An Empirical Model of Wage Dispersion with Sorting

Jesper Bagger
Royal Holloway, University of London

Rasmus Lentz
University of Wisconsin-Madison
NBER, and LMDG

June 14, 2012

Abstract

The paper studies contributions to wage dispersion in a model that allows for sorting in firm-worker matches. The model is a general equilibrium on-the-job search model with wage formation similar to that of Cahuc et al. (2006). Workers differ in their permanent skill level and firms differ with respect to productivity. As shown in Lentz (2010), in this setting, positive (negative) sorting results if the match production function is supermodular (submodular).

The model is estimated on Danish matched employer-employee data that cover the entire worker and firm population at a weekly observation frequency. The data allow a detailed view of worker and firm conditional spell hazard heterogeneity, which is at the core of the paper's identification strategy. In addition the data contain match wages which also enter the estimation, however, the estimation does not employ the direct strategy of estimating the correlation between worker and firm wage fixed effects. s shown in previous versions of this paper as well as de Melo 2008 and Lise et al. 2008, this approach fails to identify sorting on worker-firm types in models where wages are possibly non-monotone functions of the fundamental worker skill and firm productivity heterogeneity.

Preliminary estimates point to positive sorting between worker skill and firm productivity, although with modest efficiency gains if the estimated population of jobs and workers are allocated efficiently.

1 Introduction

Recent evidence suggests that worker skill and firm productivity heterogeneity are both important contributions to dispersion in observed wages.[1] It is also a well documented fact that at any point in time the labor market is characterized by a large amount of worker reallocation through job transitions where workers are chasing higher wages by moving away from jobs with lower wages into jobs with higher wages.[2] Hence, by its role in directing labor flows, one can view wage dispersion as a central component of the labor market's allocation mechanism. Therefore, the study of wage dispersion must include an understanding of the actual allocation of workers to firms that the labor market is implementing. In particular, this includes the issue of sorting.

Sorting may play an important role as a source of wage dispersion. Clearly, a given distribution of worker and firm types can produce very different output and wage distribution outcomes depending on how matches are formed. Previous work on the estimation of sources of wage dispersion in Abowd et al. (1999), Postel-Vinay and Robin (2002), and Cahuc et al. (2006) has adopted the maintained and identifying assumption that match formation is independent of the types of the agents involved. The analysis of Abowd et al. (1999) directly estimates individual worker and firm fixed effects. Subsequent to the estimation the authors test whether the estimated fixed effects are correlated in the data and find little correlation. This has been taken as evidence that sorting is not an important issue in the labor market. It is however problematic to test the hypothesis of sorting within a framework where the maintained identifying assumptions rule out key mechanisms that can produce sorting in models with production function complementarities.

This paper puts forth a general equilibrium on-the-job search model with both firm and worker heterogeneity. The analysis is based on an off-the-shelf model of on-the-job search with endogenous search intensity as in Christensen et al. (2005) combined with wage determination as in Cahuc et al. (2006). The model is analyzed in detail in Lentz (2010). Depending on the production function, the worker's search intensity can be type dependent and sorting will result. The subsequent empirical analysis will decompose wage dispersion into 4 sources; worker het-

[1] See for example Postel-Vinay and Robin (2002) and Abowd et al. (1999).

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

erogeneity, firm heterogeneity, friction, and sorting. Postel-Vinay and Robin (2002) decompose dispersion into the first three components.[3]

Abowd et al. (1999) perform a decomposition of observed wage dispersion in French matched employer-employee data into unobserved worker and firm fixed effects. The panel structure of such data sets allows the continued observation of a single worker matched with different employers which is the basis of identification of individual fixed effects. The identification of the fixed effects is done under the maintained assumption that job transitions are not correlated with either worker or firm fixed effects. This precludes natural sorting mechanisms in job transition models.

Postel-Vinay and Robin (2002) and Cahuc et al. (2006) make the point that the identification of unobserved fixed effects in Abowd et al. (1999) can be biased in the presence of frictions. Specifically, the contribution of worker fixed effect dispersion to overall wage dispersion can be upward biased if the estimation does not specifically control for the particular properties of the wage process in an on-the-job search model. In these papers, wage dispersion is explained through a structural estimation of a general equilibrium on-the-job search model. Friction is given a role separate from dispersion in worker and firm effects in the explanation of overall wage dispersion. Both of the papers assume that the distribution of worker types is independent of the type of the firm. The production function in these papers is supermodular. However, the matching technology is such that sorting does not arise. This paper proposes a search technology where sorting may arise in response to production function complementarities. The theoretical aspects of the model are described in greater detail in Lentz (2010).

2 Model

Following Lentz (2010), the framework of the model is an endogenous search intensity model with type heterogeneity on both the worker and firm side. Wages are determined through sequential bargaining similar to Dey and Flinn (2005) and Cahuc et al. (2006).

There is a continuum of firms and potential entrants with measure m, and a continuum of

[3]Lise et al. (2008) and de Melo (2008) both study wage dispersion and sorting in a Shimer and Smith (2000) style partnership model.

workers with measure normalized at unity. A worker is characterized by his or her permanent innate ability h 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 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 can search for a new job. The analysis will allow that the search technology may 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 market tightness. By assumption $\lambda'(\theta) \geq 0$. The cost of a search intensity s is given by the increasing and convex function,

$$c(s) = \frac{c_0 s^{1+\frac{1}{c_1}}}{1+\frac{1}{c_1}}, \tag{2.1}$$

where $c_0 > 0$ is a scale parameter and $c_1 > 0$ sets curvature.

A match between a type h worker and a type p firm produces value added $f(p,h)$ net of payments to capital inputs. It is assumed that $f_p(h,p) \geq 0$ 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. Hence, the labels by which types are indexed, h and p, define unambiguous rankings such that a high h indicates a placement in the top of the worker skill ranking and a high p value indicates a top placement in the firm productivity ranking. Statements on sorting then become statements about match allocation patterns in terms of worker skill and firm productivity rankings. We adopt the particular production function specification,

$$f(h,p) = f_0\big(\alpha h^{\rho} + (1-\alpha)p^{\rho}\big)^{\frac{1}{\rho}}, \tag{2.2}$$

where f_0 is a scale parameter, and $0 \leq \alpha \leq 1$. If $\rho < 1$, then the production function is supermodular. It is submodular for $\rho > 1$. The production function is modular for $\rho = 1$. As shown in

Lentz (2010) if the production function is supermodular, the equilibrium will be characterized by positive sorting between worker skill and firm productivity. If it is submodular, negative sorting will result. For $\rho = 1$ there will be no sorting between worker skill and firm productivity types.

Match separation occurs as the result of one of three distinct events. First, the worker in the match may receive an offer from an outside firm with greater productivity than the current firm which induces a quit. Second, at rate $\delta_0\lambda(\theta)$ the worker makes a job-to-job transition where the new job is drawn randomly from the vacancy offer distribution 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 offer ladder.[4] One possible explanation is that a, 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. It may also be that the worker has been given notice of a lay-off sufficiently far in advance that the worker was able to obtain a new job without an actual unemployment spell in between. The model does not take a stand on the nature of the shock. It simply allows that exogenous match separations can occur where the worker's climb up the offer ladder is reset but without the association of an actual unemployment spell. Third, for reasons unrelated to the job search process, 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, $\delta_{1h} > \delta_{1l} > 0$. The layoff rate is modeled as a possibly worker skill correlated random effect, such that worker i's layoff rate, δ_1^i, is takes value $\delta_{1j}$with probability $\Delta_j(h_i) = \Pr\left(\delta_1^i = \delta_{1j} | h_i = h\right)$.

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). An alternative bargaining protocol is presented in Yamaguchi (2006). In both cases, it is assumed that the worker can use a contact with one employer as a threat point in a bargaining game with another. An employment contract can only be re-negotiated by mutual consent. If the worker is unemployed, then the value of unemployment will be the worker's threat point. The detailed bargaining argument is presented in the appendix.

An employment contract consists of a worker's wage level and search intensity. For $j \in l, h$,

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

denote by $\tilde{V}_j(h,p,w,s)$ a skill level h, layoff rate δ_{1j} worker's asset value of a job with a type p firm and employment contract (w,s). The outcome of the employment contract bargaining as described in the appendix is such that the agreed upon search intensity maximizes the joint surplus of the match and the wage then dictates the surplus split. Hence, the search intensity depends only on the (h,δ_{1j},p) triple,

$$s_j(h,p) = \arg\max_{s\geq 0} \tilde{V}_j\big(h,p,f(h,p),s\big), \tag{2.3}$$

where $j \in l,h$.

If the worker is unemployed, the outside option in the bargaining is the value of unemployment. Denote by $\big(w_j^0(h,p), s_j(h,p)\big)$ the employment contract of a skill level h, layoff rate δ_{1j} worker who was hired out of unemployment by a type p firm. For $j \in l,h$, it satisfies,

$$\tilde{V}_j(h,p,w_j^0(h,p),s_j(h,p)) = \beta\tilde{V}_j\big(h,p,f(h,p),s_j(h,p)\big) + (1-\beta)V_j^0(h), \tag{2.4}$$

where $V_j^0(h)$ is the asset value of unemployment for a skill level h, layoff rate δ_j worker. β is the worker's bargaining power.

If an employed worker receives an outside offer, the worker will go to the most productive firm and the outcome is as if the worker bargains with the most productive firm with a threat point of going to the less productivity firm and receive full surplus. Denote by p and q the types of the two firms, where $p \geq q$. If the two firms are of equal productivity, the worker stays with the current firm. Denote the resulting wage by $w_j(h,q,p)$. It satisfies,

$$\tilde{V}_j(h,p,w_j(h,p,q),s_j(h,p)) = \beta\tilde{V}_j\big(h,p,f(h,p),s_j(h,p)\big) + (1-\beta)\tilde{V}_j\big(h,q,f(h,q),s_j(h,q)\big). \tag{2.5}$$

Denote by $q_j(h,p,w)$ the highest type a worker who is currently employed by a type p firm at wage w such that the meeting has no impact on the current employment terms. It is defined implicitly by,

$$w = w_j\big(h,q_j(h,p,w),p\big). \tag{2.6}$$

This implies that,

$$\tilde{V}_j\left(h,p,w,s_j(h,p)\right) = \beta\tilde{V}_j\big(h,p,f(h,p),s_j(h,p)\big) + (1-\beta)\tilde{V}_j\big(h,q,f(h,q),s_j(h,q)\big), \tag{2.7}$$

where $q = q_j(h, p, w)$. Equation (2.6) also illustrates that any arbitrary wage received in a match with a type p firm can be viewed as the outcome of bargaining with the type p firm given the outside option to match with a type $q_j(h, p, w)$ firm. Hence, a sufficient statistic for an employed worker's state is the record of the types of the two most productive employers that the worker has met during the past employment spell. Mostly, the value functions in the following will be stated in these terms rather than through an explicit wage. Specifically define $V_j(h, q, p) = \tilde{V}_j\big(h, p, w_j(h, q, p), s_j(h, p)\big)$.

The bargaining process assumes that search intensities can be contracted upon, which ensures that the jointly efficient search intensity level is implemented by the contract. In the alternative case, where the worker cannot commit to a search intensity choice, 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. This is because part of the incentive to generate outside offers now includes rent extraction from the current match. Presumably in this case the firm-worker pair would attempt to design an efficient mechanism to implement the jointly efficient search choice. We do not include such a mechanism in the analysis, but we conjecture that absent side payments, it will take the form of a back loaded wage profile that resembles a process where the worker initially "buys" the job in return for a subsequent receipt of the entire surplus. Such a mechanism would in part mimic the existing wage process in this paper where wages on the job increase in response to increases in the bargaining position. It is our view, that a change of the bargaining process in the direction of assuming non-commitment to the search choice will possibly impact the paper's implications for on-the-job wage growth, but insofar that the optimal mechanism is close to efficient, it will not in any substantive way change the paper's implications for the worker's search intensity. To provide robustness to this issue, our empirical analysis takes minimal empirical evidence from within job wage profiles. Rather, the core of the identification strategy is based on state conditional spell durations.

It is assumed that an unemployed type h worker receives an income stream $f(h, b)$. The

Bellman equation for the value of unemployment is given by,

$$\begin{aligned} rV_j^0(h) &= \max_{s\geq 0}\left\{f(h,b)-c(s)+(\mu+\kappa s)\lambda(\theta)E\left[\max[0,\tilde{V}_j(h,w_j^0(h,p),p)-V_j^0(h)]\right]\right\} \\ &= \max_{s\geq 0}\left\{f(h,b)-c(s)+(\mu+\kappa s)\lambda(\theta)\int_{R_j(h)}^{\bar{p}}\beta\left[V_j(h,p',p')-V_j^0(h)\right]d\Gamma(p')\right\}, \end{aligned} \tag{2.8}$$

where r is the interest rate, $\Gamma(p)$ is the cumulative firm type vacancy distribution, and $R_j(h)$ is the type h reservation productivity level defined by,

$$V_j\big(h,R_j(h),R_j(h)\big)=V_j^0(h). \tag{2.9}$$

It is straightforward to prove that $V_j(h,p,p)$ is monotonically increasing in p which establishes the reservation property of the model; that a type skill level h worker will agree to match with any employer above the productivity threshold level, $R_j(h)$. Applying integration by parts and the envelope theorem, equation (2.8) can be restated as,

$$rV_j^0(h)=\max_{s\geq 0}\left\{f(h,b)-c(s)+(\mu+\kappa s)\lambda(\theta)\int_{R_j(h)}^{\bar{p}}\frac{\beta f_p'(h,p')[1-\Gamma(p')]dp'}{r+\delta_j+\beta s_j(h,p')\lambda(\theta)[1-\Gamma(p')]}\right\}, \tag{2.10}$$

where $\delta_j\equiv\delta_0\lambda(\theta)+\delta_{1j}$.

The value of employment with a type p firm at wage $w_j(h,q,p)$ and search intensity $s_j(h,p)$ is given by,

$$\begin{aligned} rV_j(h,q,p) = {} & w_j(h,q,p)-c\big(s_j(h,p)\big)+\delta_{1j}\left[V_j^0(h)-V_j(h,q,p)\right]+ \\ & s_j(h,p)\lambda(\theta)\left[\int_p^{\bar{p}}\left[V_j(h,p,p')-V_j(h,q,p)\right]d\Gamma(p')+\int_q^p\left[V_j(h,p',p)-V_j(h,q,p)\right]d\Gamma(p')\right]+ \\ & \delta_0\lambda(\theta)\left[\Gamma(R_j(h))V_j^0(h)+\int_{R_j(h)}^{\bar{p}}V_j(h,R_j(h),p')d\Gamma(p')-V_j(h,q,p)\right]. \end{aligned} \tag{2.11}$$

Integration by parts and the envelope theorem allows the expression to be re-written as,

$$\begin{aligned} (r+\delta_j)V_j(h,q,p) = {} & w_j(h,q,p)-c(s_j(h,p))+\delta V_j^0(h)+ \\ & \delta_0\lambda(\theta)\int_{R_j(h)}^{\bar{p}}\frac{\beta f_p(h,p')\big(1-\Gamma(p')\big)dp'}{r+\delta_j+\beta s_j(h,p')\lambda(\theta)\big(1-\Gamma(p')\big)}+ \\ & s_j(h,p)\lambda(\theta)\int_p^{\bar{p}}\frac{\beta f_p(h,p')\big(1-\Gamma(p')\big)dp'}{r+\delta_j+\beta s_j(h,p')\lambda(\theta)\big(1-\Gamma(p')\big)}+ \\ & s_j(h,p)\lambda(\theta)\int_q^p\frac{(1-\beta)f_p(h,p')\big(1-\Gamma(p')\big)dp'}{r+\delta_j+\beta s_j(h,p')\lambda(\theta)\big(1-\Gamma(p')\big)}. \end{aligned} \tag{2.12}$$

The detailed derivation of equation (2.12) can be found in the appendix.

2.1 The search choices

The employment state conditional search intensity is found by use of equations (2.3) and (2.8). Together with equation (2.12), they imply the first order conditions,

$$c'(s_j^0(h)) = \kappa\lambda(\theta)\int_{R_j(h)}^{\bar{p}} \frac{\beta f_p(h,p')\big(1-\Gamma(p')\big)dp'}{r+\delta_j+\beta s(h,p')\lambda(\theta)\big(1-\Gamma(p')\big)} \tag{2.13}$$

$$c'\big(s_j(h,p)\big) = \lambda(\theta)\int_{p}^{\bar{p}} \frac{\beta f_p(h,p')\big(1-\Gamma(p')\big)dp'}{r+\delta_j+\beta s_j(h,p')\big(1-\Gamma(p')\big)}. \tag{2.14}$$

By convexity of $c(\cdot)$, differentiation of equation (2.14) with respect to p immediately yields that $s_j(h,p)$ is monotonically decreasing in p, $\forall h$. Furthermore, $s_j(h,\bar{p}) = 0$, $\forall h$. Lemma 1 establishes that the search intensity is strictly increasing in the worker type h if the production function is strictly supermodular. Also, if the production function has no complementarities between worker and firm types, then the search intensity is identical across worker types.

Lemma 1. *For any pair $(h_0,h_1) \in [\underline{h},\bar{h}] \times [\underline{h},\bar{h}]$ such that $h_0 < h_1$, and for all $p \in [b,\bar{p})$,*

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

For any $h \in [\underline{h},\bar{h}]$, $s_j(h,\bar{p}) = 0$.

Proof. See Lentz (2010). □

The reservation productivity level $R_j(h)$ defined in equation (2.9) is characterized in Lemma 2

Lemma 2. *For any $h \in [\underline{h},\bar{h}]$, if $\kappa = 1$ and $\mu = \delta_0$ then $R_j(h) = b$, and if $\kappa > 1$ and $\mu > \delta_0$ then $\bar{p} > R_j(h) > b$. Furthermore, if for any pair $(h_0,h_1) \in [\underline{h},\bar{h}]$ and for all $p \in [b,\bar{p}]$ $f_p(h_0,p) = f_p(h_1,p)$, then $R_j(h_0) = R_j(h_1)$.*

Proof. See Lentz (2010). □

In the case where $\kappa > 1$, an obvious question of interest is how $R_j(h)$ varies with h. Lemma 2 states that in the absence of production function complementarities, $R_j(h)$ is identical across worker skill levels. If $\rho \neq 1$ the model includes many of the complications associated with the classic stopping problem as analyzed in Shimer and Smith (2000). Specifically, it is straightforward to produce examples where $R_j(h)$ is not monotonically increasing in h even if the production function is supermodular.

2.2 Solving for the wage

With a solution for $s_j(h,p)$ in hand, one can immediately obtain values for the Bellman equation for the following states,

$$\begin{aligned}(r+\delta_j)V_j(h,p,p) = & f(h,p) - c(s_j(h,p)) + \delta_j V_0(h) + \\ & \delta_0 \lambda(\theta) \int_{R_j(h)}^{\bar{p}} \frac{\beta f_p(h,p')\big(1-\Gamma(p')\big)dp'}{r+\delta_j+\beta s_j(h,p')\lambda(\theta)\big(1-\Gamma(p')\big)} + \\ & s_j(h,p)\lambda(\theta) \int_{p}^{\bar{p}} \frac{\beta f_p(h,p')\big(1-\Gamma(p')\big)dp'}{r+\delta_j+\beta s_j(h,p')\lambda(\theta)\big(1-\Gamma(p')\big)}, \ \forall p \geq b. \end{aligned} \tag{2.15}$$

The value of the unemployed state is,

$$rV_j^0(h) = f(h,b) - c\big(s_j^0(h)\big) + (\mu + \kappa s_j^0(h))\lambda(\theta) \int_{R_j(h)}^{\bar{p}} \frac{\beta f_p(h,p')\big(1-\Gamma(p')\big)dp'}{r+\delta_j+\beta s_j(h,p')\lambda(\theta)\big(1-\Gamma(p')\big)}. \tag{2.16}$$

Given the wage determination mechanism in equation (2.5) combined with equation (2.15), one obtains,

$$V_j(h,q,p) = \beta V_j(h,p,p) + (1-\beta)V_j(h,q,q). \tag{2.17}$$

It then directly follows from equation (2.12) that,

$$\begin{aligned} w_j(h,q,p) = & (r+\delta_j)V_j(h,q,p) + c\big(s_j(h,p)\big) - \delta_0\lambda(\theta) \int_{R_j(h)}^{\bar{p}} \frac{\beta f_p(h,p')\big(1-\Gamma(p')\big)dp'}{r+\delta_j+\beta s_j(h,p')\lambda(\theta)\big(1-\Gamma(p')\big)} - \\ & \delta_j V_j^0(h) - s_j(h,p)\lambda(\theta)\Bigg[\int_{p}^{\bar{p}} \frac{\beta f_p(h,p')\big(1-\Gamma(p')\big)dp'}{r+\delta_j+\beta s_j(h,p')\lambda(\theta)\big(1-\Gamma(p')\big)} + \\ & \int_{q}^{p} \frac{(1-\beta) f_p(h,p')\big(1-\Gamma(p')\big)dp'}{r+\delta_j+\beta s_j(h,p')\lambda(\theta)\big(1-\Gamma(p')\big)}\Bigg]. \end{aligned} \tag{2.18}$$

2.3 Vacancy creation

Each firm is characterized by a permanent productivity p that applies to all of its matches. Firm types are distributed according to the cumulative distribution function $\Phi(\cdot)$. A firm's total output is the sum of the output of all its matches. Hence, a firm with n workers produces,

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

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 operation,

$$\begin{aligned} J_0(p) = \max_{\nu \geq 0} \Bigg[-c_\nu(\nu) + \eta\nu \sum_{j \in l,h} \int_{\underline{h}}^{\bar{h}} \Bigg\{ \left[V_j(h, p, p) - V_j\left(h, R_j(h), p\right) \right] \Lambda_j^0(h) + \\ \int_{R_j(h)}^{p} \left[V_j(h, p, p) - V_j\left(h, p', p\right) \right] \Lambda_j\left(h, p'\right) dp' \Bigg\} dh \Bigg], \end{aligned} \tag{2.19}$$

where

$$\Lambda_j(h, p) = \frac{\Delta_j \left(1 - u_j\right) s_j(h, p) g_j(h, p)}{\sum_{j' \in l,h} \Delta_{j'} \int_{\underline{h}}^{\bar{h}} \left\{ u_{j'}[\mu + \kappa s_{j'}^0(h')] v_{j'}(h') + \left(1 - u_{j'}\right) \int_b^{\bar{p}} [\delta_0 + s_{j'}(h', p')] g_{j'}(h', p') dp' \right\} dh'}$$

and

$$\Lambda_j^0(h) = \frac{\Delta_j \left\{ u_j[\mu + \kappa s_j^0(h)] v_j(h) + \left(1 - u_j\right) \delta_0 \int_b^{\bar{p}} g_j(h, p) dp \right\}}{\sum_{j' \in l,h} \Delta_{j'} \int_{\underline{h}}^{\bar{h}} \left\{ u_{j'}[\mu + \kappa s_{j'}^0(h)] v_{j'}(h) + \left(1 - u_{j'}\right) \int_b^{\bar{p}} [\delta_0 + s_{j'}(h, p)] g_{j'}(h, p) dp \right\} dh}.$$

Conditional on a meeting, $\Lambda_j(h, p)$ is the likelihood of meeting an employed skill level h, layoff rate δ_j worker who is currently employed with a productivity p firm. $\Lambda_j^0(h)$ 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. The expressions reflect a proportionality assumption in matching; a worker is represented in the pool of searchers proportionally to his or her search intensity. $g_j(h, p) = \int_b^p g_j(h, q, p) dq$ is the density of matches between skill h, layoff rate δ_j workers and productivity p firms, where $g_j(h,q,p)$ is the joint pdf of matches. u_j is layoff rate conditional unemployment rate and $\Upsilon_j(h)$ is the CDF of worker skill in the layoff rate conditional unemployment pool. $\Delta_j \equiv \int_h \Delta_j(h)\, d\Psi(h)$, $j \in l,h$ is the probability that a given worker has layoff rate δ_j,

It follows from equation (2.19) that the first order condition on the productivity conditional vacancy intensity choice is,

$$c_v'(v(p)) = \eta(1-\beta)\sum_{j\in l,h}\int_{\underline{h}}^{\bar{h}}\Big\{\left[V_j(h,p,p) - V_j(h,R_j(h),R_j(h))\right]\Lambda_j^0(h) + \int_{R_j(h)}^{p}\left[V_j(h,p,p) - V_j(h,p',p')\right]\Lambda_j(h,p')\,dp'\Big\}dh\Big]. \tag{2.20}$$

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,

$$\eta(p) = \eta v(p)\sum_{j\in l,h}\int_{\underline{h}}^{\bar{h}} I(R_j(h)\le p)\left[\Lambda_j^0(h) + \int_{R_j(h)}^{p}\Lambda_j(h,p')\,dp'\right]dh. \tag{2.21}$$

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

$$d(p) = \sum_{j\in l,h}\Delta_j\delta_j + \lambda(\theta)\left[1-\Gamma(p)\right]\frac{\sum_{j\in l,h}\Delta_j\int_{\underline{h}}^{\bar{h}} s_j(h,p)\,g_j(h,p)\,dh}{\sum_{j\in l,h}\Delta_j\int_{\underline{h}}^{\bar{h}} g_j(h,p)\,dh}. \tag{2.22}$$

2.4 Steady state

The steady state condition on the joint CDF of matches, $G_j(h,q,p)$, is,

$$\begin{aligned}\delta G_j(h,q,p) = &\int_{\underline{h}}^{h} I(R_j(h')\le q)\lambda(\theta)\left[\Gamma(p)-\Gamma(R_j(h'))\right]\Big[\frac{u_j}{1-u_j}[\mu+\kappa s_j^0(h')]v_j(h') + \\ &\delta_0\int_{R_j(h')}^{\bar{p}}\int_{q'}^{\bar{p}} g_j(h',q',p')dp'dq'\Big]dh' - \\ &\int_{\underline{h}}^{h}\int_{R_j(h')}^{q}\lambda(\theta)\Big\{(1-\Gamma(p))\int_{q'}^{q} s_j(h',p')dG_j(h',q',p') + \\ &(1-\Gamma(q))\int_{q}^{p} s_j(h',p')dG_j(h',q',p')\Big\},\end{aligned} \tag{2.23}$$

where $I(\cdot)$ is an indicator function that equals one if its expression is true, zero if false. The equation is a simple statement that the flows in and out of the $G_j(h,q,p)$ mass must balance in steady state. Equation (2.23) implies that the steady state unemployment rate for the population of layoff rate δ_{1j} workers satisfies,

$$u_j = \left[\int_{\underline{h}}^{\bar{h}} \left(1 + \frac{[1 - \Gamma(R_j(h'))][\mu + \kappa s_j^0(h')]\lambda(\theta)}{\delta_0 \lambda(\theta)\Gamma(R_j(h')) + \delta_{1j}} \right) d\Upsilon_j(h') \right]^{-1}. \tag{2.24}$$

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,

$$0 = \eta(p)\, m_{n-1}(p) + d(p)\,(n+1) m_{n+1}(p) - (\eta(p) + d(p)\, n)\, m_n(p), \tag{2.25}$$

for all $n \geq 1$ and p. It can be shown 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 must equal the number of deaths,

$$\eta(p) m_0(p) = d(p) m_1(p). \tag{2.26}$$

Furthermore, it is given that

$$\sum_{n=0}^{\infty} m_n(p) = m\phi(p), \tag{2.27}$$

where $\phi(p)$ is the firm productivity distribution pdf. Equations (2.25)-(2.27) imply that the type conditional firm size distribution $m_n(p)/(m\phi(p))$ is Poisson with arrival rate $\eta(p)/d(p)$,

$$m_n(p) = \left(\frac{\eta(p)}{d(p)} \right)^n \frac{1}{n!} \exp\left(-\frac{\eta(p)}{d(p)} \right) m\phi(p), \tag{2.28}$$

for all $n \geq 0$.

2.5 Steady state equilibrium

The equilibrium vacancy offer distribution is given by,

$$\Gamma(p) = \frac{\int_b^p \nu(p')\, d\Phi(p')}{\int_b^{\bar{p}} \nu(p')\, d\Phi(p')}. \tag{2.29}$$

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

$$\lambda(\theta) = \theta \eta(\theta), \tag{2.30}$$

where

$$\theta = \frac{m \int_b^{\bar{p}} \nu(p')\, d\Phi(p')}{\sum_{j \in l,h} \Delta_j \left[u_j \int_{\underline{h}}^{\bar{h}} [\mu + \kappa s_j^0(h)] d\Upsilon_j(h) + (1 - u_j) \int_{\underline{h}}^{\bar{h}} \int_b^{\bar{p}} [\delta_0 + s_j(h,p)] dG_j(h,p) \right]}. \tag{2.31}$$

Define the layoff rate conditional worker skill distribution, $\Psi_j(h)$, such that for any $j \in l, h$,

$$\Psi_j(h) = \frac{\int_{\underline{h}}^{h} \Delta_j(h')\, d\Psi(h')}{\int_{\underline{h}}^{\bar{h}} \Delta_j(h')\, d\Psi(h')}. \tag{2.32}$$

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

$$\Upsilon_j(h) = \frac{\int_{\underline{h}}^{h} \frac{\delta_0 \Gamma(R_j(h')) + \delta_{1j}/\lambda(\theta)}{\delta_0 \Gamma(R_j(h')) + \delta_{1j}/\lambda(\theta) + [1-\Gamma(R_j(h'))][\mu + \kappa s_j^0(h')]} d\Psi_j(h')}{\int_{\underline{h}}^{\bar{h}} \frac{\delta_0 \Gamma(R_j(h')) + \delta_{1j}/\lambda(\theta)}{\delta_0 \Gamma(R_j(h')) + \delta_{1j}/\lambda(\theta) + [1-\Gamma(R_j(h'))][\mu + \kappa s_j^0(h')]} d\Psi_j(h')}. \tag{2.33}$$

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 (2.9), (2.13), (2.14), (2.18), (2.24), (2.23), (2.29), (2.31), (2.32), and (2.33).

Lentz (2010) provides proof of existence and uniqueness of equilibrium in a somewhat simpler version of the model where vacancy intensities are constant across firm types.

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 $\mu = \delta_0$. Proposition 1 states sufficient conditions for positive sorting to occur. First, define the worker type conditional CDF of firm types by,

$$\Omega_j(p|h) = \frac{\int_b^p g_j(h,p')dp'}{\int_b^{\bar{p}} g_j(h,p')dp'}, \; j \in l, h. \tag{3.1}$$

One can then state the central characterization of sorting in steady state equilibrium.[5]

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

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

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

- $f_{hp}(h,p) < 0\,\forall(h,p) \Rightarrow \Omega_j(p|h_0) < \Omega_j(p|h_1)$ (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). □

It is worth emphasizing that the stochastic dominance results in Proposition 1 do not cleanly extend to the firm productivity conditional worker skill distribution,

$$\Omega_j(h|p) = \frac{\int_{\underline{h}}^{h} g_j(h',p)dh'}{\int_{\underline{h}}^{\bar{h}} g_j(h',p)dh'}. \tag{3.2}$$

4 Data

Our empirical analysis is conducted using a comprehensive Danish register-based Matched Employer-Employee (MEE) panel dataset. In this section we describe the data sources and the selection of the analysis data and present some basic summary statistics. We estimate our structural model using indirect inference which entails disciplining the structural parameter vector to fit a wide array of features of the data (in the indirect inference terminology: auxiliary statistics).

4.1 Data sources

The main building block of our MEE data is a dataset with individual level labor market spells recorded at a weekly frequency in 1985-2003 and effectively covering the entire Danish population aged 15-70. Workers and firms are identified via a unique person ID and firm and establishment IDs.[6] The spell data are constructed from administrative registers with information on public transfers, earnings as well as start and end dates for all jobs reported by firms to the Danish Tax Authorities, and mandatory employer pension contributions.

The spell data to identify five labor market states: employment, unemployment, retirement, self employment and non-participation. Employment spells are split up into firm-specific job

[6] In the empirical analysis we take the firm as the employing unit, but establishment IDs are retained to facilitate the merging of the spell data with other data sources; see below.

spells. By construction, nonparticipation 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, nonparticipation 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. We recode nonparticipation spells as unemployment spells.

The spell data is supplemented with background information on individuals from IDA, an *annual* population-wide (age 15-70) Danish MEE panel constructed and maintained by Statistics Denmark from several administrative registers.[7] We are able to merge the spell data with IDA data through individual, firm and establishment IDs. From IDA we retain information on age, gender, education, wages and establishment ownership.[8] Our wage measure is an estimate of the average hourly wage for jobs that are active in the last week of November. Since we obtain wage data from the IDA files, job spells that do not overlap with the last week of November in any year will have no wage information. Likewise, if the worker was unemployed in the last week of November there is no wage record for that worker in the corresponding year. We amend our estimation protocol to account for this data structure.

The final source of data that goes into the construction of our MEE analysis panel is a sequence of annual surveys on firms' financial statements collected by Statistics Denmark in 1999-2003 (the accounting data). These surveys were initiated in 1995 for a few industries and was gradually expanded until its 1999 coverage included most industries with a few exceptions such as agriculture, public services and parts of the financial sector (*source*: Statistics Denmark). The accounting data is merged onto to the spell data using firm and establishment identifiers and essentially contain the sampled firms' balance sheets from which we can compute value added and size of workforce in terms of full time equivalents (FTE). The survey has a rolling panel structure and covers approximately 9,000 firms which are selected based on the size of their November workforce.[9] Accounting data is available only at the level of the firm and cannot be attributed to

[7]IDA: Integreret Database for Arbejdsmarkedsforskning.

[8]Ownership allow us to distinguish between public and private establishments. Establishment-level information is available for all Danish establishments with at least one employee in the last week of November in each year.

[9]Specifically, no firms with 0-4 employees were sampled, while 10% of firms with 5-9 employees (included 1 year, excluded 9 years), 20% of firms with 10-19 employees (included 2 years, excluded 5 years), 50% of firms with 20-49

establishments or individual employees.

The merging of the three data sources, the spell data, the IDA data and the accounting data, yields a very comprehensive and long MEE panel with detailed individual- and firm-level information on employees and employers, labor market transitions and wages. At this stage, before selection of the analysis panel, the data contains 112,684,867 observations on 4,954,649 individuals, 454,395 firms (of which 21,290 firms carries accounting data at some point during 1999-2003) and 58,855,003 spells.

4.2 Sample selection and data manipulations

The selection criterias and data manipulations are imposed in order to rid the data of invald observations and to reduce unmodelled heterogeneity and other features of the data that our model is not designed to deal with. The following criterias are imposed:

- We truncate individual labor market histories at age 55 and discard any labor market history that pre-dates labor market entry. Labor market entry occurs at the end of the month where the individual graduate from his/her highest completed education. We discard all individuals who are observed in education or whose highest completed education code changes after their designated labor market entry. We also discard individuals with missing education information or with "too much" education relative to their age (which we define to mean that age minus years of education is less than 5 years).

- We discard individuals ever observed working (employment or selfemployment) in firms with missing firm ID or missing background information and individuals with gaps in their labor market histories. Firms with missing IDs have not been assigned any IDA information or accounting data

- We recode temporary unemployment spells as employment. A temporary unemployment spell is defined to be an unemployment spell of duration less than 13 weeks in-between employment spells with the same employer. Likewise, we recode unemployment spells of

employees (included 3 years, excluded 3 years) and 100% of firms with 50 or more employees were samples. In addition, all firms that reported a revenue exceeding DKK 100 mill. (DKK 200 mill. in the wholesale sector) in the previous year were sampled. Our estimation precedure takes this sampling scheme into account.

duration 2 weeks or less in between two employment spells with different employers as employment.

- We select the period 1994-2003 for our analysis. Our structural model relies heavily on firms and workers being characterized by firm and worker fixed effects. We believe the fixed effect assumptions are less restrictive in shorter panels.

- We discard all workers ever observed in employment in the public sector, in selfemployment, in retirement or in an industry for which we do not have any accounting data information.[10] Hence, our analysis panel contains individual labor market histories characterized by three states: Private sector employment, unemployment and nonparticipation.

- We trim the annual individual hourly wage and the (non-employment weighted) hourly value added distributions at the 1st and 99th percentiles. Moroever, we trend hourly wages and hourly value added to 2003 levels using their implicit deflators.[11]

Table 8 provides basic summary statistics on the final analysis data and also shows statistics for the first (1994) and last (2003) annual cross section in the data to assertain that the sample selection rules imposed on the data does not induce a significant amount of nonstationarity on the sample. The next section describes how we compute the moments of the data that we include in the estimation, including a detailed discussion of identification of the structural parameter vector.

5 Auxiliary models and structural identification

The auxiliary models we use in the estimation fall, broadly speaking, into one of three categories: moments that specifically identifies the sorting pattern in the data, moments that characterize labor market transitions, and moments that characterize cross sectional distributions of wages and productivity in the data.

[10] The public sector has already been discarded. In this step we effectively discard the agricultural sector and parts of the financial sector

[11] Note that value added is trended using the trend computed from employment weighted hourly value added.

Table 1: Summary statistics on the analysis data

	All years	1994	2003
Number of observations	6,930,436	669,896	713,643
Number of individuals	795,180	563,667	598,003
Number of job spells	1,738,981	501,054	520,889
Number of unemployment spells	611,464	168,841	192,753
Number of firms	117,942	53,571	58,249
Number of firms with accounting data information	19,671	-	8,265
Number of firm-years	560,291	-	-
Number of firm-years with accounting data information	39,780	-	-

5.1 Labor market sorting

A key question of interest is the identification of the production function, in particular the ρ coefficient which determines the sign and strength of complementarity between firm productivity and worker skill in production. In a partnership model, Eeckhout and Kircher (2008) argue that an identification strategy based on an Abowd et al. (1999) style wage fixed effects equation fails to identify sorting. Specifically, while one can identify the strength of sorting by comparing the within firm distribution of worker fixed effects to the full population, the strategy fails to identify whether sorting is positive or negative.[12] The lack of identification in the Eeckhout and Kircher (2008) setup follows from the result that even though worker skill and firm productivity map monotonically and strictly positively into match output, the match wage is not a monotone mapping in worker and firm types. This fundamentally breaks the link between estimated wage fixed effects and the identification of underlying worker and firm types.

We will argue that the Eeckhout and Kircher (2008) result can be generalized to our frame-

[12]Given the maintained identifying assumption of production function supermodularity, de Melo (2008) identifies the strength of the positive complementarity by the correlation between worker fixed effects within the firm.

work. The argument again rests on a result that wages may be non-monotone in agent types. However, it is not a trivial extension since the cause of the non-monotonicity differs substantially from that of the partnership model. The result provides some credibility to the argument that the Eeckhout and Kircher (2008) results extend significantly beyond their somewhat specialized setup. We subsequently offer an identification strategy that within our framework does identify both the strength and sign of sorting.

5.1.1 The wage function

Abowd et al. (1999) assume a log wage equation where worker and firm fixed effects enter additively,

$$w_{it} = x_{it}\beta + \chi_i + \varphi_{J(i,t)} + \varepsilon_{it}, \tag{5.1}$$

where $J(i,t)$ is the firm ID that worker i is matched with at time t, x_{it} is the set of worker i characteristics at time t, and χ_i and φ_j are the worker and firm fixed effects. 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, $cor[\chi_i, \varphi_{J(i,t)}]$, identify sorting in the matching between worker skill and firm productivity? Eeckhout and Kircher (2008) 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 mis-specified with respect to the worker and firm heterogeneity contributions to wages. Specifically, wages are generally not monotonically increasing in skill and productivity.

It is well known that given the wage posting setup in Postel-Vinay and Robin (2002), wages can initially decrease as a worker moves from a less to a more productive firm if the move is associated with an expectation of a higher wage growth rate. This is the key intuition for why worker skill conditional wages can be non-monotone in firm productivity in the model. Firm productivity conditional wages can furthermore also be non-monotone in worker skill as a result of differential search intensities across worker skill levels and differential returns to job offer accumulation that both map into different wage growth expectations.

It is worthwhile to contrast the wage non-monotonicity result in this model with that of the classic partnership model. In the partnership model, the non-monotonicity extends to the agent's match value functions. If for example the equilibrium is characterized by positive sorting, a high type agent tends to be matched with another high type agent in equilibrium. A low type agent may find that even though match output would increase by matching with a high type agent relative to another low type, the outside option of the high type is so high that the low type would have to deliver enough surplus to make the match acceptable to the high type, that the low type agent would actually prefer to match with another low type.

Our model does not exhibit this feature. Any worker regardless of skill level always prefers to match with a higher productivity firm. Furthermore, a worker would always prefer to have more skill regardless of the firm they are matched with. This result is stated in Lemma 3. The non-monotonicity of wages in skill and productivity is a result of the feature that the productivity of today's firm impacts the growth rate of future wages which is driven by the accumulation of outside offers.

Lemma 3. *The worker's valuation of a match $V_j(h, q, p)$ is strictly increasing in all three arguments.*

Proof. By equation (2.17), the match value satisfies $V(h, q, p) = \beta V(h, p, p) + (1 - \beta)V(h, q, q)$. For notational convenience, define $V(h, p) \equiv V(h, p, p)$. By equation (.2) it is already established that $V_p(h, p) > 0$. Hence, to establish the result in Lemma 3, it only remains to establish that $V(h, p)$ is increasing in h. $V(h, p)$ can be written as,

$$\begin{aligned}
rV(h,p) &= f(h,p) - c\big(s(h,p)\big) + \big[\delta_1 + \delta_0\lambda(\theta)\Gamma(R(h))\big]V_0(h) + s(h,p)\lambda(\theta)\int_p^{\bar{p}} V(h,p,p')d\Gamma(p') \\
&\quad + \delta_0\lambda(\theta)\int_{R(h)}^{\bar{p}} V(h,R(h),p')d\Gamma(p') - \big[\delta_0\lambda(\theta) + \delta_1 + s(h,p)\lambda(\theta)[1-\Gamma(p)]\big]V(h,p) \\
&= f(h,p) - c\big(s(h,p)\big) + \big[\delta_1 + \delta_0\lambda(\theta)[1-\beta+\beta\Gamma(R(h))]\big]V_0(h) \\
&\quad + \delta_0\lambda(\theta)\beta\int_{R(h)}^{\bar{p}} V(h,p')d\Gamma(p') + s(h,p)\lambda(\theta)\beta\int_p^{\bar{p}} V(h,p')d\Gamma(p') \\
&\quad - \big[\delta_0\lambda(\theta) + \delta_1 + \beta s(h,p)\lambda(\theta)[1-\Gamma(p)]\big]V(h,p).
\end{aligned}$$

By the assumption of jointly efficient search intensity, this can then be written as,

$$V(h,p) = \max_{s\geq 0, R\in[b,\bar{p}]} \left\{ \frac{f(h,p) - c(s) + \delta_1 V_0(h) + \beta s\lambda(\theta)\int_p^{\bar{p}} V(h,p')d\Gamma(p')}{r + \delta_0\lambda(\theta) + \delta_1 + \beta s\lambda(\theta)[1-\Gamma(p)]} \right.$$
$$\left. + \delta_0\lambda(\theta)\frac{V_0(h) + \beta\int_R^{\bar{p}}\left[V(h,p') - V_0(h)\right]d\Gamma(p')}{r + \delta_0\lambda(\theta) + \delta_1 + \beta s\lambda(\theta)[1-\Gamma(p)]} \right\}, \tag{5.2}$$

where

$$rV_0(h) = \max_{s\geq 0, R\in[b,\bar{p}]} \left\{ f(h,b) - c(s) + (\mu + \kappa s)\lambda(\theta)\beta\int_R^{\bar{p}} \left[V(h,p') - V_0(h)\right] d\Gamma(p') \right\}. \tag{5.3}$$

It is straightforward to show that the fixed point of the mapping in equation (5.3) satisfies,

$$V_0(h) = \max_{s\geq 0, R\in[b,\bar{p}]} \left\{ \frac{f(h,b) - c(s) + (\mu+\kappa s)\lambda(\theta)\beta\int_R^{\bar{p}} V(h,p')d\Gamma(p')}{r + (\mu+\kappa s)\lambda(\theta)\beta[1-\Gamma(R)]} \right\}. \tag{5.4}$$

This then establishes a unique solution to equation (5.3). Furthermore, inspection of equation (5.5) reveals that if $V(h,p)$ is increasing in h,then $V_0(h)$ is strictly increasing in h. Equation (5.2) is a contraction. Denote the mapping $T : \mathcal{F} \to \mathcal{F}$, where $\mathcal{F}$ is the set of bounded, continuous functions. For the purpose of showing that T maps the set of weakly increasing functions into the set of strictly increasing functions, consider any $h_0 < h_1$ where both h_0 and h_1 belong to the support of worker skill types. Now, take any function $V(h,p)$ that is weakly increasing in h for any p. Furthermore, let $s(h,p)$ be the maximizer of the right hand side of equation (5.2) for $V(h,p)$ and any h in the support of $\Psi(\cdot)$. Finally, let $V_0(h)$ be defined by equation (5.5) for the value of employment given by $V(h,p)$. It then follows that,

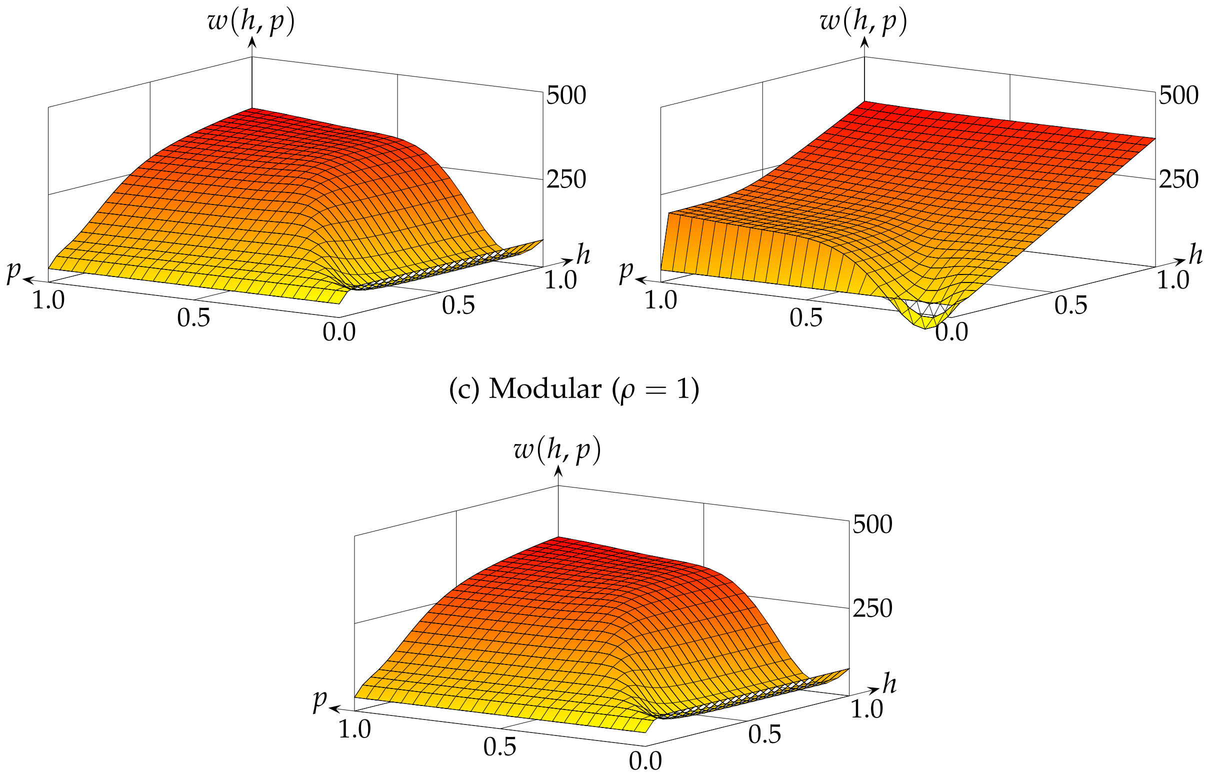
$$\begin{aligned}
(TV)(h_0,p) = {} & \frac{f(h_0,p) - c\big(s(h_0,p)\big) + \delta_1 V_0(h_0) + \beta s(h_0,p)\lambda(\theta)\int_p^{\bar{p}} V(h_0,p')d\Gamma(p')}{r + \delta_0\lambda(\theta) + \delta_1 + \beta s(h_0,p)\lambda(\theta)[1-\Gamma(p)]} \\
& + \delta_0\lambda(\theta)\frac{V_0(h_0) + \beta\int_{R(h_0)}^{\bar{p}}\left[V(h_0,p') - V_0(h_0)\right]d\Gamma(p')}{r + \delta_0\lambda(\theta) + \delta_1 + \beta s(h_0,p)\lambda(\theta)[1-\Gamma(p)]} \\
< {} & \frac{f(h_0,p) - c\big(s(h_0,p)\big) + \delta_1 V_0(h_1) + \beta s(h_0,p)\lambda(\theta)\int_p^{\bar{p}} V(h_1,p')d\Gamma(p')}{r + \delta_0\lambda(\theta) + \delta_1 + \beta s(h_0,p)\lambda(\theta)[1-\Gamma(p)]} \\
& + \delta_0\lambda(\theta)\frac{V_0(h_1) + \beta\int_{R(h_0)}^{\bar{p}}\left[V(h_1,p') - V_0(h_1)\right]d\Gamma(p')}{r + \delta_0\lambda(\theta) + \delta_1 + \beta s(h_0,p)\lambda(\theta)[1-\Gamma(p)]} \\
\leq {} & \frac{f(h_0,p) - c\big(s(h_1,p)\big) + \delta_1 V_0(h_1) + \beta s(h_1,p)\lambda(\theta)\int_p^{\bar{p}} V(h_1,p')d\Gamma(p')}{r + \delta_0\lambda(\theta) + \delta_1 + \beta s(h_1,p)\lambda(\theta)[1-\Gamma(p)]} \\
& + \delta_0\lambda(\theta)\frac{V_0(h_1) + \beta\int_{R(h_1)}^{\bar{p}}\left[V(h_1,p') - V_0(h_1)\right]d\Gamma(p')}{r + \delta_0\lambda(\theta) + \delta_1 + \beta s(h_1,p)\lambda(\theta)[1-\Gamma(p)]} \\
= {} & (TV)(h_1,p).
\end{aligned}$$

Hence, by the contraction mapping theorem, since T maps the set of function $V(h,p)$ that are increasing in h into the set of functions that are strictly increasing in h, it must be that the fixed point of equation (5.2) is strictly increasing in h. This establishes Lemma 3. $\square$

To illustrate the points, we simulate wages for the following model specification: Simplify vacancy creation so that $\nu(p) = 1$, $\forall p$. The model parameters are set as follows; $c_0 = 1$, $c_1 = 0.5$, $r = 0.05$, $\mu = 0.08$, $\delta_0 = 0.08$, $\delta_1 = 0.06$, $b = 0.1$, $\alpha = 0.5$, and $m = 0.1$. The worker skill distribution is a truncated Weibull with support $[0,1]$, shape parameter 1.5, scale parameter 0.45, and origin 0.05. The firm productivity distribution is also a truncated Weibull with support $[0,1]$, shape parameter 2.0, scale parameter 0.25, and origin 0.1. We will present results for different values of the worker's bargaining power. For any given choice of ρ, the production function scale parameter f_0 and the base offer arrival rate λ are set so as to obtain an equilibrium steady state unemployment rate of $u = 0.05$ and an average wage of $w = 180$.

Figure 5.1 presents wage function results for a worker bargaining power of $\beta = 0.2$. The wage function is defined as $w(h,p) \equiv \int_b^p w(h,q,p)g(h,q,p)dq$, where $g(h,q,p)$ is the steady state match pdf. Hence, $w(h,p)$ is the average wage realization for skill h worker with a productivity p firm. The figure presents the wage function for three different ρ values, representing the supermodular, submodular, and modular cases. All three cases illustrate that wages may be non-monotone in firm productivity. In particular there exists regions where the average wage realization for a given firm productivity type is decreasing in firm productivity. The higher productivity firm is valuable to the worker because it increases the worker's ability to extract surplus from the next high productivity firm the worker meets. The firm can consequently extract rents from the match through a lower wage. The worker accepts the lower wage with the expectation of high future wage growth and in these cases it so happens that the wage growth tends to be realized through a move to an even higher productivity firm which keeps the average wage realization with the current firm type low. The supermodular case also illustrates that wages can be non-monotone in worker skill. In this case, for relatively low firm productivity types, the search intensity choices and expected gains from upward movement on the offer ladder are so much higher for high skilled workers than low skilled workers that a given firm may be so much more valuable to a

Figure 5.1: Wage Function ($\beta = 0.2$)



(a) Supermodular ($\rho = -10$)

(b) Submodular ($\rho = 10$)

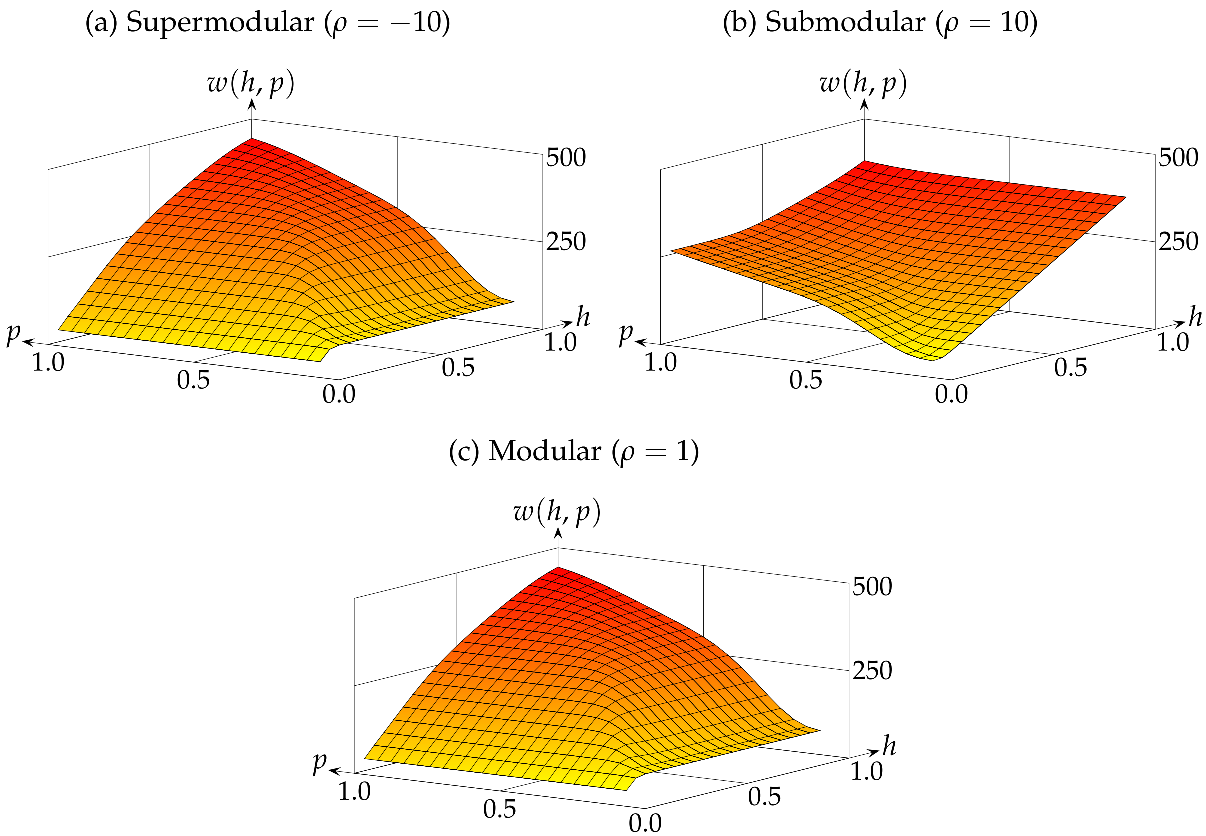
(c) Modular ($\rho = 1$)

Note: 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 $\mathrm{E}[w(h, p)] = 180.0$. The wage function is defined by $w(h, p) \equiv \int_b^p w(h, q, p) g(h, q, p) dq$.

high skilled worker than a low skilled worker in terms of increased wage growth expectations that the firm's rent extraction actually results in lower current wages for the high skilled worker.

Once the worker's bargaining power is increased, the non-monotonicity results begin to disappear. In the limit where $\beta = 1$, the productivity of the current firm does not impact future wage negotiations with other firms, because the worker extracts full match surplus regardless. In this case, the monotonicity results on the value function in Lemma 3 carry through to the wage function. Figure 5.2 shows the wage functions for the case where $\beta = 0.5$. Already at this point, the wage function is fundamentally reflecting the underlying characteristics of the production function $f(h, p)$ which is of course monotone in both h and p.

Figure 5.2: Wage Function ($\beta = 0.5$)



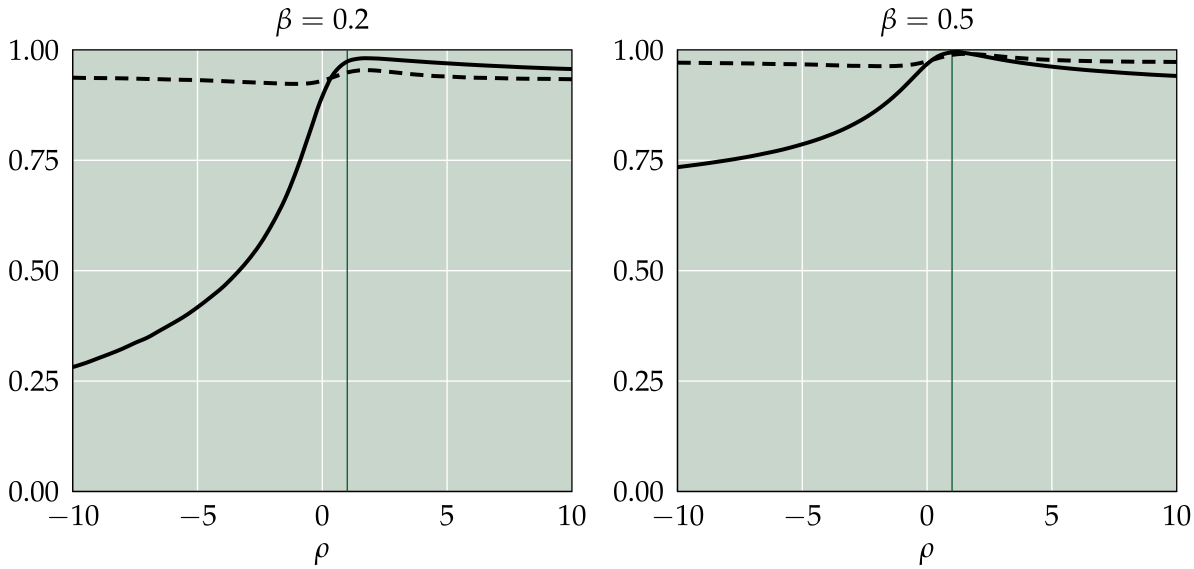
Note: For the given model specification, the production function scale parameter (f_0) and the base offer arrival rate (λ) are set such that the the steady state equilibrium solution satisfies $u = 0.05$ and $\mathrm{E}[w(h,p)] = 180.0$. The wage function is defined by $w(h,p) \equiv \int_b^p w(h,q,p)g(h,q,p)dq$.

In Figures 5.3 and 5.4 we relate estimates of worker and firm fixed effects from the wage equation (5.1) to the true underlying worker skill and firm productivity heterogeneity in simulations of steady state equilibria for different (ρ, β) combinations.

Figure 5.3 shows $\mathrm{cor}[\hat{\chi}, h]$ and $\mathrm{cor}[\hat{\varphi}, p]$. 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 does increase the correlation.

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 de-

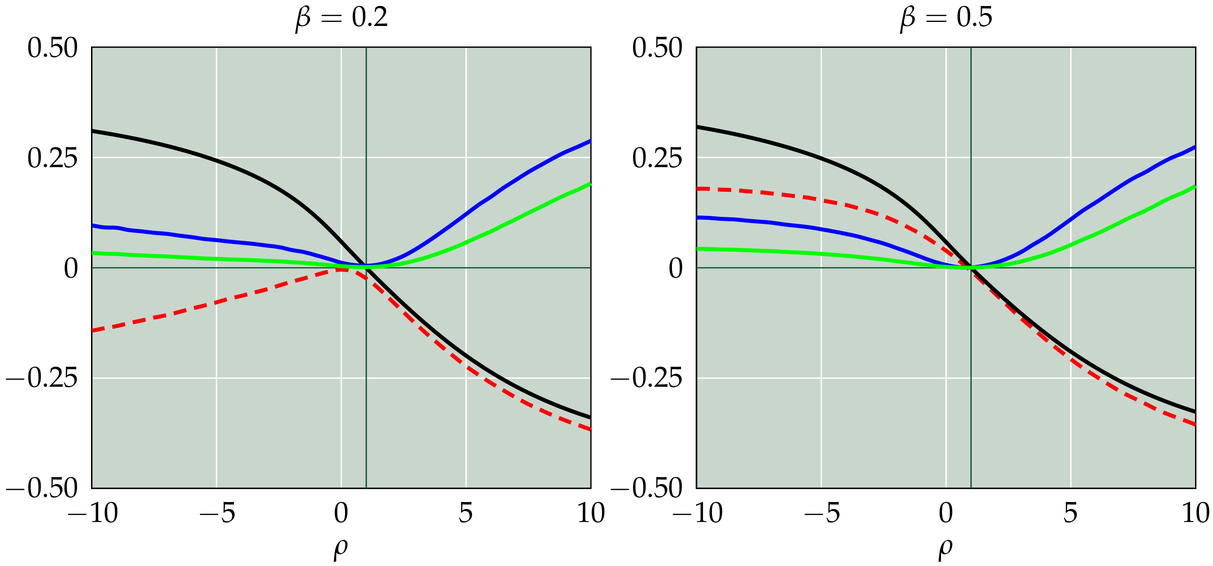
Figure 5.3: The correlation between wage fixed effects and true agent heterogeneity for given (ρ, β) combinations.



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 $\rho > 1$.

termination 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 5.1 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 iid over time error process, ε_{it} 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 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

Figure 5.4: The correlation between skill and productivity for given (ρ, β) combinations.



Note: The solid line is cor$[h, p]$. The dashed line is cor$[\chi_i, \varphi_{J(i,t)}]$. The blue solid line is cor$[\chi_i, \bar{\chi}_{-i}]$. The green line shows ν as defined in equation 5.6. 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 $\mathrm{E}[w(h, p)] = 180.0$. The dashed red line at $\rho = 1$ divides the model specifications with positive sorting for $\rho < 1$ and negative sorting for $\rho > 1$.

effects. This tendency towards negative bias in the correlation between worker and firm fixed effects in the wage equation is a general point emphasized by Postel-Vinay and Robin (????). Recent work by ?? allows for a match specific effect in the wage equation which could alleviate the within firm worker effect ranking problem problem somewhat.

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, cor$[\chi, h]$ is higher because wages are moving towards being monotone in worker skill and firm productivity.

Figure 5.4 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 $\rho = 1$. The figure also presents indicators for the distribution of worker wage fixed effects within firms relative to the overall population. One such moment suggested by de Melo (2008) is the correlation between the worker fixed effect and the average worker fixed effect of the co-workers within the firm at the time of the match. Worker i's average co-worker fixed effect at time t is given by,

$$\bar{\chi}_{-it} = \sum_{n \neq i} \mathcal{I}[J(n,t) = J(i,t)]\chi_n \Big/ \sum_{n \neq i} \mathcal{I}[J(n,t) = J(i,t)]. \tag{5.5}$$

A similar moment suggested by Eeckhout and Kircher (2008) is the population variance relative to the average within firm worker fixed effect variance,

$$F = \mathrm{E}_t\left[\frac{\mathrm{Var}_i[\chi_{it}]}{\mathrm{E}_j[\mathrm{Var}_i[\chi_{it}|J(i,t) = j]]}\right] - 1. \tag{5.6}$$

It is seen that when $\beta = 0.2$ and there is negative sorting, the correlation between wage equation worker and firm fixed effects, $\mathrm{E}_t[\mathrm{cor}[\chi_i, \varphi_{J(i,t)}]]$ is very close to equilibrium steady state $\mathrm{cor}[h, p]$. This is consistent with the results in Figure 5.4 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 $\mathrm{E}_t[\mathrm{cor}[\chi_i, \varphi_{J(i,t)}]]$ and $\mathrm{cor}[h, p]$ 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, $\mathrm{E}_t[\mathrm{cor}[\chi_i, \varphi_{J(i,t)}]]$ is negative both when sorting is positive and negative for this case.

In the case where $\beta = 0.5$, the fixed effects correlation $\mathrm{E}_t[\mathrm{cor}[\chi_i, \varphi_{J(i,t)}]]$ 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 $\mathrm{E}_t[\mathrm{cor}[\chi_i, \varphi_{J(i,t)}]]$ indicates that sorting between skill and productivity is positive. In general, the correlation coefficient between h and p is always greater than $\mathrm{E}_t[\mathrm{cor}[\chi_i, \varphi_{J(i,t)}]]$. It is also worth emphasizing that the oft 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.

The comparisons of the within firm distribution of χ_i relative to the population distribution, $\mathrm{E}_t[\mathrm{cor}[\chi_i t, \chi_{-it}]]$ and $\digamma$, both suggest that a positive observed value indicate the presence of sorting between worker skill and firm productivity, but not the sign of the sorting since both measures are positive for both positive and negative sorting.

Based on the results so far, we are short of an identification strategy for ρ. Of course, in practice, should the observed value of $\mathrm{E}_t[\mathrm{cor}[\chi_i, \varphi_{J(i,t)}]]$ be positive, identification would be obtained. However, as a general proposition we do not have a one-to-one mapping between empirical moments and ρ. In the following section, we propose an identification strategy that will identify not only strength of sorting but also the sign of it.

5.1.2 Identifying the type of sorting

As argued above, wage observations from matched employer-employee data can identify the strength of sorting. Only if observed $\mathrm{E}_t[\mathrm{cor}[\chi_i, \varphi_{J(i,t)}]]$ is positive does it also identify the type of sorting. In this section we propose an additional moment that will generally allow the identification of the type of sorting, positive or negative.

The identification strategy is focused on the correlation between inferred worker skill and unemployment durations. If sorting is positive, then high skill workers experience shorter unemployment spells than less skilled workers, and consequently the correlation between worker skill and unemployment duration should be negative. In the case where sorting is negative, less skilled workers search faster out of unemployment and the correlation should be positive. This argument uses the simple comparative statics of $s_0(h)$ with respect to h in the model.

Unemployment duration is easily observed in data. However, as shown in detail in the previous section, worker skill is not. In particular, firm type conditional wages are generally not necessarily monotone in worker skill. There is however a subset of matches where the observed wage does reveal the worker's skill level. Workers hired by the most productive firms directly out of unemployment receive the following wage,

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

Since by Lemma 3 $V_0(h)$ is strictly increasing in h it trivially follows that $w(h, b, \bar{p})$ is strictly increasing in h. Hence, the workers in the group hired directly out of unemployment into top firms can be ranked according to skill directly through the wage ranking. The identification strategy then reduces to correlating the observed wage within this group with the duration of the previous unemployment spell. Thus, if unemployment duration is negatively correlated with wages within the group, sorting is positive. And if the correlation is positive, then sorting is negative.

Implementation this identification strategy requires identification of top productivity firms.[13] Assuming the case of equally efficient search off and on the job, $\kappa = 1$, we identify a firm's type through observation of the composition of its worker inflow. Since all firm draw from the same worker pool, firms lower on the ladder will experience more rejections from workers employed with more productive firms. Hence, the fraction of a firm's labor inflow that comes from other firms is increasing in its position in the firm productivity hierarchy.[14] Conditional on a hire, the probability that the hire comes directly from another firm is given by,

$$\iota(p) = \frac{\sum_{j \in l,h} \Delta_j \left(1 - u_j\right) \int_0^1 \left[\delta_0 \int_0^1 g_j(h,p)dp + \int_0^p s_j(h,p')g_j(h,p')dp'\right] dh}{\sum_{j \in l,h} \Delta_j \int_0^1 \left[\left\{u_j[\mu + s_j^0(h)]v_j(h) + \left(1 - u_j\right)\delta_0 \int_0^1 g_j(h,p)dp\right\} + \left(1 - u_j\right)\int_0^p s_j(h,p')g_j(h,p')dp'\right] dh}$$

Straightforward differentiation yields,

$$\frac{\partial \iota(p)}{\partial p} > 0. \tag{5.8}$$

Therefore, this empirical measure can be used to identify the firm productivity ranking of firms.

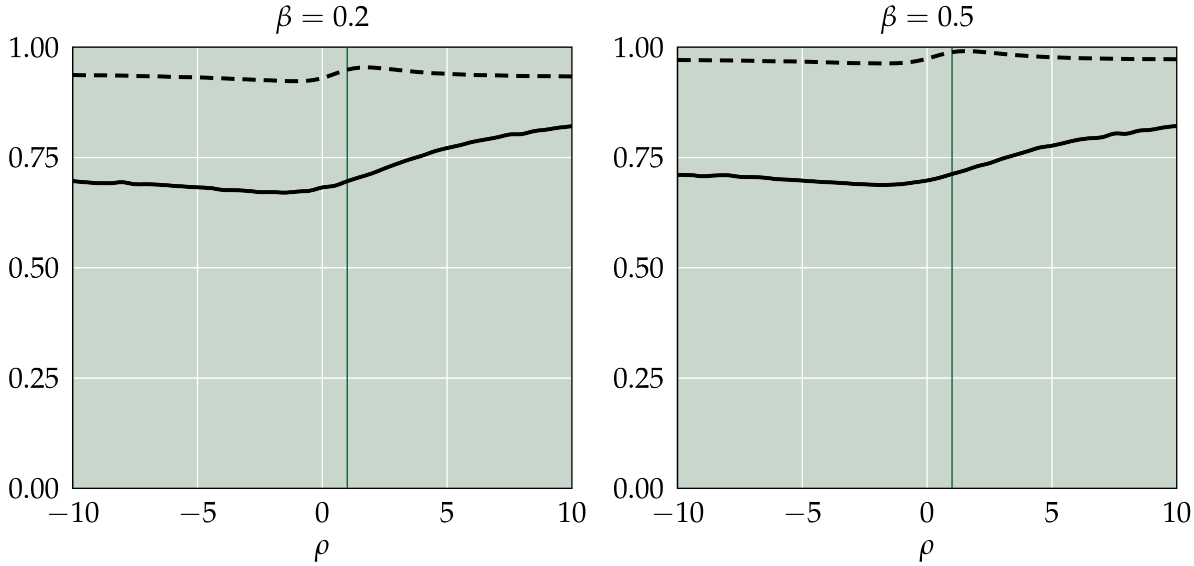
Figure 5.5 plots steady state equilibrium $\mathrm{cor}[\varphi, p]$ and $\mathrm{cor}\left[\iota(p), p\right]$ for different (β, ρ) combinations. It is seen that both empirical measures of the firm productivity ranking perform well with the wage equation firm fixed effect doing somewhat better than the inflow measure.

Figure 5.6 plots the correlation between wages and unemployment spell duration for the group of workers hired directly out of unemployment into the top 5% of firms, where the firm ranking is done either by the wage equation firm fixed effect or the job-to-job inflow relative to

[13]It is worth emphasizing that an obvious measure like the firm's labor productivity cannot be used for identification of firm type in the model, because a firm's average worker type need not be a monotone function of firm productivity regardless of the production function characteristics.

[14]In the case where $\kappa > 1$, firms may face offer rejection from unemployed workers as well, which complicates the theoretical identification argument. However the estimation is performed subject to $\kappa = 1$, and so the issue does not arise

Figure 5.5: Identification of firm productivity for given (ρ, β) combinations.

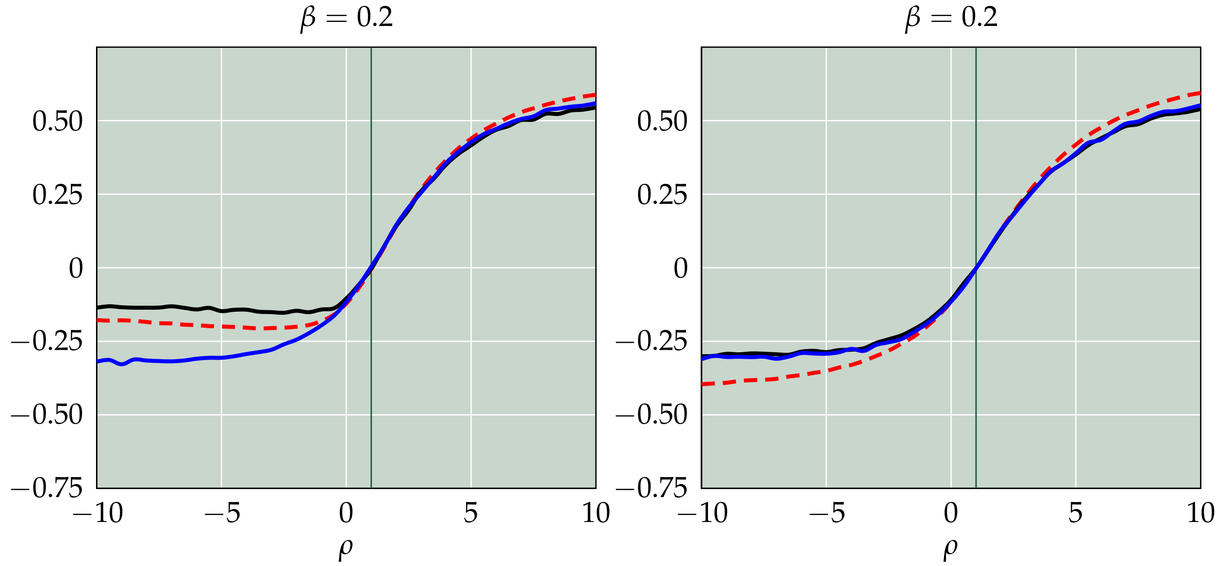


Note: The solid and dashed lines show $\text{cor}\left[\iota(p)/\left(\iota(p)+o(p)\right),p\right]$ and $\text{cor}[\hat{\varphi}, p]$, respectively. For the given model specification, the production function scale parameter (f_0) and the base offer arrival rate (λ) are set such that the the steady 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 $\rho > 1$.

outflow measure. Furthermore, the figure also shows the correlation between unemployment duration and the wage equation worker fixed effect. All three measures perform well in terms of identifying the sign of sorting - the correlation is negative when sorting is positive and vice versa. The use of the wage equation firm and worker fixed effects to make inference about the underlying skill and productivity indices works quite well for the cases that we have presented. But as emphasized before, we do not have proof that this will be the case for any model specification. The job-to-job inflow to outflow measure does identify the firm productivity ranking but in practice it is somewhat noisy.

The moments in Figure 5.6 in combination with the correlation between worker and firm fixed effects as well as the comparison of the within firm wage equation worker fixed effect relative to the overall population provide a successful foundation for identification not only of the presence of sorting between skill and productivity but also the sign of the sorting.

Figure 5.6: Identification of type of sorting for given (ρ, β) combinations.



Note: The solid black line is the corrrelation between wages and unemployment duration for workers hired directly into the top 5% of firms ranked by the job-to-job inflow to outflow measure. The solid blue line present the same correlation but using the wage equation firm fixed effect to identify the top 5% of the firms. The dashed red line shows the correlation between unemployment duration and the wage equation worker fixed effect. 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 $\mathrm{E}[w(h,p)] = 180.0$. The dashed red line at $\rho = 1$ divides the model specifications with positive sorting for $\rho < 1$ and negative sorting for $\rho > 1$.

Empirical implementation From the analysis data described in section 4 we select a data set containing all transitions from unemployment to employment, where an individual starting wage is available[15] and where the employing firm has hired at least three workers during the time it is observed in the data, and where at least one of these hires is not a job-to-job transition.[16] The first three panels in Table 2 tabulates moments of the distribution of $\iota(p))$, unemployment durations,[17] and log starting wages. The bottom three panels of Table 2 reports moments of the distribution of unemployment durations, starting wages and the covariance and correlation between unemployment durations and subsequent starting wages involving transitions into firms

[15]Starting wages are measured as the first wage observation on a job spell. If no wage observation is available within the first year of a job, the job is not used in the computation of the$\iota(p)$.

[16]These conditions ensures that very small firms, or firms that happen to only a few workers, but all through job-to-job transitions, are not treated as very productive firms with $\iota(p) = 1$.

[17]Unemployment durations are worker specific and are measured as the worker-specific average duration of any unemployment spell a worker may have in the data period. We only use non-left-censored unemployment spells that end in a transition to employment.

Table 2: Identifying the type of sorting: Empirical moments

Transitions involving firms with more than 3 hires and valid starting wages	
Mean $\iota(p)$	0.57
Standard deviation of $\iota(p)$	0.19
Skewness of $\iota(p)$	−0.54
Kurtosis of $\iota(p)$	4.04
Mean duration of preceding unemployment spell (in weeks)	63.66
Standard deviation of duration of preceding unemployment spell	68.49
Skewness of duration of preceding unemployment spell	2.60
Kurtosis of duration of preceding unemployment spell	11.97
Mean hourly starting wage (in log DKK)	5.10
Standard deviation of duration of hourly starting wage	0.28
Skewness of duration of hourly starting wage	0.46
Kurtosis of hourly starting wage	4.04
Transitions involving firms with $\iota(p)$ in upper 5$^{\text{th}}$ percentile	
Mean duration of preceding unemployment spell (in weeks)	56.86
Standard deviation of duration of preceding unemployment spell	65.45
Skewness of duration of preceding unemployment spell	2.71
Kurtosis of duration of preceding unemployment spell	12.76
Mean hourly starting wage (in log DKK)	5.21
Standard deviation of duration of hourly starting wage	0.30
Skewness of duration of hourly starting wage	0.21
Kurtosis of hourly starting wage	3.49
Covariance, duration of preceding unemployment spell and starting wage	−2.30
Correlation, duration of preceding unemployment spell and starting wage	−0.12

in the top 5 percent of the distribution of productivity (according to the firm level distribution of $\iota(p)$). We note here that the data exhibits a significant negative correlation of $-.12$ between unemployment durations and starting wages. In isolation, this indicates that the allocation of workers to firms in the labor market involves a degree of positive sorting.

5.2 Labor market transitions

Our on-the-job search model describes a labor market with two labor market states, employment and unemployment, and three possible labor market transitions: unemployment-to-employment (UE), employment-to-unemployment (EU) and employment-to-employment (EE). The structure of the model with endogenous search and vacancy intensity and the implied labor market sorting, and the presence of reallocation shocks, induces a complicated mapping between the structural parameter vector and observed labor market transitions. We include a rich set of moments ranging from simple unconditional transtion rates to more complicated wage and productivity dependent transtion rates.

All labor market transition moments are computed from estimation data extracted from the analysis panel dexcribed in section 4. The unit of observation for the labor market transition data is a spell. We select all unemployment spells not intiated in the final year of our data period and all employment spells not intiated in the final year of our data period with non-missing productivity index $\iota(p)$ and non-missing firm level average wage. Hence, we estimate the transition moments on a *flow* dataset (Ridder, 1723).

5.2.1 Kaplan-Meier hazard function estimates

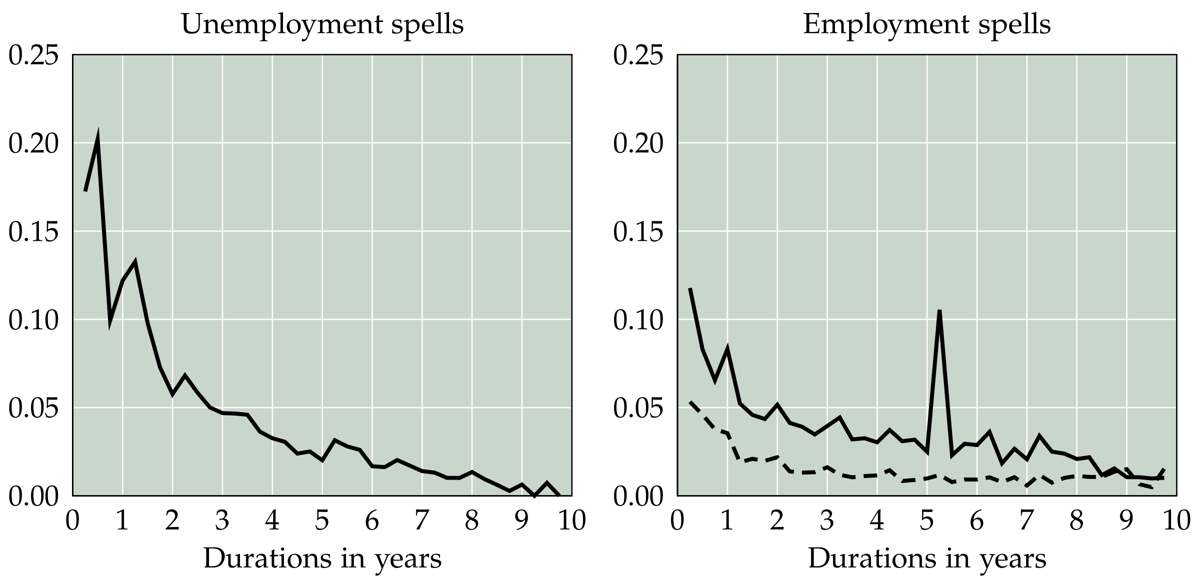
Worker and firm heterogeneity is central to our model. It is well known that the dynamic selection induced by heterogeneity in hazard rates introduces duration dependence in unconditional hazard functions. Hence, hazard functions related to both unemployment and employment spells contains information on worker and firm heterogeneity. We characterize duration dependence in part using Kaplan-Meier estimates of the unconditional UE-, EU- and EE-transition specific hazard functions, and in part using parametric duration models to be introduced below. Of course, the UE-transition specific hazard pertains to unemployment spells whereas EU- and EE-transition specific hazard pertains to employment spells.[18] The estimated hazard functions and the associated survivor functions are presented in Figures 5.7and5.8.

5.2.2 Parametric duration models

To further characterize the relationship between labor market transitions and type heterogeneity on both side of the labor market, we also include a set of (parametric) duration models in the set of auxilliary models. Specifically, we consider a Exponential hazard model frailty at the worker level that follows a unit mean Gamma distribution. The hazard model is in fact a parametric propotional hazard model with one (firm-specific) regressor, namely the firm level productivity

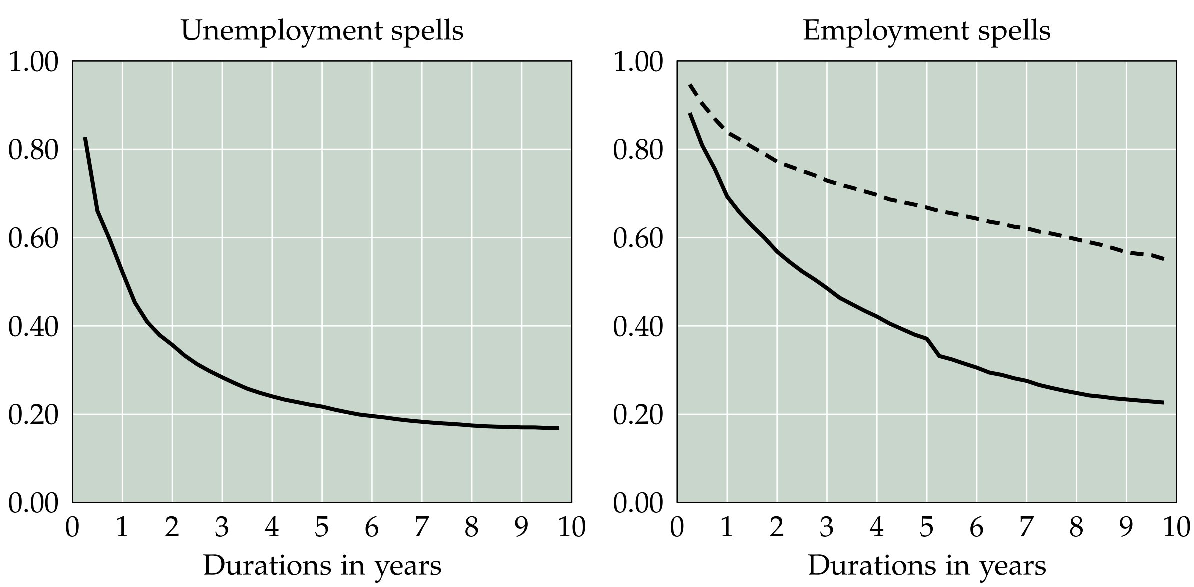
[18] In section4 we noted that the data in fact contains two types of nonemployment states, proper unemployment spells and a residual state which, for lack of a better term, we denote nonparticipation. When computing the UE transition rate we treat unemployment spells that end in a transition to nonemployment as censored. When computing EU transition rates, the empirical analogue to a job destruction rate, we consider both transitions to unemployment and transitions to nonemployment as a job destruction.

Figure 5.7: Quarterly Kaplan-Meier hazard rates (UE, EE- and EU-transitions)



Note: In the right panel the solid and dashed lines show EE- and EU separation rates, respectively.

Figure 5.8: Quarterly Kaplan-Meier survivor functions (UE, EE- and EU-transitions)



Note: In the right panel the solid and dashed lines show EE- and EU separation rates, respectively.

index $\iota(p)$. As argued further above, this statistic identifies a firm's rank in the type distribution of employers. Hence, this reduced form duration model captures the two-sided heterogeneity in the structural model, while imposing the assumption of no duration dependence in the underlying hazard function, an assumption that is also consistent with our structural model. Before we proceed to a more detailed description of the auxiliary duration model it should be noted that we estimate the model on the auxiliary dataset described above which implies that we (typically) observe multiple spells (be it unemployment or employment) per individual.[19]

We present the auxiliary duration model for employment spells. The corresponding model for unemployment durations is constructed analogously, but does not feature competing risks and does not have regressors. Let j index firms, i index workers, and let n index spells on worker i, and let T be a nonnegative random variable (duration) with realization t. A duration may be realized via two competing risks (job-to-job and job-to-unemployment transitions) or via censoring; that is $T = \min\{T_{EU}, T_{EE}, T_C\}$. The censoring process is assumed independent of T_{EU} and T_{EE}. Suppose that the latent durations T_{EU} and T_{EE} each follows a (conditional) exponential distribution; that is, it has survivor function

$$S(t_{ink}|\iota_{\mathrm{J}(i,n)}, u_{ik}) = \exp(-\lambda_{\mathrm{J}(i,n)k} u_{ik} t_{ink})$$

for $k \in EU, EE$ and where, $\mathrm{J}(i,n) = j$ if the n'th spell of worker i is a job at firm j. Define $\lambda_{jk} \equiv \exp(\lambda_{k0} + \lambda_{k1}\iota_j)$ where λ_{k0} and λ_{k1} are parameters, ι_j is the job-to-job flow index of firm j and u_{ik} is the realization of a random variable U_k which we will refer to as individual frailty. Frailty is individual specific, unobserved (to the econometrician) and is assumed to follow a unity mean Gamma distribution; that is,

$$f(u_{ik}) = \frac{\eta_k^{\eta_k} u_{ik}^{\eta_k - 1} e^{-\eta_k u_{ik}}}{\Gamma(\eta_k)}$$

[19] Estimating hazard models in the presence of unobserved heterogeneity is not free from identification problems, and there is a host of identification results in the econometrics literature, in particular pertaining to the class of Mixed Proportional Hazard Models in which our auxiliary duration model is also contained (see e.g. van den Berg (2001) for an overview of these results.). Much of this literature is devoted to establishing conditions under which true duration dependence can be *nonparametrically* identified from the dynamic selection induced by heterogeneity. Our duration model is fully parametric, and as such, is identified even if the concept of identification is somewhat mute here as the auxiliary duration model is embedded in an indirect inference procedure and is inherently mis-specified. We do require that the auxiliary statistic converges to a unique pseudo-true value (see Gourieroux, Monfort and Renault, (1993) for details).

where $\eta_k > 0$ is a parameter and $\Gamma(\cdot)$ is the imcomplete Gamma function. Frailties u_{ik} are taken to be independent across transitions, across individuals and of $\iota_{J(i,n)}$.

Since our parameterization of the auxiliary model implies that the competing risks share no parameters we may estimate transition specific models separately. Consider the likelihood contribution from individual i who is represented by N job spells. Define $d_{ink} = 1$ if the nth spell of worker i ends due to a type-k transition (that is, $d_{inEE} = d_{inEU} = 0$ indicates censoring). It is straight forward, albeit cumbersome, to show that for k transitions the log-likelihood contribution of worker i is

$$\ell_{ik} = \sum_{n=1}^{N} d_{ink} \left[\ln \sigma_k^2 + \ln \lambda_{J(i,n)} \right] - \left(1 + \sigma_k^{-2}\right) \ln \left\{ 1 + \sigma_k^2 \left[\sum_{n=1}^{N} \lambda_{J(i,n)} t_{in} \right] \right\} + \ln \Gamma \left(\left[\sum_{n=1}^{N} d_{ink} \right] + \sigma_k^{-2} \right) - \ln \Gamma \left(\sigma_k^{-2} \right) \quad (5.9)$$

where $\sigma_k^2 = \eta_k^{-1}$, the variance of U_k such that the case of no frailty arises as $\sigma_k^2 \to 0$. The fact that the likelihood function yields a closed form solution implies that estimation of the parameters of the auxiliary duration models can be carried out without resorting to computational intensive numerical intergration routines. To further lighten the computation burden in estimating the parametric duration models we estimate them using only 10 percent of the available workers. The estimated parameter vector is $(\lambda_{k0}, \lambda_{k1}, \sigma_k^2)$ for employment spells ($k \in EU, EE$) and $(\lambda_{UE0}, \sigma_{UE}^2)$ for unemployment spells. Table 3 presents estimated parametric duration models for employment and unemployment spells, with and without unobserved firm heterogeneity, and with no worker heterogeneity, non-shared worker heterogeneity, and shared worker heterogeneity.[20]

Figure5.9 plots employment hazard functions implied by our preferred parametric duration model specifications, “Firm heterogeneity, shared Gamma worker heterogeneity”. Figure5.9 plots hazard functions *unconditional* on worker heterogeneity, which generates duration dependence, but *conditional* on firm heterogeneity, i.e. the productivity index $\iota(p)$. Specifically, the twp top

[20] The likelihood function stated in equation (5.9) is for “shared” worker heterogeneity. That is, the worker-specific random effect is shared between all spells for that particular worker. When worker heterogeneity is “non-shared” the worker specific random effect is i.i.d. across workers and across spells for a given worker. Hence, non-shared worker heteroegeneity is in effect spell heterogeneity. The likelihood function simplifies somewhat when heterogeneity is non-shared compared to equation (5.9).

Table 3: Parameter estimates, Gamma Mixed Exponential duration models: Unemployment and employment spells

	λ_0	λ_1	σ^2
Unemployment spells, UE transitions (quaterly durations)			
No worker heterogenety ($\sigma^2 = 0$)	-2.213		
Non-shared Gamma worker heterogeneity	-1.455		1.145
Shared Gamma worker heterogeneity	-1.638		1.137
Employment spells, EU transitions (quarterly durations)			
No firm heterogeneity, no worker heterogeneity ($\lambda_1 = \sigma^2 = 0$)	-3.603		
Firm heterogeneity, no worker heterogeneity ($\sigma^2 = 0$)	-1.722	-2.898	
No firm heterogeneity, non-shared Gamma worker heterogeneity ($\lambda_1 = 0$)	-2.741		4.415
Firm heterogeneity, Non-shared Gamma worker heterogeneity	$-.348$	-4.000	2.964
No firm heterogeneity, shared Gamma worker heterogeneity ($\lambda_1 = 0$)	-3.212		2.683
Firm heterogeneity, shared Gamma worker heterogeneity	-1.497	-2.819	1.848
Employment spells, EE transitions (quarterly durations)			
No firm heterogeneity, no worker heterogeneity ($\lambda_1 = \sigma^2 = 0$)	-2.779		
Firm heterogeneity, no worker heterogeneity ($\sigma^2 = 0$)	-2.433	$-.501$	
No firm heterogeneity, non-shared Gamma worker heterogeneity ($\lambda_1 = 0$)	-2.132		1.420
Firm heterogeneity, non-shared Gamma worker heterogeneity	-1.441	-1.004	1.422
No firm heterogeneity, shared Gamma worker heterogeneity ($\lambda_1 = 0$)	-2.715		.651
Firm heterogeneity, shared Gamma worker heterogeneity	-1.985	-1.047	.701

panels in Figure5.9 plots destination specific employment hazards (EE- and EU-transitions) conditional on the productivity index equal to the 10th percentile, the 50th percentile and the 90th percentile in the distribution of the productivity index across workers. In the bottom panel, which shows a plot the unemployment hazard function, firm heterogeneity is not relevant. For comparions we superimpose the Kaplan-Meier estimate of the hazard function as presented in Figure 5.7 in Figure5.9 as well.

5.2.3 Wage dependent job separation rates

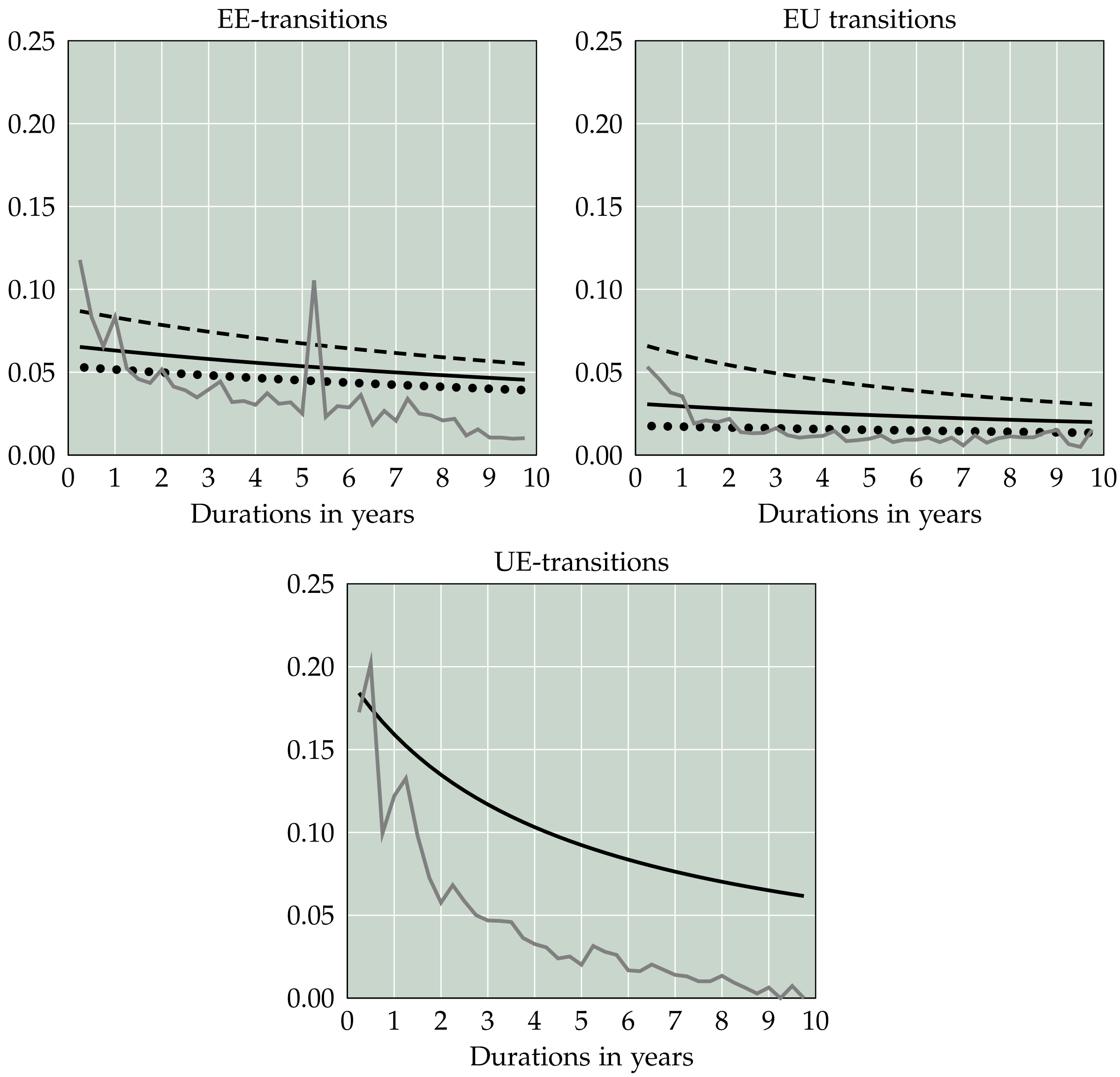
Our structural model also implies an intricate relationship between wages and labor market transitions between jobs and between employment and unemployment. Even though we are not able to show this formally, our simulation studies reveals that the strenght and sign of sorting in the labor market impacts on the wage dependence of the EE transition rates through the compositional effects it has on the wage distribution: low wage workers are predeominatly made up of workers of low type-h (with the caveats described in the preceding subsection).[21] Under positive sorting, low type-h workers also search less intensely, which implies that low-wage workers tend to leave their jobs at a relatively slow rate, inducing a relatively flat EE-transition rate-wage profile. Under negative sorting, low type-h workers search intensly and therefore leave their jobs at a faster rate. This implies a relatively steep and declining EE transition rate-wage profile.

For that reason we condition the structural estimation on a set of nonparametric regressions of EU- and EE-transition (and total job separation indicators) indicators on wages in the moment vector. Naturally, here we only use employment spells. Let ι^{EE} and ι^{EU} be quarterly EU- and EE-transition indicators and let $\hat{G}(w)$ be the empirical distribution function of a wage measure w. We then (nonparametrically) regress ι^{EE}, respectively ι^{EU}, onto $\hat{G}(w)$ and use $\hat{\mathrm{E}}(\iota^{EE}|\hat{G}(w) = a)$, respectively $\hat{\mathrm{E}}(\iota^{EU}|\hat{G}(w) = a)$, as moments to be matched where a takes on ten equi-distanced values in $[0, 1]$.[22] Again, these regressions are computed using flow data, i.e. in the population

[21] Indeed, it is the fact that wages are not monotonic in worker and firm types that prevents us from obtaining an analytical relationship between the EE-transition rate and wage paid that can be exploited in a formal identification analysis of the strength and sign of labor market sorting.

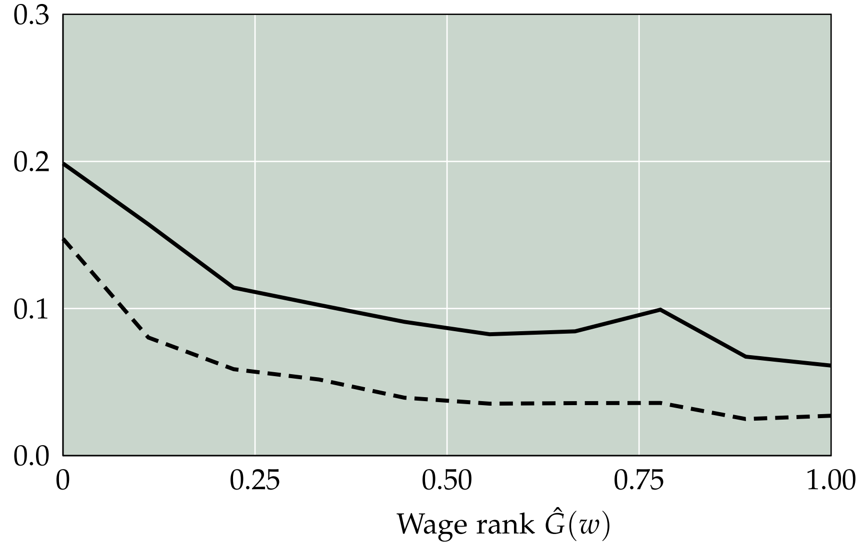
[22] By conditioning on $\hat{G}(w)$, i.e. the rank of the wage measure, we aviod that the location of the wage distribution affects the fit of the regression.

Figure 5.9: Auxiliary parametric quartely hazard functions (EE- , EU- and UE-transitions)



Note: In the two top panels the dashed, solid and dotted lines show hazard functions conditional on the productivity index $\iota(p)$ equal to the 10th percentile, the 50th percentile and the 90th percentile in the distribution of the productivity index across workers, respectively. The gray lines show Kaplan-Meier hazard function estimates.

Figure 5.10: Quarterly wage rank conditional EE- and EU-separation rates



Note: The solid and dashed lines show EE- and EU separation rates, respectively.

Table 4: Jobs per employment cycle and unemployment to employment ratio

Mean number of jobs per employment cycle (incl. left and right censored cycles)	2.20
S.d. number of jobs per employment cycle (incl. left and right censored cycles)	1.17
Mean number of jobs per employment cycle (excl. left and right censored cycles)	1.99
S.d. number of jobs per employment cycle (excl. left and right censored cycles)	1.14
Average fraction of jobs ending in unemployment	0.29

of all non-left-censored employment spells that are not initiated in the last calender year of the observation period. As a wage measure w we use firm level average wages. The resulting EE- and EU-transition profiles are rendered graphically in Figure 5.10. We defer further comments until assessing the model's ability to fit the observed transition pattern.

5.2.4 Employment cycles

The final set of moments related to labor market transitions are measures of the number of jobs per employment cycle and the average fraction of employment spells in annual cross sections that terminates with a transition to unemployment. An employment cycle is a cycle of job spells with no intervening unemployment spells. These moments are reported in Table 4 below.

5.3 Wages and productivity

The third set of auxiliary models that conditions our estimation relates to cross sectional heterogeneity in in wages, productivity and firm sizes.

5.3.1 Log wage regression

Given our model's focus on two-sided heterogeneity and sorting, the two-way (firm and worker) fixed effect model popularized by Abowd et al. (1999) seems to be a natural descriptive model for wages. However, as shown in section XXX, moments from this model that may at first sight appear central to our analysis such as the correlation between worker and firm fixed effects, may provide entirely misleading information on the strenght and sign of labor market sorting. Moreover, estimation of the the full Abowd et al. (1999) log wage regression is computational intensive and therefore not ideally suite for an Indirect Inference procedure. Still, a regression model in the spirit of Abowd et al. (1999) provides a parsimonious representation of wage heterogeneity in the model and 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

$$\ln w_{in} = \alpha_i + \xi_{\mathrm{J}(i,n)} + \varepsilon_{in}$$

where i index individuals, n index observatations on individual i and $\mathrm{J}(i,n) = j$ if worker i is employed by firm j in the n'th observations of worker i. Hence, the auxiliary log wage regression decomposes wages into a fixed worker specific component α_i, a fixed firm specific component ξ_j, and a residual component ε_{in}. When estimating the auxilry wage regression we impose the following restrictions: $\mathrm{E}\left[\varepsilon_{in}|i, \mathrm{J}(i,n)\right] = 0$ and $\mathrm{E}\left[(\alpha_i - \overline{\alpha})\left(\xi_{\mathrm{J}(i,n)} - \overline{\xi}\right)\right] = 0$. The former is usually referred to as an assumption of "exogenous mobility" while the latter is sometimes referred to as an assumption of "no sorting". The notion of "exogenous mobility" and "no sorting" in a purposefully misspecified auxiliary model embedded in an Indirect Inference procedure is not meaningful in an economic sense, but the restrictions does have some useful econometric implications. The first restrictions, also imposed in Abowd et al. (1999), allows for least squares estimation of the parameters of the auxiliary log wage regression (i.e. the worker and firm fixed effects). The

Table 5: Distribution of firm and worker effects from auxiliary log wage regression

	Mean	S. d.	Skewness	Kurtosis	$\frac{Var(x)}{Var(\ln w)}$
Log wage	5.21	.31	.45	3.54	1.00
Worker effects	5.21	.28	.58	3.73	.83
Firm effects	–	.05	−.17	25.07	.03
Residual	–	.12	−.13	7.36	.14

second restriction was not imposed in Abowd et al. (1999). Under $E\left[(\alpha_i - \overline{\alpha})\left(\xi_{J(i,n)} - \overline{\xi}\right)\right] = 0$ (and no covariates) we may treat worker and firm effects as *random effects* rather than *fixed effects*. In other words, we may ignore firm fixed effect when computing the worker fixed effects and *vice verca*.

We estimate the auxiliary log wage regresssion on an annual 10 year long panel of wages extracted from the master panel described in section 4 and include the first four moments of the distribution of worker fixed effects, and the second, third and fourth moment of the firm fixed effects and residual log wages in the set of auxiliary statistics.[23] Notice that this set of moments include a decomposition of the total log-wage variance as observed in the data. Table 5 reports the estimated moments and the implied log wage variance decomposition

5.3.2 Mean-min ratios

To further tie our model to the data and to address an ongoing discussion related to the ability of job search models to deliver the amount of wage dispersion that is observed in the data we include in our set of auxiliary statistics the Mean-Min ratio proposed as a useful and parsimonious measure of wage dispersion proposed by Hornstein, Krussel and Violante (2011) . Table 6 reports the empirical Mean-Min ratios for three different minimum wage measurements (1st, 5th and 10th percentile in the wage distribution).

5.3.3 Productivity, wages and firm size

The auxiliary firm fixed effects estimated from the auxiliary log wage regression captures firm specific effects in remuneration, such as firm productivity. Our data however also provides direct information on firm productivity, as measured by value added per worker hour. As described in

[23] The mean firm fixed effect and the mean residual log wage is of course normalized to zero.

Table 6: Mean-Min ratios

Mean wage	198.55
Minimum wage (1st percentile)	97.54
Minimum wage (5th percentile)	120.62
Minimum wage (10th percentile)	131.67
Mean-Min ratio (1st percentile)	2.04
Mean-Min ratio (5th percentile)	1.65
Mean-Min ratio (10th percentile)	1.51

Table 7: Moments of the joint distribution of productivity, wages and firm size

	Mean	S. d.	Skew.	Kurt.	(1)	(2)	(3)	(4)	(5)
(1) Value added/FTE	222.40	99.61	2.15	10.49	1.00	0.62	0.13	0.08	0.04
(2) Wage cost/FTE	166.73	49.90	2.26	17.27		1.00	0.09	0.10	0.04
(3) Value added (mill. DKK)	21.59	105.04	25.31	887.63			1.00	0.95	0.91
(4) Wage cost (mill. DKK)	15.75	66.48	22.02	710.80				1.00	0.98
(5) FTE	89.81	392.48	25.94	991.48					1.00

section4, this piece of information originates from a survey conducted by Statistics Denmark that is subject to a specific sampling scheme also described in section4. The structural model allow us to reproduce this sampling scheme in the simulation thus allowing precise replication of the actual data.

We include the first four moments of the distribution of value added per worker-hour, firm-level wage cost per worker-hour and firm size, measured by the annual work force size (full time equivalents) and total value added, as well as the correlation matrix of these four variables. The moments pertains to the distribution of the four variables across firms and is subject to the sampling rules set out in section4.

Table 7 reports the empirical moments relating to the joint distribution of productivity, wages and firm size.

6 Model Estimation

The model is estimated by indirect inference on a selection of the statistics described in the previous section. The auxiliary model is detailed in the tables that describe the fit of the model estimate. We make the following specifications. The worker skill distribution Ψ is assumed to be a beta distribution with parameters $\left(\beta_0^\psi, \beta_1^\psi\right)$. Conditional on the worker's skill level h, the probability that she is of the low lay-off rate type is given by, $\Pr\left(\delta_1 = \delta_l | h\right) = \frac{\exp(\xi_0 + \xi_1 h)}{1 + \exp(\xi_0 + \xi_1 h)}$. As is common in this type of search equilibrium estimation, we directly estimate the vacancy offer distribution $\Gamma\left(\cdot\right)$, and then subsequently back out the firm type distribution $\Phi\left(\cdot\right)$, that is consistent with the estimated offer distribution. For the case of $\kappa \leq 1$, it is trivial to show that there exist a unique firm type distribution for any Γ estimate on the $[0, 1]$ support. We specify the offer distribution to be a beta distribution with parameters $\left(\beta_0^{\gamma}, \beta_1^{\gamma}\right)$. We maintain the specification of $\kappa = 1$ in the estimation. The interest rate is set at $r = 0.05$, hence all rates are stated at annual frequency. The vacancy posting cost function is specified to be $c_{\nu}\left(\nu\right) = \frac{c_0^{\nu}}{1 + \frac{1}{c_1^{\nu}}} \nu^{1 + \frac{1}{c_1^{\nu}}}$. In the estimation of the search frictions, we again employ a short cut and normalize $\lambda\left(\theta\right) = 1$ as well as $c_0^{\nu} = 1$. λ and the constant terms in the search and recruitment cost functions are not separably identified. For the purpose of counterfactuals, one will then require a specification of the matching function consistent with $\lambda\left(\theta\right) = 1$ and the search cost estimate.

The indirect inference procedure is a simulated minimum distance estimation where the model is simulated to produce a data set of exactly the same structure as that of the real data. The model estimate is the particular model parameter combination that minimizes the distance between the simulated data and the real data where the distance is determined by the auxiliary model. 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 200,000 over 10 years, 360 times. The simulated auxiliary model statistics are the average over those 360 simulation repetitions.

The parameters to be estimated are, $\left(c_0, c_1, c_1^{\nu}, \delta_0, \delta_l, \delta_h, \xi_0, \xi_1, \beta_0^{\psi}, \beta_1^{\psi}, \beta_0^{\gamma}, \beta_1^{\gamma}, \beta, f_0, \rho\right)$. The estimate is presented in Table 8. First of all, it is seen that the estimate implies that the production is supermodular, hence more skilled workers are on average with more productive firms. This sorting pattern is reinforced by the estimated negative correlation between worker skill and worker specific layoff rate. More skilled workers are more likely to be of the low lay-off rate type. As a result, less skilled workers a laid off more often and do not on average climb as high on the ladder as their more skilled peers. Specifically, the correlation between worker skill and lay off

Table 8: Model estimate

	Estimate	Std. error
c_0	205.044	
c_1	6.069	
c_1^ν	1.478	
δ_0	0.064	
δ_l	0.006	
δ_h	0.431	
ξ_0	-5.881	
ξ_1	9.438	
β_0^ψ	0.940	
β_1^ψ	14.194	
β_0^γ	12.510	
β_1^γ	4.652	
β	0.656	
f_0	778.793	
ρ	-2.857	

rate is such that the probability of being a low lay off rate type is 0.19 for the lowest skilled 3 percent of the workers. The highest skilled 3 percent of workers are low lay off rate with probability 0.94. The low lay off rate is 0.006 and the high lay off rate is estimated at 0.431. The high lay off rate workers' incentives to search for job opportunities are greatly affected by their high employed effective discount rate which combined with their fast exit out of employment results in a high unemployment rate of 43% as opposed to an unemployment rate of less than 1% of the low lay off rate workers. The correlation between worker skill and firm productivity in the match distribution is 0.118. It is useful to quantify the estimated complementarities by performing a simple counterfactual of taking all the existing matches in the equilibrium and rearrange the workers and firms to maximize output. In this case, this means matching the highest skilled worker with the highest productivity job, the second highest skilled worker with the second most productive job, and so on.[24] We find very modest efficiency improvements from an optimal assignment of the low layoff worker matches: Output increases by 0.2%. This low estimate is probably tied to the fairly low estimate of heterogeneity in the model, and one would expect that

[24]This is of course not quite right in that the optimal assignment is a problem of matching a two-dimensional worker type with a single dimensional firm type. However, the employment rate of high layoff rate workers is sufficiently low, that it is a reasonable approximation to just consider the optimal assignment of the low layoff rate workers, which is what we do.

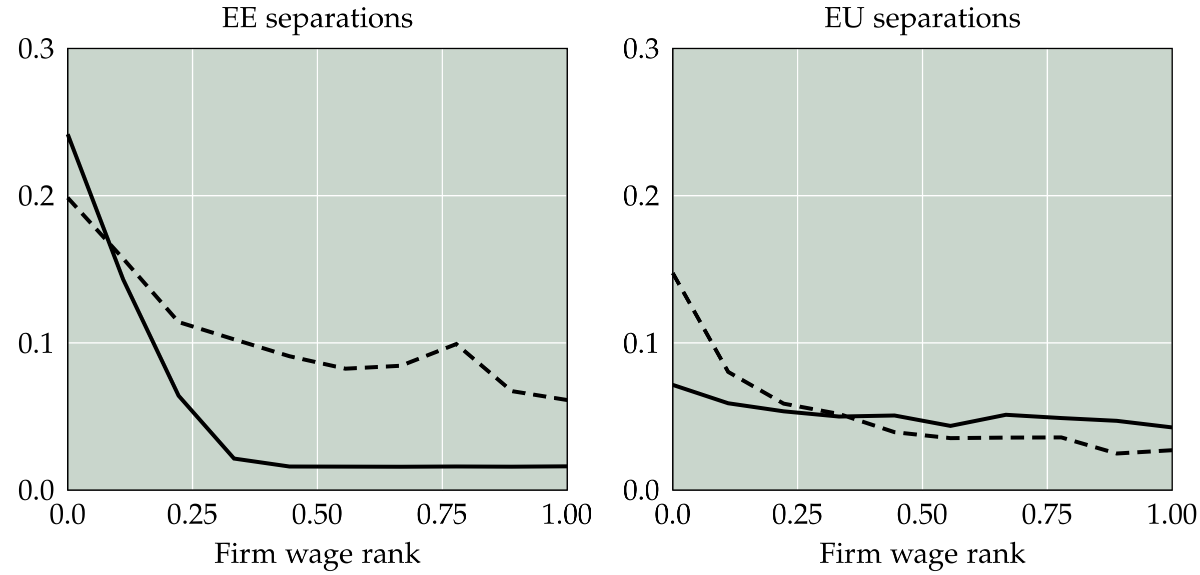
Table 9: Model fit.

	Data	Simulation
Quarterly UE hazard	0.162	0.154
Quarterly EU hazard	0.047	0.052
Quarterly EE hazard	0.101	0.046
Parametric duration models		
-UE, λ_0	-1.638	-1.306
-UE, σ^2	1.137	1.137
-EU, λ_0	-1.497	-1.232
-EU, λ_1	-2.819	-1.907
-EU, σ^2	1.848	1.848
-EE, λ_0	-1.985	-0.823
-EE, λ_1	-1.047	-0.921
-EE, σ^2	0.701	0.701
Avg(VA/FTE)	222.404	211.836
Std(VA/FTE)	99.611	36.897
Avg(Wages/FTE)	166.727	181.670
Std(Wages/FTE)	49.904	27.307
Avg(FTE)	89.810	86.979
Std(FTE)	392.481	52.229
Cov(VA/FTE,Wages/FTE)	3,080.860	994.776
Cov(VA/FTE,FTE)	1,612.566	1,881.282
Avg(U dur | top 5% firm)	56.864	69.950
Avg(log start wage | top 5% firm)	5.208	4.259
Cov(log start wage,U dur | top 5% firm)	-2.298	-2.302
Mean/(bottom 5th percentile) wage ratio	1.646	1.926
Avg JU fraction	0.287	0.397
Avg(# jobs in emp cycle)	2.204	1.420
Std(# jobs in emp cycle)	1.425	0.592

if the model did a better job of matching the variance in firm wages and labor productivity, then the counterfactual efficiency improvement would increase as well.

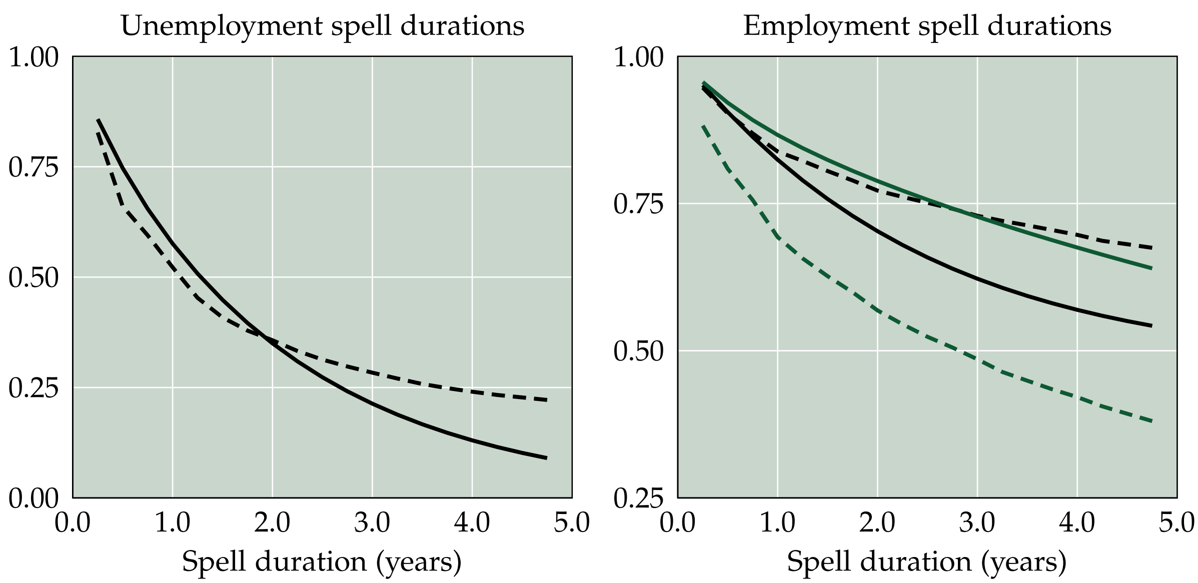
The model fit is described in Table 9 and Figures 6.1 and 6.2. Generally, the model fits the qualitative patterns in the data. Both the EE and EU hazard rates are declining in firm wages. The Kaplan-Meier duration conditional hazard rates are declining in duration as a result of the estimated search intensity heterogeneity across workers. Firm size and labor productivity are positively related, and so are wages and firm labor productivity. The overall level of job-to-job transitions is a bit too low. As a result, the model under estimates the number of jobs in an employment cycle and over estimates the JU fraction. The model hits the levels of firm

Figure 6.1: Model fit: Firm wage rank conditional separation rates



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

Figure 6.2: Model fit: Kaplan-Meier survivor functions.



Note: Data in dashed line. Model estimate in solid line. Employment spells ending in an EE transtion drawn in green. Employment spells ending in an EU transition drawn in black.

labor productivity, wages and labor force size fine, but cannot capture the full variation of those measures across firms in the data. The model, on the other hand, generates a bit too much wage dispersion across workers according to the mean-min wage ratio. The model does well in capturing the negative correlation between unemployment duration and wages among worker newly hired workers in top firms.

Detailed derivations Consider an employed worker of type h who is employed with a type p firm at employment contract (w,s). Denote by $q = q(h,w,p)$, the threshold type such that a meeting of an outside firm with type less than q has no impact on the worker's wage. Furthermore, adopt the short hand $V(h,q,p)$ as the value of employment to a type h worker who is employed with a type p firm subject to an employment contract set through bargaining where the worker had the threat point to accept outside employment with a type q firm. The value function, $\tilde{V}(h,w,p,s)$, for the employed worker is,

$$\begin{aligned}
r\tilde{V}(h,p,w,s) &= w - c(s) + [\delta_1 + \Gamma(R(h))\,\delta_0\lambda(\theta)]\,[V_0(h) - V(h,q,p)] + \\
&\quad s\lambda(\theta)\int_p^{\bar{p}} \left[V(h,p,p') - V(h,q,p)\right] d\Gamma(p') + \\
&\quad s\lambda(\theta)\int_q^{p} \left[V(h,p',p) - V(h,q,p)\right] d\Gamma(p') + \\
&\quad \delta_0\lambda(\theta)\int_{R(h)}^{\bar{p}} \left[V(h,R(h),p') - V(h,q,p)\right] d\Gamma(p') \\
&= w - c(s) + [\delta_1 + \Gamma(R(h))\,\delta_0\lambda(\theta)]\,V_0(h) - \left[\delta_0\lambda(\theta) + \delta_1 + s\lambda(\theta)\big(1-\Gamma(q)\big)\right] V(h,q,p) + \\
&\quad s\lambda(\theta)\int_p^{\bar{p}} \left[\beta V(h,p',p') + (1-\beta)V(h,p,p)\right] d\Gamma(p') + \\
&\quad s\lambda(\theta)\int_q^{p} \left[\beta V(h,p,p) + (1-\beta)V(h,p',p')\right] d\Gamma(p') + \\
&\quad \delta_0\lambda(\theta)\int_{R(h)}^{\bar{p}} \left[\beta V(h,p',p') + (1-\beta)V_0(h)\right] d\Gamma(p')
\end{aligned}$$

Integration by parts yields,

$$\begin{aligned}
(r + \delta_0\lambda(\theta) + \delta_1)\,\tilde{V}(h,p,w,s) &= w - c(s) + [\delta_1 + \Gamma(R(h))\,\delta_0\lambda(\theta)]\,V_0(h) - s\lambda(\theta)\big(1-\Gamma(q)\big)V(h,q,p) + \\
&\quad s\lambda(\theta)(1-\beta)\big(1-\Gamma(p)\big)V(h,p,p) + s\lambda(\theta)\beta\big(1-\Gamma(p)\big)V(h,p,p) + \\
&\quad s\lambda(\theta)\beta\int_p^{\bar{p}} \big(1-\Gamma(p')\big)V_p'(h,p',p')dp' + \\
&\quad s\lambda(\theta)\beta\big(\Gamma(p)-\Gamma(q)\big)V(h,p,p) - s\lambda(\theta)(1-\beta)\big(1-\Gamma(p)\big)V(h,p,p) + \\
&\quad s\lambda(\theta)(1-\beta)\big(1-\Gamma(q)\big)V(h,q,q) + s\lambda(\theta)(1-\beta)\int_q^{p} \big(1-\Gamma(p')\big)V'(h,p',p')dp' \\
&\quad \delta_0\lambda(\theta)(1-\beta)\,[1-\Gamma(R(h))]\,V_0(h) + \delta_0\lambda(\theta)\beta\,[1-\Gamma(R(h))]\,V_0(h) + \\
&\quad \delta_0\lambda(\theta)\beta\int_{R(h)}^{\bar{p}} \left[1-\Gamma(p')\right] V'(h,p',p')dp'.
\end{aligned}$$

By $V(h,q,p) = \beta V(h,p,p) + (1-\beta)V(h,q,q)$, one obtains.

$$\begin{aligned}
\left(r+\delta_0\lambda(\theta)+\delta_1\right)\tilde{V}(h,p,w,s) \quad = \quad & f(h,p) - c\,(s) + \left(\delta_0\lambda(\theta)+\delta_1\right)V_0(h) + \\
& s\lambda(\theta)\beta\int_p^{\bar{p}} V'\left(h,p',p'\right)\left[1-\Gamma(p')\right]dp' + \\
& s\lambda(\theta)(1-\beta)\int_q^p V'(h,p',p')[1-\Gamma(p')]dp' + \\
& \delta_0\lambda(\theta)\beta\int_{R(h)}^{\bar{p}} V'\left(h,p',p'\right)\left[1-\Gamma(p')\right]dp'. \qquad (.1)
\end{aligned}$$

By the envelope theorem it follows that,

$$\begin{aligned}
(r+\delta_0\lambda(\theta)+\delta_1)V_p'(h,p,p) \quad &= \quad f_p'(h,p) - s(h,p)\lambda(\theta)\beta\big(1-\Gamma(p)\big)V_p'(h,p,p) \\
&\Updownarrow \\
V_p'(h,p,p) \quad &= \quad \frac{f_p'(h,p)}{r+\delta_0\lambda(\theta)+\delta_1+\beta s(h,p)\lambda(\theta)\big(1-\Gamma(p)\big)}. \qquad (.2)
\end{aligned}$$

Hence, equation (D.1) can be written as,

$$\begin{aligned}
\left(r+\delta_0\lambda(\theta)+\delta_1\right)\tilde{V}(h,p,w,s) \quad = \quad & w - c\,(s) + \left(\delta_0\lambda(\theta)+\delta_1\right)V_0(h) + \\
& s\lambda(\theta)\beta\int_p^{\bar{p}} \frac{f_p'(h,p')[1-\Gamma(p')]dp'}{r+\delta_0\lambda(\theta)+\delta_1+\beta s(h,p')\lambda(\theta)[1-\Gamma(p')]} + \\
& s\lambda(\theta)(1-\beta)\int_q^p \frac{f_p'(h,p')[1-\Gamma(p')]dp'}{r+\delta_0\lambda(\theta)+\delta_1+\beta s(h,p')\lambda(\theta)[1-\Gamma(p')]} + \\
& \delta_0\lambda(\theta)\beta\int_{R(h)}^{\bar{p}} \frac{f_p'(h,p')[1-\Gamma(p')]dp'}{r+\delta_0\lambda(\theta)+\delta_1+\beta s(h,p')\lambda(\theta)[1-\Gamma(p')]}.
\end{aligned}$$

Steady state $G(h,q,p)$ The steady state condition on $G(h,q,p)$ is given by,

$$\begin{aligned}
(1-u)\delta G(h,q,p) + (1-u)\lambda(\theta)\int_{\underline{h}}^{h}\int_{R(h')}^{q}\Big\{\big(1-\Gamma(p)\big)\int_{q'}^{q} s\big(h',p'\big)dG\big(h',q',p'\big) \\
+\big(1-\Gamma(q)\big)\int_q^p s\big(h',p'\big)dG\big(h',q',p'\big)\Big\} \quad = \\
\int_{\underline{h}}^{h} I(R(h') \le q)\big[\Gamma(p)-\Gamma(R(h'))\big]\lambda(\theta)\Big[u[\delta_0+\kappa s_0(h')]v(h') + \\
(1-u)\delta_0\int_{R(h')}^{\bar{p}}\int_{q'}^{\bar{p}} g(h',q',p')dp'dq'\Big]dh'. \qquad (.3)
\end{aligned}$$

Evaluate at $(h, \bar{p}, \bar{p})$ and differentiate with respect to h to obtain,

$$
\begin{aligned}
(\delta_0\lambda(\theta)+\delta_1)(1-u)\int_{R(h)}^{\bar{p}}\int_{q'}^{\bar{p}} g(h,q',p')dp'dq' &= [1-\Gamma(R(h))]\lambda(\theta)\Big\{u[\mu+\kappa s_0(h)]v(h')+ \\
&\quad (1-u)\delta_0\int_{R(h)}^{\bar{p}}\int_{q'}^{\bar{p}} g(h,q',p')dp'dq'\Big\} \\
&\Updownarrow \\
(\delta_0\lambda(\theta)\Gamma(R(h))+\delta_1)(1-u)\int_{R(h)}^{\bar{p}}\int_{q'}^{\bar{p}} g(h,q',p')dp'dq' &= u[1-\Gamma(R(h))]\lambda(\theta)[\mu+\kappa s_0(h)]v(h) \\
&\Updownarrow \\
\delta_0(1-u)\int_{R(h)}^{\bar{p}}\int_{q'}^{\bar{p}} g(h,q',p')dp'dq' &= \frac{\delta_0\lambda(\theta)[1-\Gamma(R(h))]}{\delta_0\lambda(\theta)\Gamma(R(h))+\delta_1}u[\mu+\kappa s_0(h)]v(h) \qquad (.4)
\end{aligned}
$$

Insert this into equation (.3),

$$
\begin{aligned}
\frac{\delta_0\lambda(\theta)+\delta_1}{\lambda(\theta)}G(h,q,p)+\int_{\underline{h}}^{h}\int_{R(h')}^{q}\Big[[1-\Gamma(p)]\int_{q'}^{q} s(h',p')dG(h',q',p') \\
+[1-\Gamma(q)]\int_{q}^{p} s(h',p')dG(h',q',p')\Big] &= \\
\frac{u}{1-u}\int_{\underline{h}}^{h} I(R(h')\le p)[\Gamma(p)-\Gamma(R(h'))][\mu+\kappa s_0(h')]v(h')\frac{\delta_1+\delta_0\lambda(\theta)}{\delta_0\lambda(\theta)\Gamma(R(h))+\delta_1}dh'. \qquad (.5)
\end{aligned}
$$

Evaluate (.5) at $(\bar{h}, \bar{p}, \bar{p})$ to obtain,

$$
\begin{aligned}
\frac{\delta_0\lambda(\theta)+\delta_1}{\lambda(\theta)} &= \frac{u}{1-u}\int_{\underline{h}}^{\bar{h}}[1-\Gamma(R(h'))]\frac{\delta_1+\delta_0\lambda(\theta)}{\delta_0\lambda(\theta)\Gamma(R(h))+\delta_1}[\mu+\kappa s_0(h')]v(h')dh' \\
&\Updownarrow \\
\frac{u}{1-u} &= \left[\int_{\underline{h}}^{\bar{h}}\frac{[1-\Gamma(R(h'))]\,[\mu+\kappa s_0(h')]}{\delta_0\Gamma(R(h'))+\delta_1/\lambda(\theta)}v(h')dh'\right]^{-1} \\
&\Updownarrow \\
u &= \left[\int_{\underline{h}}^{\bar{h}}\left(1+\frac{[1-\Gamma(R(h'))]\,[\mu+\kappa s_0(h')]}{\delta_0\Gamma(R(h'))+\delta_1/\lambda(\theta)}\right)d\Upsilon(h')\right]^{-1}.
\end{aligned}
$$

One then obtains,

$$
\begin{aligned}
\int_{\underline{h}}^{h}\int_{R(h')}^{q}\Big[\int_{q'}^{q}[\delta/\lambda(\theta)+[1-\Gamma(p)]s(h',p')]g(h',q',p')\,dp' \\
+\int_{q}^{p}[\delta/\lambda(\theta)+[1-\Gamma(q)]s(h',p')]g(h',q',p')dp'\Big]dq'dh' &= \\
\frac{\delta}{\lambda(\theta)}\frac{\int_{\underline{h}}^{h} I(R(h')\le q)[\Gamma(p)-\Gamma(R(h'))]\frac{\mu+\kappa s_0(h')}{\delta_0\Gamma(R(h'))+\delta_1/\lambda(\theta)}d\Upsilon(h')}{\int_{\underline{h}}^{\bar{h}}\frac{[1-\Gamma(R(h'))][\mu+\kappa s_0(h')]}{\delta_0\Gamma(R(h'))+\delta_1/\lambda(\theta)}d\Upsilon(h')}. \qquad (.6)
\end{aligned}
$$

Steady state equilibrium solution for $\Upsilon(h)$ Consider the equilibrium condition,

$$\Psi(h) = u\Upsilon(h) + (1-u)\, G(h, \bar{p}) .$$

Differentiate with respect to h to obtain,

$$\begin{aligned}\psi(h) &= uv(h) + (1-u)\int_b^{\bar{p}} g(h, p')\, dp' \\ &= \left[1 + \frac{[1-\Gamma(R(h))][\mu + \kappa s_0(h)]}{\delta_0\Gamma(R(h)) + \delta_1/\lambda(\theta)}\right] uv(h),\end{aligned}$$

where the last equality follows from equation (.4). By the steady state unemployment rate expression in equation (**??**), it follows that,

$$\psi(h) = \frac{\left[1 + \frac{[1-\Gamma(R(h))][\mu+\kappa s_0(h)]}{\delta_0\Gamma(R(h))+\delta_1/\lambda(\theta)}\right] v(h)}{\int_{\underline{h}}^{\bar{h}} \left(1 + \frac{[1-\Gamma(R(h'))][\mu+\kappa s_0(h')]}{\delta_0\Gamma(R(h'))+\delta_1/\lambda(\theta)}\right) v(h')\, dh'}, \tag{.7}$$

which is an integral equation for $\Upsilon(h)$ as a function of $\Psi(h)$. Define,

$$\Delta(h) = \frac{[1-\Gamma(R(h))][\mu + \kappa s_0(h)]}{\delta_0\Gamma(R(h)) + \delta_1/\lambda(\theta)}.$$

Then restate equation (.7),

$$v(h) = \left[1 + \int_{\underline{h}}^{\bar{h}} \Delta(h')\, v(h')\, dh'\right] \frac{\psi(h)}{1 + \Delta(h)}.$$

Use equation (.7) to solve for $1 + \int_{\underline{h}}^{\bar{h}} \Delta(h')\, v(h')\, dh'$. First, some minor manipulation,

$$\begin{aligned}\psi(h) + \psi(h)\int_{\underline{h}}^{\bar{h}} \Delta(h')\, v(h')\, dh' &= [1 + \Delta(h)]\, v(h) \\ &\Updownarrow \\ v(h) - \frac{\psi(h)}{1+\Delta(h)}\int_{\underline{h}}^{\bar{h}} \Delta(h')\, v(h')\, dh' &= \frac{\psi(h)}{1+\Delta(h)} \\ &\Updownarrow \\ \Delta(h)\, v(h) - \frac{\psi(h)\,\Delta(h)}{1+\Delta(h)}\int_{\underline{h}}^{\bar{h}} \Delta(h')\, v(h')\, dh' &= \frac{\psi(h)\,\Delta(h)}{1+\Delta(h)}.\end{aligned}$$

Now, integrate from $\underline{h}$ to $\bar{h}$ to obtain,

$$\begin{aligned}
\int_{\underline{h}}^{\bar{h}} \Delta\left(h'\right) v\left(h'\right) dh' \left[1 - \int_{\underline{h}}^{\bar{h}} \frac{\psi\left(h'\right)\Delta\left(h'\right)}{1+\Delta\left(h'\right)} dh'\right] &= \int_{\underline{h}}^{\bar{h}} \frac{\psi\left(h'\right)\Delta\left(h'\right)}{1+\Delta\left(h'\right)} dh' \\
&\Updownarrow \\
1 + \int_{\underline{h}}^{\bar{h}} \Delta\left(h'\right) v\left(h'\right) dh' &= 1 + \frac{\int_{\underline{h}}^{\bar{h}} \frac{\psi(h')\Delta(h')}{1+\Delta(h')} dh'}{1 - \int_{\underline{h}}^{\bar{h}} \frac{\psi(h')\Delta(h')}{1+\Delta(h')} dh'} \\
&= \frac{1}{1 - \int_{\underline{h}}^{\bar{h}} \frac{\Delta(h')}{1+\Delta(h')} \psi\left(h'\right) dh'} \\
&= \frac{1}{\int_{\underline{h}}^{\bar{h}} \left[1 - \frac{\Delta(h')}{1+\Delta(h')}\right] \psi\left(h'\right) dh'} \\
&= \frac{1}{\int_{\underline{h}}^{\bar{h}} \frac{1}{1+\Delta(h')} \psi\left(h'\right) dh'}.
\end{aligned}$$

Hence, one obtains the solution,

$$v\left(h\right) = \frac{\left[1+\Delta\left(h\right)\right]^{-1} \psi\left(h\right)}{\int_{\underline{h}}^{\bar{h}} \left[1+\Delta\left(h'\right)\right]^{-1} \psi\left(h'\right) dh'},$$

which can also be written as,

$$\Upsilon\left(h\right) = \frac{\int_{\underline{h}}^{h} \frac{\delta_0 \Gamma(R(h')) + \delta_1/\lambda(\theta)}{\delta_0 \Gamma(R(h')) + \delta_1/\lambda(\theta) + [1-\Gamma(R(h'))][\mu + \kappa s_0(h')]} d\Psi\left(h'\right)}{\int_{\underline{h}}^{\bar{h}} \frac{\delta_0 \Gamma(R(h')) + \delta_1/\lambda(\theta)}{\delta_0 \Gamma(R(h')) + \delta_1/\lambda(\theta) + [1-\Gamma(R(h'))][\mu + \kappa s_0(h')]} d\Psi\left(h'\right)}.$$

A Firm labor force composition is independent of firm size

Consider a labor force that consists of k types. For the purpose of this argument, a type i worker is characterized by a hire rate h_i and a separation rate d_i. Firm entry and exit takes place through the zero labor force size pool. Each worker i size process is independent. Hence, the distribution of the number of type i workers employed by the firm will be Poisson distributed,

$$m_n^i = \frac{\left(\frac{h_i}{d_i}\right)^n \exp\left(-\frac{h_i}{d_i}\right)}{n!}.$$

Denote by $\vec{n} = (n_1, n_2, \ldots, n_k)$ the composition of the firm's labor force. The mass of size n firms is formed based on the sum of the individual worker type distributions,

$$\begin{aligned}
m_n &= \sum_{\{\vec{n} \geq 0 | \sum n_i = n\}} \prod_{i=1}^{k} m_{n_i}^i \\
&= \frac{\left[\sum_{i=1}^{k} \frac{h_i}{d_i}\right]^n \exp\left(-\sum_{i=1}^{k} \frac{h_i}{d_i}\right)}{n!},
\end{aligned}$$

which is just a Poisson in the sum of the individual hiring and separation rate fraction. Consider the expectation of the share of type i workers in the firm's labor force conditional on the firm having n workers,

$$\begin{aligned}
E\left[\frac{n_i}{n}|n\right] &= \frac{\sum_{\left\{\vec{n}\geq 0|\sum n_j = n\right\}} \frac{n_i}{n} \prod_{j=1}^{k} m_{n_j}^{j}}{m_n} \\
&= \frac{\sum_{\left\{\vec{n}\geq 0|\sum n_j = n\right\}} n! \frac{n_i}{n} \frac{\prod_{j=1}^{k}\left(\frac{\eta_j}{\delta_j}\right)^{n_j}}{\prod_{j=1}^{k} n_j!}}{\sum_{\left\{\vec{n}\geq 0|\sum n_j = n\right\}} n! \frac{\prod_{j=1}^{k}\left(\frac{\eta_j}{\delta_j}\right)^{n_j}}{\prod_{j=1}^{k} n_j!}} \\
&= \frac{\left(\frac{\eta_i}{\delta_i}\right)\left[\sum_{i=1}^{k} \frac{h_i}{d_i}\right]^{n-1}}{\left[\sum_{i=1}^{k} \frac{h_i}{d_i}\right]^{n}} \\
&= \frac{\frac{h_i}{d_i}}{\sum_{i=1}^{k} \frac{h_i}{d_i}}
\end{aligned}$$

where the second to last step applies the multinomial theorem. Hence, the share of type i workers in the firm's labor force is independent of the size of the firm's labor force. Consequently, the firm's overall worker separation rate is not size dependent.

B Equilibrium market tightness and vacancies

We have the following relationships.

$$\lambda(\theta) = \theta\eta(\theta), \tag{B.1}$$

where

$$\theta = \frac{m \int_{b}^{\bar{p}} v\left(p'\right) d\Phi\left(p'\right)}{\sum_{j\in l,h} \Delta_j \left[u_j \int_{\underline{h}}^{\bar{h}} [\mu + \kappa s_j^0\left(h\right)] d\Upsilon_j\left(h\right) + \left(1-u_j\right) \int_{\underline{h}}^{\bar{h}} \int_{b}^{\bar{p}} [\delta_0 + s_j\left(h,p\right)] dG_j\left(h,p\right)\right]}.$$

$$c_v'\left(v\left(p\right)\right) = \eta\left(1-\beta\right) \sum_{j\in l,h} \int_{\underline{h}}^{\bar{h}} \int_{R_j(h')}^{p} \left[V_j\left(h',p,p\right) - V_j\left(h',p',p'\right)\right] d\Lambda_j\left(h',p'\right).$$

The latter can be written as,

$$v\left(p\right) = c_v'^{-1}\left(\frac{\lambda}{\theta}\left(1-\beta\right) \sum_{j\in l,h} \int_{\underline{h}}^{\bar{h}} \int_{R_j(h')}^{p} \left[V_j\left(h',p,p\right) - V_j\left(h',p',p'\right)\right] d\Lambda_j\left(h',p'\right)\right).$$

The hire cost function is specified is a power function, which implies

$$c_v'^{-1}\left(x\right) = \left(\frac{x}{c_{0v}}\right)^{c_{1v}}.$$

It then follows that,

$$\theta^{1+c_{1v}} = \frac{m \int_b^{\bar{p}} c_v'^{-1} \left(\lambda \Theta\left(p\right)\right) d\Phi\left(p'\right)}{\sum_{j \in l,h} \Delta_j \left[u_j \int_{\underline{h}}^{\bar{h}} [\mu + \kappa s_j^0\left(h\right)] d\Upsilon_j\left(h\right) + \left(1 - u_j\right) \int_{\underline{h}}^{\bar{h}} \int_b^{\bar{p}} [\delta_0 + s_j\left(h, p\right)] dG_j\left(h, p\right) \right]},$$

where,

$$\Theta\left(p\right) = \left(1 - \beta\right) \sum_{j \in l,h} \int_{\underline{h}}^{\bar{h}} \int_{R_j(h')}^{p} \left[V_j\left(h', p, p\right) - V_j\left(h', p', p'\right)\right] d\Lambda_j\left(h', p'\right).$$

It then remains only to spell out the calculation of $\Theta\left(p\right)$.

$$\Theta\left(p\right) = \left(1 - \beta\right) \sum_{j \in l,h} \Delta_j \frac{\int_{\underline{h}}^{\bar{h}} \left[V_j\left(h, p, p\right) - V_j\left(h, R_j(h), R_j(h)\right)\right] \left[u_j[\mu + \kappa s_j^0(h)] v_j(h) + \left(1 - u_j\right) \delta_0 \int_b^{\bar{p}} g_j\left(h, p'\right) dp' \right] + \int}{\sum_{j' \in l,h} \Delta_{j'} \int_{\underline{h}}^{\bar{h}} \left\{ u_{j'}[\mu + \kappa s_{j'}^0(h')] v_{j'}(h') + \left(1 - u_{j'}\right) \int_b^{\bar{p}} [\delta_0 + \right.}$$

C Other stuff

$$c'\left(s\right) = \left(1 + \frac{1}{c_1}\right) c\left(s\right)/s = c_0 s^{1/c_1} = \kappa \lambda\left(\theta\right) \int_{R_j(h)}^{\bar{p}} \frac{\beta f_p'(h, p')[1 - \Gamma(p')] dp'}{r + \delta_j + \beta s_j(h, p') \lambda(\theta)[1 - \Gamma(p')]}$$

$$rV_j^0\left(h\right) = f(h, b) - sc'\left(s\right) c_1 / \left(1 + c_1\right) + \left(\mu + \kappa s\right) c'\left(s\right)/\kappa = f(h, b) + \kappa^{1+c_1} s\left(R_j\left(h\right)\right) c'\left(s\left(R_j\left(h\right)\right)\right) / \left(1 + c_1\right) + \mu c'$$

D Employment contract bargaining

At the beginning of an employment relationship, the firm and the worker bargain over a constant wage and worker's search intensity that will remain in effect until the relationship terminates or both parties consent to renegotiation. The bargaining game is an application of the alternating offers game of Rubinstein (1982) and most resembles the exogenous break down version as presented in Binmore et al. (1986). The following two subsections present the subgame perfect equilibrium for the case of an unemployed worker worker and a worker who is renegotiating subsequent to an outside offer, respectively. The arguments are closely related to the bargaining games described in Cahuc et al. (2006), although the bargaining is simplified to take place in artificial time with zero disagreement values and the possibility of meeting another employer during bargaining is eliminated.

The outcomes of the alternating offers games are identical to that of axiomatic Nash bargaining where the threat point of the firm is always zero for the firm, and the worker's threat point is either unemployment or full surplus extraction from the least productive of the two firms competing over the worker. This is the argument presented in Dey and Flinn (2005). Specifically, the bargaining outcome of an unemployed worker maximizes the Nash product,

$$\{w_0(h,p), s(h,p)\} = \arg\max_{w,s} \left(\tilde{V}(h,p,w,s) - V_0(h)\right)^{\beta} \tilde{J}(h,w,p,s)^{(1-\beta)}, \tag{D.1}$$

which yields the worker valuation,

$$V(h, R(h), p) = \beta V(h,p,p) + (1-\beta)V_0(h). \tag{D.2}$$

The inclusion of the reservation productivity argument implicitly states that the worker will only accept to bargain with employer types greater than $R(h)$.

The outcome of a worker bargaining with two employer types, q and p such that $p > q$ is that the worker will negotiate an employment contract with the type p firm with a threat point of full surplus extraction and efficient search intensity with the lower type firm, $V(h,q,q)$. Hence, the employment contract that results from this bargaining setting is,

$$\{w(h,q,p), s(h,p)\} = \arg\max_{w,s} \left(\tilde{V}(h,p,w,s) - V(h,q,q)\right)^{\beta} \tilde{J}(h,w,p,s)^{(1-\beta)}. \tag{D.3}$$

The bargaining outcome is,

$$V(h,q,p) = \beta V(h,p,p) + (1-\beta)V(h,q,q). \tag{D.4}$$

In both cases (D.1) and (D.3), the agreed upon search intensity $s(h,p)$ is the one that maximizes total match surplus. This is the jointly efficient search intensity level and does not depend on the specific surplus split dictated by bargaining power and threat points.

D.1 Unemployed worker

Consider an alternating offers game where the worker makes an offer (w_e, s_e) to the firm. If the firm accepts, employment starts and the worker receives payoff $\tilde{V}(h,p,w_e,s_e)$ and the firm receives $\tilde{J}(h,p,w_e,s_e) = \tilde{V}(h,p,f(h,p),s_e) - \tilde{V}(h,p,w_e,s_e)$. If the firm rejects the offer, the bargaining breaks down with exogenous probability Δ. If so, the firm receives a zero payoff and the worker goes back to unemployment and receives $V_0(h)$. If bargaining does not break down, the bargaining moves to the next round where the firm makes an offer (w_f, s_f) with probability $1-\beta$ and the worker gets to make the offer (w_e, s_e) with probability β. If the firm makes the offer and the worker accepts, the worker receives $\tilde{V}(h,p,w_f,s_f)$ and the firm receives $\tilde{J}(h,p,w_f,s_f) = \tilde{V}(h,p,f(h,p),s_f) - \tilde{V}(h,p,w_f,s_f)$. If the worker rejects, the game moves on to the next round if no break down occurs. And again, the worker will make the offer with probability β and the firm with probability $1-\beta$. The game continues like this ad infinitum or until agreement is reached. Disagreement payoffs are zero and the discount rate between rounds is zero.

Both the worker and the firm will offer the same search intensity, $s_e = s_f = s(h,p)$, where $s(h,p) = \arg\max_s \tilde{V}(h,p,f(h,p),s)$. Furthermore, consider the strategies where the worker accepts any offer (w,s) such that $\tilde{V}(h,p,w,s) \geq \tilde{V}\big(h,p,w_f,s(h,p)\big)$ and rejects any offer such that $\tilde{V}(h,p,w,s) < \tilde{V}\big(h,p,w_f,s(h,p)\big)$. Similarly, the firm accepts any offer (w,s) such that $\tilde{J}(h,p,w,s) \geq \tilde{J}\big(h,p,w_e,s(h,p)\big)$ and rejects any offer such that $\tilde{J}(h,p,w,s) < \tilde{J}\big(h,p,w_e,s(h,p)\big)$.

By definition the firm's payoff satisfies $\tilde{J}(h,p,w,s) = \tilde{V}(h,p,f(h,p),s) - \tilde{V}(h,p,w,s)$. Hence, a firm accepts any offer such that

$$\tilde{V}(h,p,w,s) \leq \tilde{V}\big(h,p,w_e,s(h,p)\big) - \tilde{V}\big(h,p,f(h,p),s(h,p)\big) + \tilde{V}\big(h,p,f(h,p),s\big). \tag{D.5}$$

It is seen that the right hand side of the firm acceptance condition (D.5) is maximized for $s = s(h,p)$ and does not depend on w. Hence, any worker deviation $s_e' \neq s_e = s(h,p)$ that will be accepted by the firm must result in a worker payoff $\tilde{V}(h,p,w,s_e') < \tilde{V}(h,p,w_e,s(h,p))$, for any w, which is not profitable.

A similar argument can be made that the firm will not want to deviate from $s_f = s(h,p)$. The worker will accept any offer such that,

$$\tilde{J}(h,p,w,s) \leq \tilde{V}(h,p,f(h,p),s) - \tilde{V}\big(h,p,w_f,s(h,p)\big). \tag{D.6}$$

It is seen that the right hand side of the worker acceptance decision (D.6) is maximized for $s = s(h,p)$ and that it does not depend on w. Hence, any firm deviation $s_f' \neq s_f = s(h,p)$ that will be accepted by the worker must result in a firm payoff $\tilde{J}(h,p,w,s_f') < \tilde{J}(h,p,w_f,s_f)$, for any w, which is not profitable.

It also follows directly from the above acceptance arguments that any strategy that prescribes $s_e \neq s(h,p)$ or $s_f \neq s(h,p)$ cannot be an equilibrium because a deviation to $s(h,p)$ will be profitable.

Now consider potential deviations in the wage. The worker's payoff $\tilde{V}(h,p,w,s_e)$ is monotonically increasing in w. It follows directly from (D.5) that any worker wage offer deviation w'_e that will be accepted by the firm is such that $w'_e \leq w_e$. This is not profitable. Any other deviation will not be accepted by the firm and is therefore also not profitable. A similar argument applies to possible firm wage offer deviations.

Sub game perfection of the acceptance strategies requires that the worker is indifferent between accepting the firm's offer (w_f, s_f) and rejecting it. A similar indifference applies on the firm side. This disciplines the acceptance levels by,

$$\hat{V}(w_f) = (1-\Delta)\big[\beta\hat{V}(w_e) + (1-\beta)\hat{V}(w_f)\big] + \Delta V_0(h) \tag{D.7}$$

$$\hat{J}(w_e) = (1-\Delta)\big[\beta\hat{J}(w_e) + (1-\beta)\hat{J}(w_f)\big] \tag{D.8}$$

where $\hat{V}(w) = \tilde{V}(h,p,w,s(h,p))$ and $\hat{J}(w) = \tilde{V}(h,p,w,s(h,p))$. Equations (D.7) and (D.8) can be rewritten as,

$$\beta\big[\hat{V}(w_f) - \hat{V}(w_e)\big] = \Delta\big[V_0(h) - \beta\hat{V}(w_e) - (1-\beta)\hat{V}(w_f)\big] \tag{D.9}$$

$$(1-\beta)\big[\hat{J}(w_f) - \hat{J}(w_e)\big] = \Delta\big[\beta\hat{J}(w_e) + (1-\beta)\hat{J}(w_f)\big]. \tag{D.10}$$

Taking the limit as $\Delta \to 0$, equations (D.7) and (D.8) imply that $w_f \to w_e$. Denote the common limit by w. Hence,

$$\frac{\partial\hat{V}(w)}{\partial w} = \lim_{\Delta\to 0}\frac{\hat{V}(w_f) - \hat{V}(w_e)}{w_f - w_e}$$

$$\frac{\partial\hat{J}(w)}{\partial w} = \lim_{\Delta\to 0}\frac{\hat{J}(w_f) - \hat{J}(w_e)}{w_f - w_e}.$$

Since changes in w only affect the match surplus split, it follows that $\partial\hat{V}(w)/\partial w = -\partial\hat{J}(w)/\partial w$. Hence, taking the limit $\Delta \to 0$ in equations (D.9) and (D.10) yields,

$$-\frac{\beta}{1-\beta} = \frac{V_0(h) - \beta\hat{V}(w) - (1-\beta)\hat{V}(w)}{\beta\hat{J}(w) + (1-\beta)\hat{J}(w)}$$

$$\Updownarrow$$

$$\hat{V}(w) = \beta\hat{V}\big(f(h,p)\big) + (1-\beta)V_0(h). \tag{D.11}$$

Hence, as the break down probability goes to zero, the outcome of the alternating offers game limits to the outcome of the axiomatic Nash bargaining outcome in equation (D.2).

D.2 Employed worker

Cahuc et al. (2006) provide a strategic bargaining foundation for the axiomatic Nash bargaining outcome in equation (D.4). The outcome is a subgame perfect equilibrium in a game based on firms submitting bids for the worker subject to a worker's option to use the bids as threat points in a subsequent strategic bargaining game. In the game between two employers of types q and p, respectively, where $q \leq p$, the higher type firm wins by submitting a contract bid $(w, s(h, p))$ as stated in equation (D.4).

References

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

Binmore, Ken, Ariel Rubinstein, and Asher Wolinsky (1986). The nash bargaining solution in economic modelling. *The RAND Journal of Economics* 17, no. 2: 176–188.

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.

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

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 (2008). Identifying sorting - in theory. *University of Pennsylvania Working Paper*.

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. *Forthcoming in Journal of Economic Theory*.

Lise, Jeremy, Costas Meghir, and Jean-Marc Robin (2008). Matching, sorting, and wages. *University College London 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.

Yamaguchi, Shintaro (2006). Job search, bargaining, and wage dynamics. *Working Paper*.