log wage regression, i.e. (5.4) with only (5.5) imposed, for real and simulated data.[24] In Table 4 we denote the correlation between worker and firm fixed effects a "wage sorting effect", not to be confused with sorting, a notion that refers to the equilibrium correlation between worker skills and firm productivity. The wage sorting

order sense, so that the inflow measure for ranking firms has about the same measurement noise between data and model simulation.

[24]The data used for this log wage variance decomposition is identical to that used for the estimation of the auxiliary log wage regression used in the estimation of the structural parameters.

Table 4: Log wage variance decomposition—The Abowd et al. (1999) approach.

	Data		Sim.	
	Value	Percent of $Var(\ln w_{in})$	Value	Percent of $Var(\ln w_{in})$
Total log wage variance, $Var(\ln w_{in})$	0.097	100%	0.087	100%
Worker effect, $Var(\chi_i)$	0.070	71%	0.056	64%
Firm effect, $Var(\varphi_{K(i,n)})$	0.014	14%	0.011	13%
Residual effect, $Var(\epsilon_{in})$	0.015	15%	0.015	17%
Wage sorting, $2Cov(\chi_i, \varphi_{K(i,t)})$	-0.002	0%	0.005	6%

effect measures the extent to which high wage workers work in high wage firms. We note that our estimated model reproduces the Abowd et al. (1999) log wage variance decomposition well, with a slight underestimation of the relative importance of worker effects, and a slight overestimation of the relative importance of the wage sorting effect. Recall that the equilibrium correlation between worker skills h and firm productivity p is 0.12. Hence, Table 4 is consistent with our conjecture that, within our framework, the wage sorting effect from a Abowd et al. (1999)-type wage regression provides a lower bound for the degree of sorting in terms of productivity types.

5.4.3 Unemployment duration and wages

Our data suggests that, among workers hired by a top ranked firm, workers with higher starting wages have shorter unemployment durations, the empirical correlation being $Corr(\bar{t}_i^u, w_i^0 | \hat{\iota} \geq 0.95) = -0.168$. The model interprets the negative correlation as a resulting from a supermodular match production function. The simulated model gets the sign of the correlation right, but overestimates the magnitude of the correlation, with the simulated correlation being $Corr(\bar{t}_i^u, w_i^0 | \hat{\iota} \geq 0.95) = -0.381$. Looking towards the fit to the marginal distributions of unemployment durations and starting wages, we see that the estimated model face some problem in fitting unemployment durations, overestimating both the average and the variance.[25] For starting wages, the model slightly

[25]We are operating with a broad definition of unemployment, encompassing both registered unemployment and some nonparticipation. Enriching the model to allow for additional type heterogeneity in the unemployment pool is likely to improve the model's ability to fit unemployment spells. It is however unlikely that the introduction of this feature into the model will have a significant impact on the model estimate.

underestimates the average, and substantially underestimates the variance.

In the online appendix we conduct an overfitting exercise, forcing a near-perfect fit to $Corr(\bar{t}_i^u, w_i^0|\hat{\iota} \geq 0.95)$.[26] As it turns out, overfitting $Corr(\bar{t}_i^u, w_i^0|\hat{\iota} \geq 0.95)$ has very little impact on the estimated structural parameters, in particular the estimate of ρ remains almost the same. Perhaps more importantly, all of our conclusions regarding sorting, mismatch and wage dispersion remains intact with overfitting of $Corr(\bar{t}_i^u, w_i^0|\hat{\iota} \geq 0.95)$.

6 Mismatch

The notion of labor market mismatch in for example Gautier and Teulings (2012) and in Shimer and Smith (2000) take the populations of jobs and workers as given and ask to what extent the equilibrium in question can be improved upon by changing the allocation of workers to firms. We determine a similar notion of mismatch in our setting: For the estimated equilibrium match distribution, $G_j(h, p)$ and employment level e_j, determine the mass of jobs, $e = e_0 + e_1$ and the distribution of productivity of those jobs, $G(p) = \left[\sum_{j=0}^{1} e_j G_j(1, p)\right] / (e_0 + e_1)$. Then for the estimated population of workers, unemployed and employed, assign them to jobs to maximize aggregate output. We will not concern ourselves with how exactly a social planner might make this assignment happen. The assignment represents a first best non-frictional assignment in the spirit of a core assignment in Becker (1973), but for a given population of jobs.[27]

The counter factual also means that we are not concerned with job loss since a worker can immediately be put into the position in question. Thus, since the production function is estimated to be supermodular, the optimal allocation result in Becker (1973) dictates that the highest skill worker be matched with the most productive job, the second highest skill worker with the second highest productivity job, and so forth.

For the estimated model, we find that taking the estimated population of jobs and

[26]Our Indirect Inference estimator measures the distance between empirical and simulated auxiliary statistics using the inverse variance-covariance matrix of the empirical moments as the metric. Hence, relatively less precisely estimated auxiliary statistics receive less weight in the estimation. By construction, $Corr(\bar{t}_i^u, w_i^0|\hat{\iota} \geq 0.95)$ is computed from the relatively small sample of workers and is relatively imprecisely estimated. The overfitting exercise simply scales up the weight put on $Corr(\bar{t}_i^u, w_i^0|\hat{\iota} \geq 0.95)$ in the estimation.

[27]This counter factual should not be confused with an exercise of eliminating frictions in our model and letting job creation respond in equilibrium. For our notion of mismatch, the estimated population of jobs is held constant.

workers and allocating them efficiently produces an output increase of 7.7%. The estimated correlation between worker skill and firm productivity is estimated to be 0.12 and therefore suggest substantial noise in the allocation suggesting scope for substantial efficiency improvements. This has to be weighted against the estimated amount of complementarity in the production function as well as the estimated dispersion in productive heterogeneity across workers and jobs. In the absence of complementarities, the efficiency gain would be zero.

7 Log wage variance decompositions

Given worker skill and firm productivity distributions may produce radically different wage distributions, depending on the allocation of workers to firms implemented by the labor market. Hence, the study of wage dispersion must include an understanding of sorting, something that has hitherto eluded the literature. In this section we use the estimated structural model to provide a novel decomposition of log wage variance into a worker effect, a firm effect, a friction effect and a sorting effect.

We conduct the analysis on simulated data as we need the structural elements (j, h, p, q) which are unobserved in the real data. It is straightforward to simulate steady state data for I workers. The simulated data is $\{w_i, j_i, h_i, p_i, q_i\}$ where $i = 1, 2, ..., I$ index workers, and where $w_i = w_{j_i}(h_i, p_i, q_i)$ according to (2.8). We take $I = 100,000$. The object of interest is $Var(\ln w_i)$. As it turns out 99.8 percent of simulated log wage variation arise within layoff-type. Given the very low share of high layoff type workers among employed workers, less than 1%, understanding total log wage variance becomes a matter of understanding log wage variation among low layoff type workers. Our analysis is consequently focused exclusively on low layoff type workers ($j = L$). To proceed, we project simulated $\ln w_i$ onto $(1, h_i, p_i, q_i)$, and base our variance decomposition on the resulting (minimum mean square) prediction of the simulated wages:

$$\widehat{\ln w_i} = \widehat{\varrho}_0 + \widetilde{h}_i + \widetilde{p}_i + \widetilde{q}_i, \tag{7.1}$$

with $\widetilde{h}_i \equiv \widehat{\varrho}_h h_i$, $\widetilde{p}_i \equiv \widehat{\varrho}_p p_i$, and $\widetilde{q}_i \equiv \widehat{\varrho}_q q_i$, and where $(\widehat{\varrho}_0, \widehat{\varrho}_h, \widehat{\varrho}_p, \widehat{\varrho}_q)'$ are projection parameters.[28] We refer to $\widetilde{h}_i$ as the worker skill factor, to $\widetilde{p}_i$ as the firm productivity factor, and

[28]The predicted log wages based on (7.1) are monotone in h_i, p_i and q_i. The issues with nonmonotone

to $\widetilde{q}_i$ as the outside option factor. The factors $(\widetilde{h}_i, \widetilde{p}_i, \widetilde{q}_i)$ are not independent. $\widetilde{p}_i$ and $\widetilde{q}_i$ are mechanically related because $p_i \leq q_i$, and are both increasing in a worker's search capital. Labor market sorting implies dependence between $\widetilde{h}_i$ and $\widetilde{p}_i$ and $\widetilde{q}_i$.[29] We now drop subscript i to avoid clutter.

Separating between- and within-firm variance in (7.1) such that

$$Var(\widehat{\ln w_i}) = \underbrace{Var(E[\widetilde{h}_i + \widetilde{p}_i + \widetilde{q}_i | \widetilde{p}_i])}_{\text{Between-firm}} + \underbrace{E(Var[\widetilde{h}_i + \widetilde{p}_i + \widetilde{q}_i | \widetilde{p}_i])}_{\text{Within-firm}},$$

expanding the variance expressions, and rearranging terms yields

$$\begin{aligned} Var(\widehat{\ln w}) \quad = \quad & \underbrace{Var(\widetilde{p}) + Var(E[\widetilde{q}|\widetilde{p}]) + 2Cov(\widetilde{p}, E[\widetilde{q}|\widetilde{p}])}_{\text{Firm effect}} + \underbrace{E(Var[\widetilde{h}|\widetilde{p}])}_{\text{Worker effect}} \\ & + \underbrace{E(Var[\widetilde{q}|\widetilde{p}])}_{\text{Friction effect}} + \underbrace{Var(E[\widetilde{h}|\widetilde{p}]) + 2Cov(E[\widetilde{h}|\widetilde{p}], \widetilde{p})}_{\text{Sorting effect (to be continued)}} \\ & \underbrace{+ 2Cov(E[\widetilde{h}|\widetilde{p}], E[\widetilde{q}|\widetilde{p}]) + 2E(Cov[\widetilde{h}, \widetilde{q}, |\widetilde{p}])}_{\text{Sorting effect (continued)}}. \end{aligned} \tag{7.2}$$

In (7.2), the firm effect contains between-firm variation in $\widetilde{p} + \widetilde{q}$. This is a natural definition as $\widetilde{p}_{it}$ is the firm productivity factor, and $\widetilde{q}_{it}$ is the outside option factor, which according to the structural model, correlates with firm productivity. The worker effect contains the within-firm variation in the worker skill factor $\widetilde{h}_{it}$. For given firm productivity and outside option, the only source of wage variation is worker skills. The friction effect, a notion introduced in Postel-Vinay and Robin (2002), reflects within-firm variation in outside options. Within a firm, variation in outside options arise due to the frictions, i.e. stochastic, arrival of job opportunities to workers. The sorting effect is driven by covariance between the worker skill factor $\widetilde{h}_{it}$ on the one side, and firm productivity and outside option factor $\widetilde{p}_{it}$ and $\widetilde{q}_{it}$ on the other side.

Table 5 reports (7.2) based on our estimated model. The linear projection comprises 92% of the variation in simulated log wages. Worker effects account for 51% of predicted log wage variation, firm effects for 11%, and friction effects for 23%. Labor market

wages discussed in section 2.2 occur when the outside option q_i is integrated out of the wage equation. This highlights the importance of having the estimated structural model for making correct inferences; without it, we would not know the equilibrium distribution of (h_i, p_i, q_i), and would be unable to appropriately control for the effect of outside options in (7.1).

[29] $\widetilde{h}_i$, $\widetilde{p}_i$ and $\widetilde{q}_i$ are also related in that they depend on the projection parameters $\widehat{\varrho}_h$, $\widehat{\varrho}_p$ and $\widehat{\varrho}_q$, which are reduced form parameters. We ignore this channel in the interpretation of our results.

Table 5: Log-Wage Variance Decomposition—Accounting for Labor Market Sorting

	Value	% of $Var(\ln w_{it})$ (predicted)
Worker effect	0.037	51%
Firm effect	0.008	11%
Friction effect	0.017	23%
Sorting effect	0.011	15%
Total predicted variance	0.073	100%
Total simulated variance	0.079	108%

sorting effects account for 15%. Hence, labor market sorting contributes substantially to the dispersion of wages across workers.

If there is no sorting, i.e. if $E[\widetilde{h}_{it}|\widetilde{p}_{it},\widetilde{q}_{it}] = E[\widetilde{h}_{it}]$, (7.2) simplifies. First, we have $Var(E[\widetilde{h}_{it}|\widetilde{p}_{it}]) = 0$ and $E(Var[\widetilde{h}_{it}|\widetilde{p}_{it}]) = Var(\widetilde{h}_{it})$, so that the worker effect is simply the variance of worker skills in the population. Second, all the covariance terms in the sorting effect vanishes. Without sorting, our variance decomposition thus resembles that of Postel-Vinay and Robin (2002).[30] Using French data, they find that worker effects account for 0-35% of the variance of log wages, depending on occupation, with worker effects being more important in higher occupations. Firm effects accounts for 20-50% of log wage variation, with wage variation in higher occupations being less dependent on firm heterogeneity. Finally, Postel-Vinay and Robin (2002) find that search frictions accounts for 40-61% of total log wage variation, depending on the occupation under consideration.[31]

[30] Their structural wage equation may be written as $\ln w_{it} = h_{it} + p_{it} + \Xi_{it}$ where $\Xi_{it} = \Xi(p_{it}, q_{it})$ is the log share of output transferred to the worker (see Postel-Vinay and Robin (2002, p. 2305, equation (5))). A within- and between-firm log wage variance decomposition then yields

$$Var(\ln w_{it}) = \underbrace{Var(h_{it})}_{\text{Worker effect}} + \underbrace{Var(p_{it}) + 2Cov(p_{it}, \Xi_{it}) + Var(E[\Xi_{it}|p_{it}])}_{\text{Firm effect}} + \underbrace{E[Var(\Xi_{it}|p_{it})]}_{\text{Friction effect}}.$$

With no sorting our decomposition is not exactly identical to that of Postel-Vinay and Robin (2002) due to a number of other differences between the two models, including an unrestricted bargaining power parameter and endogenous search intensities.

[31] Our decomposition (7.2) also bears some resemblance to the decompositions presented in Abowd et al. (1999) and Abowd et al. (2002). These latter decompositions are reduced form, based on a log wage regression like our auxiliary log wage regression (5.4) but where the correlation between worker and firm fixed effects is left unrestricted, i.e. without assumption (5.6) imposed. Such a regression gives rise to a decomposition of wage dispersion into worker effects, firm effects, residual effects, and "wage sorting" effects. As noted in relation to Table 4, our structural model does a good job in reproducing Abowd et al.

8 Concluding Remarks

A labor market addresses mismatch through worker reallocation. The greater the mismatch, the greater the urgency of the reallocation. Indeed, empirical evidence documents that reallocation is a common occurrence in most labor markets and that reallocation directly between jobs more often than not are associated with wage increases. This paper quantifies the contribution of labor market heterogeneity to wage dispersion in a frictional labor market setting where assortative matching may be present.

The model is estimated on Danish matched employer-employee data. The estimation can be viewed to be a production function estimation study. But output is not observed at the match level and wages fail to be a reliable reflection of the match production function because they are possibly non-monotone functions of the productivity indices. Hence, inference is necessarily indirect and emphasizes systematic reallocation rate differences across different worker and firm type matches.

In addition to the variation in the data directly related to identification of sorting, the model is also disciplined to fit a large number of other statistics from the data through an indirect inference estimation. In this sense, the paper also serves as a robustness check on previous wage dispersion measurement papers that are estimated on a narrower view of the data. In general, the model does quite well in fitting the data. The estimation exercise is heavily over identified. It leaves some unexplained reallocation rate heterogeneity which is probably not all that unattractive given the stylized nature of the model. We check whether forcing a perfect fit to the moments that are primarily related to sorting and find that it has a minimal impact on the sorting implications.

In the estimated model wage variation is decomposed into four sources: Worker heterogeneity (51%), firm heterogeneity (11%), friction (23%), and sorting (15%). The match production function is estimated to be supermodular implying positive assortative matching. Through the model's wage determination mechanism it incents more skilled workers to search with greater urgency to reallocate to better firms. The correlation coefficient between worker skill and firm productivity is 0.12 in the steady state match distribution. The associated mismatch implies that output could be increased by 7.7% if the estimated set of matched workers and firms were perfectly sorted.

(1999)-type wage regressions, and the associated variance decompositions.

Appendix

A Wage regressions and monotonicity

Ignoring the role of observable covariates, and subsuming the constant term into, say, the firm effect, Abowd et al. (1999) assume a log wage equation where worker and firm fixed effects enter additively,

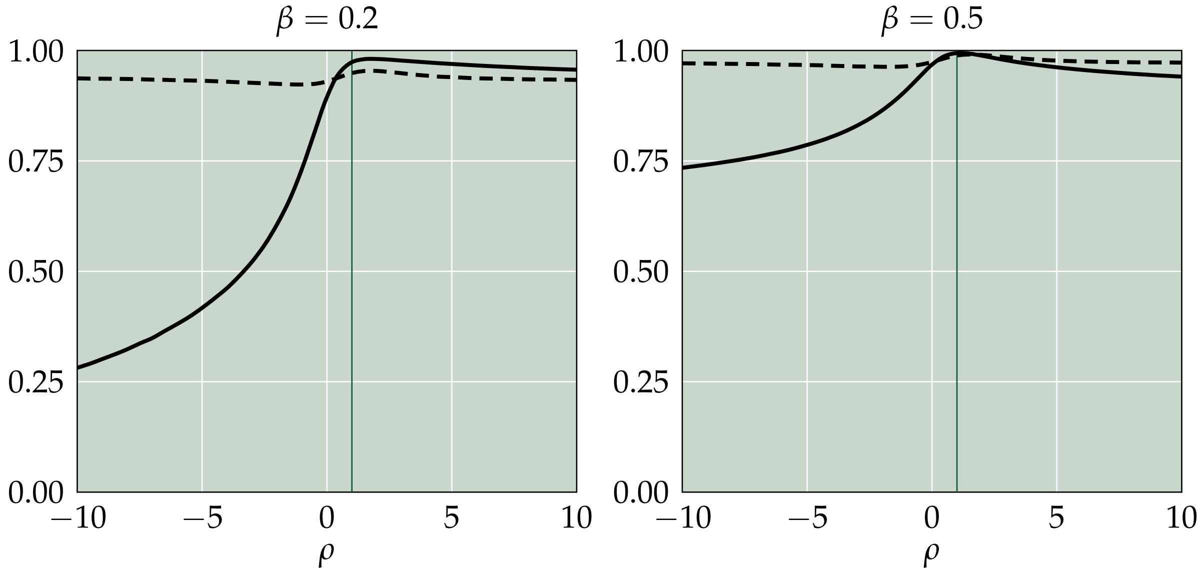
$$w_{in} = \chi_i + \varphi_{\mathrm{K}(i,n)} + \epsilon_{in}, \tag{A.1}$$

where $\mathrm{K}(i,n)$ is the firm ID that worker i is matched with at observation time n, and χ_i and φ_k are the worker and firm fixed effects. The notation as in the main part of the paper, see the description related to our auxiliary log wage regression (5.4) with (5.5) imposed. The identification of the fixed effects from matched employer-employee data relies on this additive structure. Consider a class of models where workers differ by skill and firms by productivity. An agent's type is permanent. Furthermore, match output is increasing in both skill and productivity. Can the estimated worker and firm fixed effects from the log-linear wage equation be used as the basis for identification of the underlying worker skill and firm productivity heterogeneity? In particular, does the correlation between the estimated worker and firm fixed effects, $corr(\widehat{\chi}_i, \widehat{\varphi}_{\mathrm{K}(i,n)})$, identify sorting in the matching between worker skill and firm productivity? Eeckhout and Kircher (2011) provide a negative answer for their model. We will generally provide a negative answer as well. Both answers are based on the insight that for the model structures in question, the log additive wage equation is fundamentally misspecified with respect to the worker and firm heterogeneity contributions to wages. Specifically, wages are generally not monotonically increasing in skill and productivity.

In Figures A.1 and A.2 we relate estimates of worker and firm fixed effects from the wage equation (A.1) to the true underlying worker skill and firm productivity heterogeneity in simulations of steady state equilibria for different (ρ, β) combinations.

Figure A.1 shows $corr(\widehat{\chi}_i, h_i)$ and $corr(\widehat{\varphi}_k, p_k)$. It is seen that the wage equation firm fixed effect is strongly correlated with firm productivity regardless of the type and strength of sorting and worker's bargaining power. Not surprisingly, higher bargaining power increases the correlation.

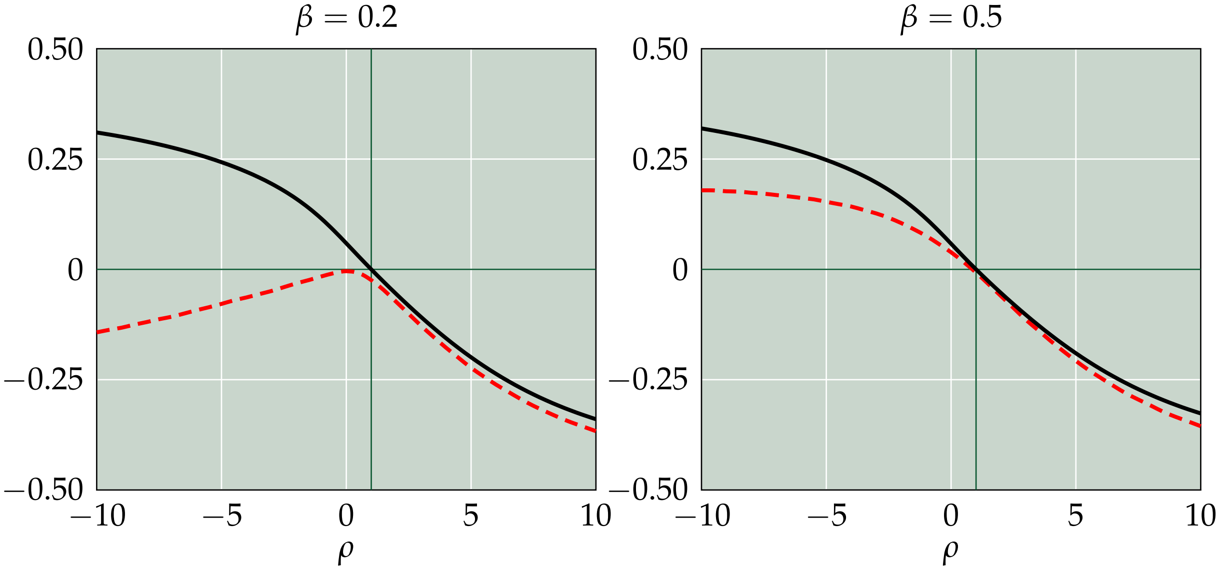
Figure A.1: The correlation between wage fixed effects and true agent heterogeneity for given (ρ, β) combinations.



Note: The solid and dashed lines show $\text{cor}[\hat{\chi}, h]$ and $\text{cor}[\hat{\varphi}, p]$, respectively. For the given model specification, the production function scale parameter (f_0) and the base offer arrival rate (λ) are set such that the the steady state equilibrium solution satisfies $u = 0.05$ and $\text{E}[w(h, p)] = 180.0$. The dashed red line at $\rho = 1$ divides the model specifications with positive sorting for $\rho < 1$ and negative sorting for $\rho > 1$.

The correlation between the wage equation worker fixed effect and worker skill is on the other hand quite sensitive to the specification of the model. If sorting is positive and wage determination is primarily set by wage posting, then the correlation is low. In this case, the wage profiles of more skilled workers are characterized by substantial wage growth over an employment spell, and consequently, the notion of a wage equation worker fixed effect is misplaced. As documented in Figure 2.1 it is in this type of equilibrium also perfectly possible to observe more skilled workers receive lower wages than less skilled workers within a given firm. In such a case, the estimation will tend to rank the less skilled worker with a higher fixed effect than the more skilled worker. This mechanism is strengthened by the assumption that the wage equation has an i.i.d. over time error process, ϵ_{in} and the fact that even for the high skilled workers, the wage process has some permanence to it. Since the more skilled worker's realized wage growth is often associated with an actual job-to-job transition, the estimation will be allowed to

Figure A.2: The correlation between skill and productivity for given (ρ, β) combinations.



Note: The solid line is $\text{cor}[h, p]$. The dashed line is $\text{cor}[\chi_i, \varphi_{K(i,n)}]$. The wage equation fixed effects are estimated on simulated data from the given steady state equilibrium. For the given model specification, the production function scale parameter (f_0) and the base offer arrival rate (λ) are set such that the the steady state equilibrium solution satisfies $u = 0.05$ and $\text{E}[w(h, p)] = 180.0$. The dashed red line at $\rho = 1$ divides the model specifications with positive sorting for $\rho < 1$ and negative sorting for $\rho > 1$.

explain the substantial observed wage growth of the high skilled worker by increasing the wage equation fixed effect differential between the two firms involved in the job-to-job transition, thereby laying a foundation for a negative bias in the correlation between wage equation worker and firm fixed effects.

In the negative sorting case, low skilled workers are the ones taking temporary current wage hits with the expectation of future gains. As a result, in this type of equilibrium wages are monotonically increasing in worker skill within a given firm and the ranking of wage equation worker fixed effects will be aligned with the skill ranking. This accounts for the strong positive correlation between the estimated wage equation worker fixed effects and worker skill for the negative sorting cases, $\rho > 1$.

For higher β, where wage determination is to a greater extent set by bargaining rather than posting, $corr(\widehat{\chi}_i, h_i)$ is higher because wages are moving towards being monotone in worker skill and firm productivity.

Figure A.2 presents the correlation between the wage equation fixed effects in relation to the correlation between the skill and productivity indices in the equilibrium steady state match distribution. The correlation between h and p based on $G(h, p)$ reveals the basic property of the model that sorting is positive for $\rho < 1$, negative for $\rho > 1$, and there is no sorting when $\rho = 1$. It is seen that when $\beta = 0.2$ and there is negative sorting, the correlation between wage equation worker and firm fixed effects, $E_n[corr(\widehat{\chi}_i, \widehat{\varphi}_{K(i,n)})]$ is very close to equilibrium steady state $corr(h, p)$. This is consistent with the results in Figure A.2 that the estimated wage equation worker and firm fixed effects are closely correlated with the skill and productivity indices in this case. When sorting is positive and $\beta = 0.2$, we see that $E_n[corr(\widehat{\chi}_i, \widehat{\varphi}_{K(i,n)})]$ and $corr(h, p)$ diverge. In this case, the worker fixed effects are so poorly related to the skill ranking that the resulting negative bias drives the correlation between χ and φ negative. As a result, $E_n[corr(\widehat{\chi}_i, \widehat{\varphi}_{K(i,n)})]$ is negative both when sorting is positive and negative for this case.

In the case where $\beta = 0.5$, the fixed effects correlation $E_n[corr(\widehat{\chi}_i, \widehat{\varphi}_{K(i,n)})]$ does quite well in capturing the steady state match correlation between skill and productivity. There is some negative bias in the positive sorting case, but in this case, the correlation coefficients share the same signs.

The above results suggest that an observed positive value of $E_n[corr(\widehat{\chi}_i, \widehat{\varphi}_{K(i,n)})]$ indicates that sorting between skill and productivity is positive. In general, the correlation coefficient between h and p is always greater than $E_n[corr(\widehat{\chi}_i, \widehat{\varphi}_{K(i,n)})]$. It is also worth emphasizing that the often observed small and negative correlation between χ and φ is consistent with anything from mild negative sorting to strong positive sorting between h and p.

References

Abowd, John M., Robert H. Creecy, and Franics Kramarz (2002). Computing person and firm effects using linked longitudinal employer-employee data. *Working paper*.

Abowd, John M., Francis Kramarz, and David N. Margolis (1999). High wage workers and high wage firms. *Econometrica* 67, no. 2: 251–334.

Bagger, J., B. J. Christensen, and D. T. Mortensen (2014). Wage and labor productivity dispersion: The roles of total factor productivity, labor quality, capital intensity, and rent sharing. *Working paper*.

Bagger, Jesper, François Fontaine, Fabien Postel-Vinay, and Jean-Marc Robin (2014). Tenure, experience, human capital and wages: A tractable equilibrium search model of wage dynamics. *American Economic Review*: forthcoming.

Barlevy, Gadi and H. N. Nagaraja (2013). Estimating mobility rates in search models with initial condition problems. *Working paper*.

Bartolucci, Christian and Francesco Devicienti (2013). Better worker move to better firms: A simple test to identify sorting. *Working paper*.

Becker, Gary S. (1973). A theory of marriage: Part I. *The Journal of Political Economy* 81, no. 4: 813–46.

Bunzel, Henning, Bent Jesper Christensen, Peter Jensen, Nicholas M. Kiefer, Lars Korsholm, Lars Muus, George Neumann, and Michael Rosholm (2001). Specification and estimation of equilibrium search models. *Review of Economic Dynamics 4*: 90–126.

Burdett, Kenneth and Dale T. Mortensen (1998). Wage differentials, employer size, and unemployment. *International Economic Review* 39, no. 2: 257–273.

Cahuc, Pierre, Fabien Postel-Vinay, and Jean-Marc Robin (2006). Wage bargaining with on-the-job search: Theory and evidence. *Econometrica* 74, no. 2: 323–64.

Christensen, Bent Jesper, Rasmus Lentz, Dale T. Mortensen, George Neumann, and Axel Werwatz (2005). On the job search and the wage distribution. *Journal of Labor Economics* 23, no. 1: 31–58.

Dey, Matthew S. and Christopher J. Flinn (2005). An equilibrium model of health insurance provision and wage determination. *Econometrica* 73, no. 2: 571–627.

Eeckhout, Jan and Philipp Kircher (2011). Identifying sorting - in theory. *Review of Economic Studies* 78, no. 3: 872–906.

Fallick, Bruce C. and Charles A. Fleischman (2001). The importance of employer-to-employer flows in the u.s. labor market. *Federal Reserve Board Finance and Economics Discussion Paper 2001-18*.

Fujita, Shigeru and Guiseppe Moscarini (2014). Recall and unemployment. *Working paper*.

Gautier, Pieter A. and Coen N. Teulings (2012). Sorting and the output loss due to search frictions. *Working paper*.

Gourieroux, Christian, Alain Monfort, and Eric Renault (1993). Indirect inference. *Journal of Applied Econometrics* 8, no. 0: S85–S118.

Hagedorn, Marcus, Tzuo Hann Law, and Iourii Manovskii (2013). Identifying sorting. *Working paper*.

Hornstein, Andreas, Per Krusell, and Giovanni L. Violante (2011). Frictional wage dispersion in search models: A quantitative assessment. *American Economic Review* 101, no. 4: 2873–2898.

Jolivet, Grégory, Fabien Postel-Vinay, and Jean-Marc Robin (2006). The empirical content of the job search model: Labor mobility and wage distributions in europe and the US. *European Economic Review* 50, no. 4: 877–907.

Lentz, Rasmus (2010). Sorting by search intensity. *Journal of Economic Theory* 145, no. 4: 1436–1452.

Lentz, Rasmus (2014). Optimal employment contracts with hidden search. *NBER Working Papers*, no. 19988.

Lise, Jeremy, Costas Meghir, and Jean-Marc Robin (2013). Matching, sorting, and wages. *Working paper*.

Lopes de Melo, Rafael (2008). Sorting in the labor market: Theory and measurement. *Working Paper*.

Nagypál, Éva (2005). On the extent of job-to-job transitions. *Working paper*.

Postel-Vinay, Fabien and Jean-Marc Robin (2002). Equilibrium wage dispersion with worker and employer heterogeneity. *Econometrica* 70, no. 6: 2295–2350.

Rubinstein, Ariel (1982). Perfect equilibrium in a bargaining model. *Econometrica* 50, no. 1: 97–109.

Shimer, Robert and Lones Smith (2000). Assortative matching and search. *Econometrica* 68, no. 2: 343–369.

van den Berg, Gerard J. and Geert Ridder (1999). An empirical equilibrium search model of the labor market. *Econometrica* 66, no. 5: 1183–1221.

Economics Working Papers

2013-26:	Torben M. Andersen, Jonas Maibom, Michael Svarer and Allan Sørensen: Do Business Cycles Have Long-Term Impact for Particular Cohorts?
2013-27:	Martin Paldam: Simulating publication bias
2013-28:	Torben M. Andersen and Allan Sørensen: Product market integration, tax distortions and public sector size
2014-01:	Leonie Gerhards and Neele Siemer: Private versus Public Feedback - The Incentive Effects of Symbolic Awards
2014-02:	Casper Worm Hansen, Peter Sandholt Jensen and Lars Lønstrup: The Fertility Transition in the US: Schooling or Income?
2014-03:	Mette Trier Damgaard and Christina Gravert: Now or never! The effect of deadlines on charitable giving: Evidence from a natural field experiment
2014-04:	Christina Gravert: Pride and Patronage - The effect of identity on pay-what-you-want prices at a charitable bookstore
2014-05:	Julia Nafziger: Packaging of Sin Goods - Commitment or Exploitation?
2014-06:	Sylvanus Kwaku Afesorgbor and Peter A.G. van Bergeijk: Measuring multi-membership in economic integration and its trade-impact. A comparative study of ECOWAS and SADC
2014-07:	Hristos Doucouliagos and Martin Paldam: Finally a breakthrough? The recent rise in the size of the estimates of aid effectiveness
2014-08:	Martin Paldam: The public choice of university organization. A stylized story with some explanation
2014-09:	Tor Eriksson, Zhihua Qin and Wenjing Wang: Firm-level Innovation Activity, Employee Turnover and HRM Practices – Evidence from Chinese Firms
2014-10:	Erik Strøjer Madsen and Yanqing Wu: Advertising and concentration in the brewing industry
2014-11:	Jesper Bagger and Rasmus Lentz: An Empirical Model of Wage Dispersion with Sorting