

# Online Appendix

## An Empirical Model of Wage Dispersion with Sorting

Jesper Bagger\*      Rasmus Lentz<sup>†</sup>

March 28, 2014

Section A presents details on the structural model including the wage bargaining procedure we employ, section B contains the results of two robustness exercises. In the first, we overfit  $\text{corr}(\bar{t}_i^u, w_i^0 | \hat{\tau} \geq 0.95)$ , the correlation between unemployment durations and starting wages for workers hired into top ranked firms. In the second, we split our analysis data by education (high education, meaning some college education, and low education, meaning everyone else).

### A Model

#### A.1 The worker's valuation of a job

Consider an employed worker of type  $(j, h)$  who is employed with a type  $p$  firm at employment contract  $(w, s)$ . Denote by  $q = q_j(h, w, p)$ , 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  $V_j(h, q, p)$  as the value of employment to a type  $(j, h)$  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. Furthermore, let  $V_j(h, p) = V_j(h, p, p)$  and  $\hat{\Gamma}(p) = 1 - \Gamma(p)$  and  $\hat{\delta}_j = \delta_j + \delta_0 \lambda$ . The value function,  $\tilde{V}_j(h, w, p, s)$ , for the employed worker is,

$$\begin{aligned}
 r\tilde{V}_j(h, p, w, s) &= w - c(s) + [\delta_j + \Gamma(R_j(h)) \delta_0 \lambda] [V_j^0(h) - V_j(h, q, p)] + \\
 &\quad s\lambda \int_p^{\bar{p}} [V_j(h, p, p') - V_j(h, q, p)] d\Gamma(p') + \\
 &\quad s\lambda \int_q^p [V_j(h, p', p) - V_j(h, q, p)] d\Gamma(p') + \\
 &\quad \delta_0 \lambda \int_{R_j(h)}^{\bar{p}} [V(h, R_j(h), p') - V_j(h, q, p)] d\Gamma(p').
 \end{aligned}$$

---

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

<sup>†</sup>University of Wisconsin-Madison; NBER; LMDG; CAP; E-mail: rlentz@ssc.wisc.edu

This can be rewritten as,

$$\begin{aligned}
r\tilde{V}_j(h, p, w, s) &= w - c(s) + [\delta_j + \Gamma(R_j(h)) \delta_0 \lambda] V_j^0(h) - [\hat{\delta}_j + s\lambda \hat{\Gamma}(q)] V_j(h, q, p) + \\
&\quad s\lambda \int_p^{\bar{p}} [\beta V(h, p') + (1 - \beta)V(h, p)] d\Gamma(p') + \\
&\quad s\lambda \int_q^p [\beta V(h, p) + (1 - \beta)V(h, p')] d\Gamma(p') + \\
&\quad \delta_0 \lambda \int_{R_j(h)}^{\bar{p}} [\beta V(h, p') + (1 - \beta)V_j^0(h)] d\Gamma(p').
\end{aligned}$$

Integration by parts yields,

$$\begin{aligned}
(r + \hat{\delta}_j) \tilde{V}_j(h, p, w, s) &= w - c(s) + [\delta_j + \Gamma(R_j(h)) \delta_0 \lambda] V_j^0(h) - s\lambda \hat{\Gamma}(q) V(h, q, p) + \\
&\quad s\lambda(1 - \beta) \hat{\Gamma}(p) V(h, p) + s\lambda \beta \hat{\Gamma}(p) V(h, p) + s\lambda \beta \int_p^{\bar{p}} \hat{\Gamma}(p') V'_p(h, p', p') dp' + \\
&\quad s\lambda \beta (\Gamma(p) - \Gamma(q)) V(h, p) - s\lambda(1 - \beta) \hat{\Gamma}(p) V(h, p) + \\
&\quad s\lambda(1 - \beta) \hat{\Gamma}(q) V(h, q) + s\lambda(1 - \beta) \int_q^p \hat{\Gamma}(p') V'(h, p') dp' + \\
&\quad \delta_0 \lambda(1 - \beta) \hat{\Gamma}(R_j(h)) V_j^0(h) + \delta_0 \lambda \beta \hat{\Gamma}(R_j(h)) V_j^0(h) + \\
&\quad \delta_0 \lambda \beta \int_{R_j(h)}^{\bar{p}} \hat{\Gamma}(p') V'(h, p') dp'.
\end{aligned}$$

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

$$\begin{aligned}
(r + \hat{\delta}_j) \tilde{V}_j(h, p, w, s) &= f(h, p) - c(s) + \hat{\delta}_j V_0(h) + \\
&\quad s\lambda \beta \int_p^{\bar{p}} V'_j(h, p') \hat{\Gamma}(p') dp' + \\
&\quad s\lambda(1 - \beta) \int_q^p V'_j(h, p') \hat{\Gamma}(p') dp' + \\
&\quad \delta_0 \lambda \beta \int_{R_j(h)}^{\bar{p}} V'_j(h, p') \hat{\Gamma}(p') dp'. \tag{A.1}
\end{aligned}$$

By the envelope theorem it follows that,

$$\begin{aligned}
(r + \hat{\delta}_j) V'_j(h, p) &= f'_p(h, p) - s(h, p) \lambda \beta \hat{\Gamma}(p) V'_j(h, p) \\
&\quad \Updownarrow \\
V'_j(h, p) &= \frac{f'_p(h, p)}{r + \hat{\delta}_j + \beta s(h, p) \lambda \hat{\Gamma}(p)}. \tag{A.2}
\end{aligned}$$

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

$$\begin{aligned}
(r + \hat{\delta}_j) \tilde{V}_j(h, p, w, s) &= w - c(s) + \hat{\delta}_j V_j^0(h) + \\
& s\lambda\beta \int_p^{\bar{p}} \frac{f'_p(h, p') \hat{\Gamma}(p') dp'}{r + \hat{\delta}_j + \beta s(h, p') \lambda \hat{\Gamma}(p')} + \\
& s\lambda(1 - \beta) \int_q^p \frac{f'_p(h, p') \hat{\Gamma}(p') dp'}{r + \hat{\delta}_j + \beta s(h, p') \lambda \hat{\Gamma}(p')} + \\
& \delta_0 \lambda \beta \int_{R_j(h)}^{\bar{p}} \frac{f'_p(h, p') \hat{\Gamma}(p') dp'}{r + \hat{\delta}_j + \beta s(h, p') \lambda \hat{\Gamma}(p')}.
\end{aligned}$$

## A.2 Wage 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 agree to renegotiate. 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 and a worker who is renegotiating with an outside offer in hand, 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). All the following arguments are for a given  $(j, h)$ -type of worker and so to save on notation, the functional dependence on  $(j, h)$  is suppressed. Specifically, the bargaining outcome of an unemployed worker maximizes the Nash product,

$$\{w_0(p), s(p)\} = \arg \max_{w, s} (\tilde{V}(p, w, s) - V^0)^\beta \tilde{J}(w, p, s)^{(1-\beta)}, \quad (\text{A.3})$$

which yields the worker valuation,

$$V(R, p) = \beta V(p) + (1 - \beta) V^0. \quad (\text{A.4})$$

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

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,  $V(q) = V(q, q)$ . Hence, the employment contract that results from this bargaining setting is,

$$\{w(q, p), s(p)\} = \arg \max_{w, s} (\tilde{V}(p, w, s) - V(q))^\beta \tilde{J}(w, p, s)^{(1-\beta)}. \quad (\text{A.5})$$

The bargaining outcome is,

$$V(q, p) = \beta V(p) + (1 - \beta)V(q). \quad (\text{A.6})$$

In both cases (A.3) and (A.5), the agreed upon search intensity  $s(p)$  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.

### A.2.1 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  $\tilde{V}(p, w_e, s_e)$  and the firm receives  $\tilde{J}(p, w_e, s_e) = \tilde{V}(p, f(p), s_e) - \tilde{V}(p, w_e, s_e)$ . If the firm rejects the offer, the bargaining breaks down with exogenous probability  $\Delta$ . If so, the firm receives a zero payoff and the worker goes back to unemployment and receives  $V^0$ . 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  $\beta$ . If the firm makes the offer and the worker accepts, the worker receives  $\tilde{V}(p, w_f, s_f)$  and the firm receives  $\tilde{J}(p, w_f, s_f) = \tilde{V}(p, f(p), s_f) - \tilde{V}(p, w_f, s_f)$ . 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  $\beta$  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,  $s_e = s_f = s(p)$ , where  $s(p) = \arg \max_s \tilde{V}(p, f(p), s)$ . Furthermore, consider the strategies where the worker accepts any offer  $(w, s)$  such that  $\tilde{V}(p, w, s) \geq \tilde{V}(p, w_f, s(p))$  and rejects any offer such that  $\tilde{V}(p, w, s) < \tilde{V}(p, w_f, s(p))$ . Similarly, the firm accepts any offer  $(w, s)$  such that  $\tilde{J}(p, w, s) \geq \tilde{J}(p, w_e, s(p))$  and rejects any offer such that  $\tilde{J}(p, w, s) < \tilde{J}(p, w_e, s(p))$ .

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

$$\tilde{V}(p, w, s) \leq \tilde{V}(p, w_e, s(p)) - \tilde{V}(p, f(p), s(p)) + \tilde{V}(p, f(p), s). \quad (\text{A.7})$$

It is seen that the right hand side of the firm acceptance condition (A.7) is maximized for  $s = s(p)$  and does not depend on  $w$ . Hence, any worker deviation  $s'_e \neq s_e = s(p)$  that will be accepted by the firm must result in a worker payoff  $\tilde{V}(p, w, s'_e) < \tilde{V}(p, w_e, s(p))$ , for any  $w$ , which is not profitable.

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

$$\tilde{J}(p, w, s) \leq \tilde{V}(p, f(p), s) - \tilde{V}(p, w_f, s(p)). \quad (\text{A.8})$$

It is seen that the right hand side of the worker acceptance decision (A.8) is maximized for  $s = s(p)$  and that it does not depend on  $w$ . Hence, any firm deviation  $s'_f \neq s_f = s(p)$  that will be accepted by the worker must result in a firm payoff  $\tilde{J}(p, w, s'_f) < \tilde{J}(p, w_f, s_f)$ , for any  $w$ , which is not profitable.

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

Now consider potential deviations in the wage. The worker's payoff  $\tilde{V}(p, w, s_e)$  is monotonically increasing in  $w$ . It follows directly from (A.7) 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,

$$\hat{V}(w_f) = (1 - \Delta)[\beta\hat{V}(w_e) + (1 - \beta)\hat{V}(w_f)] + \Delta V^0 \quad (\text{A.9})$$

$$\hat{J}(w_e) = (1 - \Delta)[\beta\hat{J}(w_e) + (1 - \beta)\hat{J}(w_f)] \quad (\text{A.10})$$

where  $\hat{V}(w) = \tilde{V}(p, w, s(p))$  and  $\hat{J}(w) = \tilde{J}(p, w, s(p))$ . Equations (A.9) and (A.10) can be rewritten as,

$$\beta[\hat{V}(w_f) - \hat{V}(w_e)] = \Delta[V_0(h) - \beta\hat{V}(w_e) - (1 - \beta)\hat{V}(w_f)] \quad (\text{A.11})$$

$$(1 - \beta)[\hat{J}(w_f) - \hat{J}(w_e)] = \Delta[\beta\hat{J}(w_e) + (1 - \beta)\hat{J}(w_f)]. \quad (\text{A.12})$$

Taking the limit as  $\Delta \rightarrow 0$ , equations (A.9) and (A.10) imply that  $w_f \rightarrow w_e$ . Denote the common

limit by  $w$ . Hence,

$$\begin{aligned}\frac{\partial \hat{V}(w)}{\partial w} &= \lim_{\Delta \rightarrow 0} \frac{\hat{V}(w_f) - \hat{V}(w_e)}{w_f - w_e} \\ \frac{\partial \hat{J}(w)}{\partial w} &= \lim_{\Delta \rightarrow 0} \frac{\hat{J}(w_f) - \hat{J}(w_e)}{w_f - w_e}.\end{aligned}$$

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  $\Delta \rightarrow 0$  in equations (A.11) and (A.12) yields,

$$\begin{aligned}-\frac{\beta}{1-\beta} &= \frac{V^0 - \beta \hat{V}(w) - (1-\beta) \hat{V}(w)}{\beta \hat{J}(w) + (1-\beta) \hat{J}(w)} \\ &\Downarrow \\ \hat{V}(w) &= \beta \hat{V}(f(p)) + (1-\beta) V^0.\end{aligned}\tag{A.13}$$

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 (A.4).

### A.2.2 Employed worker

Cahuc et al. (2006) provide a strategic bargaining foundation for the axiomatic Nash bargaining outcome in equation (A.6). 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 (A.6).

### A.3 Proof of Lemma 3

In this we suppress layoff rate heterogeneity for notational simplicity. 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 (A.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,

$$\begin{aligned}
rV(h, p) &= f(h, p) - c(s(h, p)) + [\delta_1 + \delta_0\lambda\Gamma(R(h))]V^0(h) + s(h, p)\lambda \int_p^{\bar{p}} V(h, p, p')d\Gamma(p') \\
&\quad + \delta_0\lambda \int_{R(h)}^{\bar{p}} V(h, R(h), p')d\Gamma(p') - [\hat{\delta}_j + s(h, p)\lambda\hat{\Gamma}(p)]V(h, p) \\
&= f(h, p) - c(s(h, p)) + [\delta_1 + \delta_0\lambda[1 - \beta + \beta\Gamma(R(h))]]V^0(h) \\
&\quad + \delta_0\lambda\beta \int_{R(h)}^{\bar{p}} V(h, p')d\Gamma(p') + s(h, p)\lambda\beta \int_p^{\bar{p}} V(h, p')d\Gamma(p') \\
&\quad - [\hat{\delta}_j + \beta s(h, p)\lambda\hat{\Gamma}(p)]V(h, p).
\end{aligned}$$

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

$$\begin{aligned}
V(h, p) &= \max_{s \geq 0, R \in [b, \bar{p}]} \left\{ \frac{f(h, p) - c(s) + \delta_1 V^0(h) + \beta s \lambda \int_p^{\bar{p}} V(h, p') d\Gamma(p')}{r + \hat{\delta} + \beta s \lambda \hat{\Gamma}(p)} \right. \\
&\quad \left. + \delta_0 \lambda \frac{V^0(h) + \beta \int_R^{\bar{p}} [V(h, p') - V_0(h)] d\Gamma(p')}{r + \hat{\delta} + \beta s \lambda \hat{\Gamma}(p)} \right\}, \tag{A.14}
\end{aligned}$$

where

$$rV^0(h) = \max_{s \geq 0, R \in [0, 1]} \left\{ f(h, 0) - c(s) + (\mu + \kappa s)\lambda\beta \int_R^{\bar{p}} [V(h, p') - V^0(h)] d\Gamma(p') \right\}. \tag{A.15}$$

It is straightforward to show that the fixed point of the mapping in equation (A.15) satisfies,

$$V^0(h) = \max_{s \geq 0, R \in [0, 1]} \left\{ \frac{f(h, 0) - c(s) + (\mu + \kappa s)\lambda\beta \int_R^{\bar{p}} V(h, p') d\Gamma(p')}{r + (\mu + \kappa s)\lambda\beta\hat{\Gamma}(R)} \right\}. \tag{A.16}$$

This then establishes a unique solution to equation (A.15). Furthermore, inspection of equation (A.16) reveals that if  $V(h, p)$  is increasing in  $h$ , then  $V^0(h)$  is strictly increasing in  $h$ . Equation (A.14) is a contraction. Denote the mapping  $T : \mathcal{F} \rightarrow \mathcal{F}$ , 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 (A.14) for  $V(h, p)$  and any  $h$  in the support of  $\Psi(\cdot)$ . Finally, let  $V^0(h)$  be defined by

equation (A.16) for the value of employment given by  $V(h, p)$ . It then follows that,

$$\begin{aligned}
(TV)(h_0, p) &= \frac{f(h_0, p) - c(s(h_0, p)) + \delta_1 V_0(h_0) + \beta s(h_0, p) \lambda \int_p^{\bar{p}} V(h_0, p') d\Gamma(p')}{r + \hat{\delta} + \beta s(h_0, p) \lambda \hat{\Gamma}(p)} \\
&\quad + \delta_0 \lambda \frac{V^0(h_0) + \beta \int_{R(h_0)}^{\bar{p}} [V(h_0, p') - V^0(h_0)] d\Gamma(p')}{r + \hat{\delta} + \beta s(h_0, p) \lambda \hat{\Gamma}(p)} \\
&< \frac{f(h_0, p) - c(s(h_0, p)) + \delta_1 V^0(h_1) + \beta s(h_0, p) \lambda \int_p^{\bar{p}} V(h_1, p') d\Gamma(p')}{r + \hat{\delta} + \beta s(h_0, p) \lambda \hat{\Gamma}(p)} \\
&\quad + \delta_0 \lambda \frac{V^0(h_1) + \beta \int_{R(h_0)}^{\bar{p}} [V(h_1, p') - V^0(h_1)] d\Gamma(p')}{r + \hat{\delta} + \beta s(h_0, p) \lambda \hat{\Gamma}(p)} \\
&\leq \frac{f(h_0, p) - c(s(h_1, p)) + \delta_1 V^0(h_1) + \beta s(h_1, p) \lambda \int_p^{\bar{p}} V(h_1, p') d\Gamma(p')}{r + \hat{\delta} + \beta s(h_1, p) \lambda \hat{\Gamma}(p)} \\
&\quad + \delta_0 \lambda \frac{V^0(h_1) + \beta \int_{R(h_1)}^{\bar{p}} [V(h_1, p') - V^0(h_1)] d\Gamma(p')}{r + \hat{\delta} + \beta s(h_1, p) \lambda \hat{\Gamma}(p)} \\
&= (TV)(h_1, p).
\end{aligned}$$

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 (A.14) is strictly increasing in  $h$ . This establishes Lemma 3.

#### A.4 The firm's vacancy choice

$\Lambda_j(h, p)$  is the likelihood of meeting an employed skill level  $h$ , layoff rate  $\delta_j$  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  $\delta_j$  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. By the assumption of proportionality in matching, they are defined by,

$$\Lambda_j(h, p) = \frac{\tilde{\zeta}_j (1 - u_j) s_j(h, p) g_j(h, p)}{\sum_{j' \in \{L, H\}} \tilde{\zeta}_{j'} \int_0^1 \left\{ u_{j'} [\mu + \kappa s_{j'}^0(h')] v_{j'}(h') + (1 - u_{j'}) \int_0^1 [\delta_0 + s_{j'}(h', p')] g_{j'}(h', p') dp' \right\} dh'}$$

and

$$\Lambda_j^0(h) = \frac{\tilde{\zeta}_j \left\{ u_j [\mu + \kappa s_j^0(h)] v_j(h) + (1 - u_j) \delta_0 \int_0^1 g_j(h, p) dp \right\}}{\sum_{j' \in \{L, H\}} \tilde{\zeta}_{j'} \int_0^1 \left\{ u_{j'} [\mu + \kappa s_{j'}^0(h)] v_{j'}(h) + (1 - u_{j'}) \int_0^1 [\delta_0 + s_{j'}(h, p)] g_{j'}(h, p) dp \right\} dh}$$



The first order condition for the firm's vacancy intensity choice is given by,

$$c'_v(v(p)) = \eta(1-\beta) \sum_{j \in \{L, H\}} \int_0^1 \left\{ \left[ V_j(h, p, p) - V_j(h, R_j(h), R_j(h)) \right] \Lambda_j^0(h) + \int_{R_j(h)}^p \left[ V_j(h, p, p) - V_j(h, p', p') \right] \Lambda_j(h, p') dp' \right\} dh.$$

### A.5 Steady state $G(h, q, p)$

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

$$\begin{aligned} (1-u)\delta G(h, q, p) + (1-u)\lambda(\theta) \int_{\underline{h}}^h \int_{R(h')}^q \left\{ (1-\Gamma(p)) \int_{q'}^q s(h', p') dG(h', q', p') \right. \\ \left. + (1-\Gamma(q)) \int_q^p s(h', p') dG(h', q', p') \right\} = \\ \int_{\underline{h}}^h I(R(h') \leq q) [\Gamma(p) - \Gamma(R(h'))] \lambda(\theta) \left[ u[\delta_0 + \kappa s_0(h')] v(h') + \right. \\ \left. (1-u)\delta_0 \int_{R(h')}^{\bar{p}} \int_{q'}^{\bar{p}} g(h', q', p') dp' dq' \right] dh'. \quad (\text{A.17}) \end{aligned}$$

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

$$\begin{aligned} (\delta_0 \lambda(\theta) + \delta_1)(1-u) \int_{R(h)}^{\bar{p}} \int_{q'}^{\bar{p}} g(h, q', p') dp' dq' &= [1 - \Gamma(R(h))] \lambda(\theta) \left\{ u[\mu + \kappa s_0(h)] v(h') + \right. \\ &\quad \left. (1-u)\delta_0 \int_{R(h)}^{\bar{p}} \int_{q'}^{\bar{p}} g(h, q', p') dp' dq' \right\} \\ &\Downarrow \\ (\delta_0 \lambda(\theta) \Gamma(R(h)) + \delta_1)(1-u) \int_{R(h)}^{\bar{p}} \int_{q'}^{\bar{p}} g(h, q', p') dp' dq' &= u[1 - \Gamma(R(h))] \lambda(\theta) [\mu + \kappa s_0(h)] v(h) \\ &\Downarrow \\ \delta_0(1-u) \int_{R(h)}^{\bar{p}} \int_{q'}^{\bar{p}} g(h, q', p') dp' dq' &= \frac{\delta_0 \lambda(\theta) [1 - \Gamma(R(h))]}{\delta_0 \lambda(\theta) \Gamma(R(h)) + \delta_1} u[\mu + \kappa s_0(h)] v(h). \quad (\text{A.18}) \end{aligned}$$

Insert this into equation (A.17),

$$\begin{aligned} \frac{\delta_0 \lambda(\theta) + \delta_1}{\lambda(\theta)} G(h, q, p) + \int_{\underline{h}}^h \int_{R(h')}^q \left[ [1 - \Gamma(p)] \int_{q'}^q s(h', p') dG(h', q', p') \right. \\ \left. + [1 - \Gamma(q)] \int_q^p s(h', p') dG(h', q', p') \right] = \\ \frac{u}{1-u} \int_{\underline{h}}^h I(R(h') \leq p) [\Gamma(p) - \Gamma(R(h'))] [\mu + \kappa s_0(h')] v(h') \frac{\delta_1 + \delta_0 \lambda(\theta)}{\delta_0 \lambda(\theta) \Gamma(R(h)) + \delta_1} dh'. \quad (\text{A.19}) \end{aligned}$$

Evaluate (A.19) at  $(\bar{h}, \bar{p}, \bar{p})$  to obtain,

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

One then obtains,

$$\begin{aligned}
&\int_{\underline{h}}^h \int_{R(h')}^q \left[ \int_{q'}^q [\delta / \lambda(\theta) + [1 - \Gamma(p)] s(h', p')] g(h', q', p') dp' \right. \\
&\quad \left. + \int_q^p [\delta / \lambda(\theta) + [1 - \Gamma(q)] s(h', p')] g(h', q', p') dp' \right] dq' dh' = \\
&\frac{\delta}{\lambda(\theta)} \frac{\int_{\underline{h}}^h I(R(h') \leq q) [\Gamma(p) - \Gamma(R(h'))] \frac{\mu + \kappa s_0(h')}{\delta_0 \Gamma(R(h')) + \delta_1 / \lambda(\theta)} dY(h')}{\int_{\underline{h}}^{\bar{h}} \frac{[1 - \Gamma(R(h'))] [\mu + \kappa s_0(h')]}{\delta_0 \Gamma(R(h')) + \delta_1 / \lambda(\theta)} dY(h')}. \tag{A.20}
\end{aligned}$$

## A.6 Steady state equilibrium solution for $Y(h)$

Consider the equilibrium condition,

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

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 (A.18). By the steady state unemployment rate expression it follows that,

$$\psi(h) = \frac{\left[ 1 + \frac{[1 - \Gamma(R(h))][\mu + \kappa s_0(h)]}{\delta_0 \Gamma(R(h)) + \delta_1 / \lambda(\theta)} \right] v(h)}{\int_{\underline{h}}^{\bar{h}} \left( 1 + \frac{[1 - \Gamma(R(h'))][\mu + \kappa s_0(h')]}{\delta_0 \Gamma(R(h')) + \delta_1 / \lambda(\theta)} \right) v(h') dh'}, \tag{A.21}$$

which is an integral equation for  $Y(h)$  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 (A.21),

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

Use equation (A.21) to solve for  $1 + \int_{\underline{h}}^{\bar{h}} \Delta(h') v(h') dh'$ . First, some minor manipulation,

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

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

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

Hence, one obtains the solution,

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

which can also be written as,

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

## A.7 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  $\vec{n} = (n_1, n_2, \dots, n_k)$  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,

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

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 Estimation

## B.1 Overfitting $Corr(\bar{r}_i^u, w_i^0 | \hat{\tau} \geq 0.95)$

Our Indirect Inference estimator minimizes the distance between real and simulated moments, measured as a weighted Euclidean distance, with the weighting matrix being the inverse variance-covariance matrix of the estimated vector of moments. This is the optimal weighting matrix in that it induces efficiency, but effectively, it puts relatively less weight on moments that are imprecisely estimated. Hence, precision may come at the expense of a potentially deteriorated fit, an issue that is salient for the auxiliary statistic that formally pins down production technology complementarities,  $Corr(\bar{r}_i^u, w_i^0 | \hat{\tau} \geq 0.95)$ . The computation of  $Corr(\bar{r}_i^u, w_i^0 | \hat{\tau} \geq 0.95)$  is very taxing on the data as it utilizes only a small number of firms, the most productive ones, and within these firms, only a small number of workers, those who initiate their jobs with a transition from unemployment. The implication is that the the covariance between starting wages and unemployment duration in very productive firms is relatively imprecisely estimated, and thus only carries a light weight in the main estimation. It is therefore not surprising that our main estimation face difficulties in fitting this moment precisely. As a robustness check on our estimates and subsequent analysis and conclusion, this appendix presents a set of estimates where we artificially scale up the weight put on the covariance between starting wages and unemployment duration in very productive firms. This forces a near-perfect fit to the central moment identifying production function complementarity.

Specifically, let  $\Sigma$  be the weight matrix in the Indirect inference estimator. In the main estimation, we take  $\Sigma = \widehat{\mathbf{W}}^{-1}$ , where  $\mathbf{W}$  is the variance-covariance matrix of the vector of auxiliary statistics and data moments used in the estimation, denoted  $\mathbf{a}$ . Suppose  $Corr(\bar{r}_i^u, w_i^0 | \hat{\tau} \geq 0.95)$  is the  $k$ th entry in the moment vector. We conduct a separate estimation using an augmented weight matrix  $\Sigma$  which is equal to  $\widehat{\mathbf{W}}^{-1}$  except that the  $(k, k)$ -entry has been scaled up by a factor 50,000. The Indirect Inference estimator is consistent for any choice of weighting matrix, as long as it is symmetric and positive definite. The augmented weighting matrix retains these properties, and delivers a consistent, albeit overfitted, estimate of the structural parameters. For the augmented  $\sqrt{N}(\hat{\omega} - \omega_0) \rightarrow^d \mathcal{N}(\mathbf{0}, (1 + S^{-1})[\mathbf{J}'\Sigma\mathbf{J}]^{-1}\mathbf{J}'\Sigma\widehat{\mathbf{W}}\Sigma\mathbf{J}[\mathbf{J}'\Sigma\mathbf{J}]^{-1})$ , where  $\mathbf{J} = \partial\mathbf{a}(\omega)/\partial\omega$  evaluated at  $\hat{\omega}$ , and from which we obtain the reported standard errors.

### B.1.1 Structural Parameter Estimates

The structural parameters, estimated with the augmented weighting matrix, are presented in Table 1, alongside our preferred parameter estimates reported in Table 2 for comparison. The two sets of parameters are largely identical. In the augmented estimation, the elasticity of search

cost with respect to search effort is 1.068 in comparison to 1.077 in the main estimation, with the relative off-the-job to on-the-job search intensity  $\kappa$ , the rate at which offers arrive exogenously  $\underline{s}$ , the recruitment cost function  $c_v(v)$ , and the reallocation rate  $\delta_0$  also being very similar in the two estimations. The same holds true for the estimated vacancy and worker heterogeneity distributions  $\Gamma(p)$  and  $\Psi(h)$ , see Figure B.1 for a graphical rendition, and also for the estimated parameters of the CES match production function  $f(h, p)$ . In particular, the overfitted estimates, forcing a near-perfect fit to  $\text{Corr}(\bar{t}_i^u, w_i)$ , delivers an estimate of  $\rho$  at  $-2.198$ . This is almost identical to our preferred estimate at  $-2.045$ .

The main difference between the overfitted and main estimation is in the job destruction process, and to a lesser extent in the estimated bargaining power  $\beta$ . The overfitted estimation predicts that the vast majority of workers in the population, 93%, have relatively low job destruction risk. In our preferred estimation 86% of workers faced relatively low job destruction risk. Also consistent with the main estimation, the overfitted estimates predicts that workers with high job destruction risk face have a very weak attachment to the labor market, with the average time spent in employment between unemployment spells being approximately 6 months (among employed workers high layoff rate workers account for only 1%, similar to our main estimation). However, with overfitting, the estimated model yields a job destruction rate faced by low layoff workers,  $\delta_L$ , that is three times higher than in the preferred estimate presented in Table 2. We have  $\delta_L = 0.186$  in the augmented estimation versus  $\delta_L = 0.063$  in our preferred estimation. The overfitted estimate of workers' bargaining power parameter  $\beta$  is higher than in the main estimation, with the overfitted estimation yielding  $\beta = 0.233$  versus a preferred estimate of  $\beta = 0.177$ .

### B.1.2 Model Fit

Table 2 present the fit of the model estimated with the augmented weighting matrix, and also allow comparison to the fit of our preferred estimated model. As expected, overfitting  $\text{Corr}(\bar{t}_i^u, w_i)$  result in a perfect fit. The improved fit to this particular auxiliary statistics comes at the expense of a deterioration in the fit of a closely related moment, namely the average unemployment for worker hired into top-ranked firms, and to a less extent the standard deviation of these unemployment durations. In the overfitted estimation, we overestimate the duration of unemployment spells by a factor of almost two. For the remaining moments listed in Table 2 there is no substantial difference between the fit of our preferred estimate, and the overfitted model estimate.

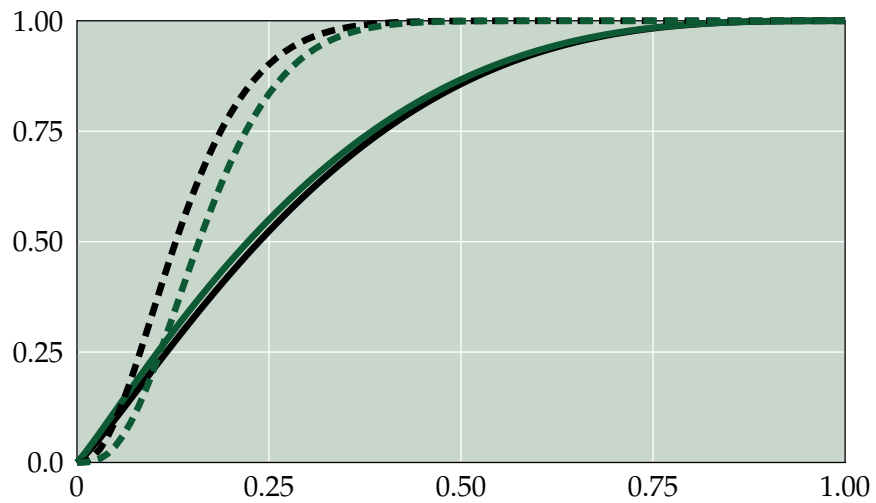
The deteriorated fit to unemployment durations is also visible in Figure B.2 which plots

Table 1: Structural Parameter Estimates—Overfitting  $Corr(\bar{t}_i^u, w_i)$

	Preferred	Overfitted
Annual job destruction rate, low type, $\delta_L$	0.063 (0.0001)	0.186 ( )
Annual job destruction rate, high type, $\delta_H$	1.905 (0.0001)	2.031 ( )
Job destruction type distribution, $\zeta_L = \Pr(\delta = \delta_L)$	0.858 (0.0001)	0.931 ( )
Search cost function $c(s) = \frac{(c_0s)^{1+1/c_1}}{1+1/c_1}$		
$c_0$	54.420 (0.0002)	77.276 ( )
$c_1$	12.911 (0.0002)	14.680 ( )
Recruitment cost function $c_v(v) = \frac{v^{1+1/c_{v1}}}{1+1/c_{v1}}$		
$c_{1v}$	0.012 (0.0001)	0.011 ( )
Exogenous search, $\underline{g}$	0.034 (0.0001)	0.034 ( )
Annual reallocation rate, $\delta_0$	0.106 (0.0001)	0.116 ( )
Off-the-job to on-the-job relative search efficiency, $\kappa$	0.845 (0.0001)	0.808 ( )
Firm productivity CDF on $p \in [0, 1]$ , $\Phi(p) = Beta(\beta_0^\Phi, \beta_1^\Phi)$		
$\beta_0^\Phi$ (scale)	1.188 (0.0001)	1.106 ( )
$\beta_1^\Phi$ (shape)	3.151 (0.0001)	3.112 ( )
Worker skill CDF on $h \in [0, 1]$ , $\Psi(h) = Beta(\beta_0^\Psi, \beta_1^\Psi)$		
$\beta_0^\Psi$ (scale)	2.638 (0.0970)	3.350 ( )
$\beta_1^\Psi$ (shape)	16.022 (0.0971)	16.352 ( )
Match production function, $f(h, p) = f_0 \left( \alpha (h + \underline{h})^\rho + (1 - \alpha)(p + \underline{p})^\rho \right)^{\frac{1}{\rho}}$		
$\rho$	-2.045 (0.0155)	-2.198 ( )
$\alpha$	0.311 (0.0001)	0.324 ( )
$f_0$	931.169 (0.0001)	909.647 ( )
Workers' bargaining power, $\beta$	0.177 (0.0001)	0.233 ( )
Std. deviation, wage measurement error, $\sigma_w$	0.094 (0.0015)	0.110 ( )

Note: Standard errors in parentheses.

Figure B.1: Firm- and Worker Heterogeneity Distributions



Firm types  $p$ , worker types  $h$

Note: The black solid line shows the estimated vacancy heterogeneity distribution  $\Gamma(p)$  for our preferred estimation. The green solid line shows  $\Gamma(p)$  for the overfitted estimation. The dashed line shows the estimated population worker heterogeneity distribution  $\Psi(h)$  for our preferred estimation. The green dashed line shows  $\Psi(p)$  for the overfitted estimation.

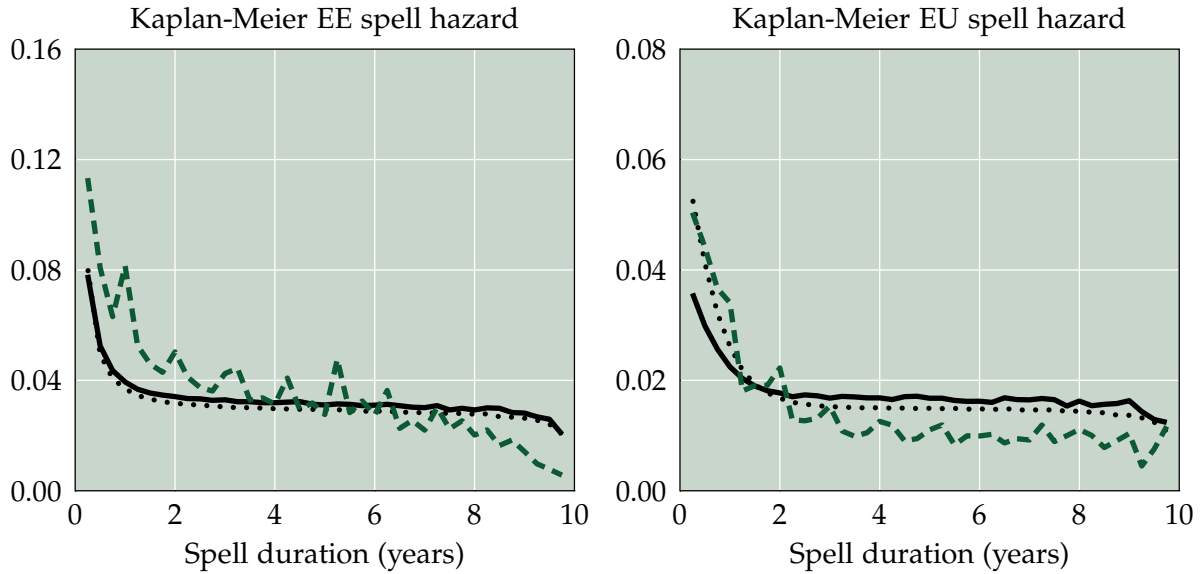


Table 2: Model Fit—Overfitting  $Corr(\bar{t}_i^u, w_i)$ 

	Data	Sim.	
		Preferred	Overfitted
<b>Labor Market Transitions</b>			
Number of jobs in employment cycle, average	2.182 ( )	2.178	2.230
Number of jobs in employment cycle, std. dev.	1.541 ( )	1.260	1.281
Average share of matches in a cross section ending in EU-transition	0.338 ( )	0.348	0.347
<b>Cross Section Heterogeneity</b>			
Log firm wage, employment weighted average	5.254 ( )	5.252	5.250
Log firm wage, employment weighted std. dev.	0.168 ( )	0.160	0.171
Log firm wage, newly hired workers average	5.167 ( )	5.166	5.170
Log firm wage, newly hired workers std. dev.	0.222 ( )	0.201	0.227
Firm size in FTE, average	8.646 ( )	8.874	8.509
Fraction of active firms to worker population	0.091 ( )	0.089	0.090
Within-job annual log wage growth, average	0.009 ( )	0.005	0.005
Firm effects from auxiliary log wage regression (??), average	5.238 ( )	5.217	5.219
Firm effects from auxiliary log wage regression (??), std. dev.	0.179 ( )	0.149	0.161
Worker effects from auxiliary log wage regression (??), std. dev.	0.218 ( )	0.220	0.199
Residuals from auxiliary log wage regression (??), std. dev.	0.134 ( )	0.129	0.135
Mean-min wage ratio	1.854 ( )	1.799	1.797
<b>Labor Market Sorting</b>			
Unemployment duration (in weeks) for workers hired into top ranked firms, average	57.395 ( )	78.466	104.774
Unemployment duration (in weeks) for workers hired into top ranked firms, std. dev.	69.853 ( )	80.484	88.066
Starting wage (in DKK) for workers hired into top ranked firms, average	186.0 ( )	172.8	180.1
Starting wage (in DKK) for workers hired into top ranked firms, std. dev.	62.3 ( )	44.8	45.5
Correlation(unemployment duration, starting wage) for workers hired into top ranked firms	-0.168 ( )	-0.381	-0.168

Note: Standard errors obtained by block-bootstrap in parentheses.

Figure B.2: Model Fit with Overfitting—Quarterly Kaplan-Meier Employment Hazards.



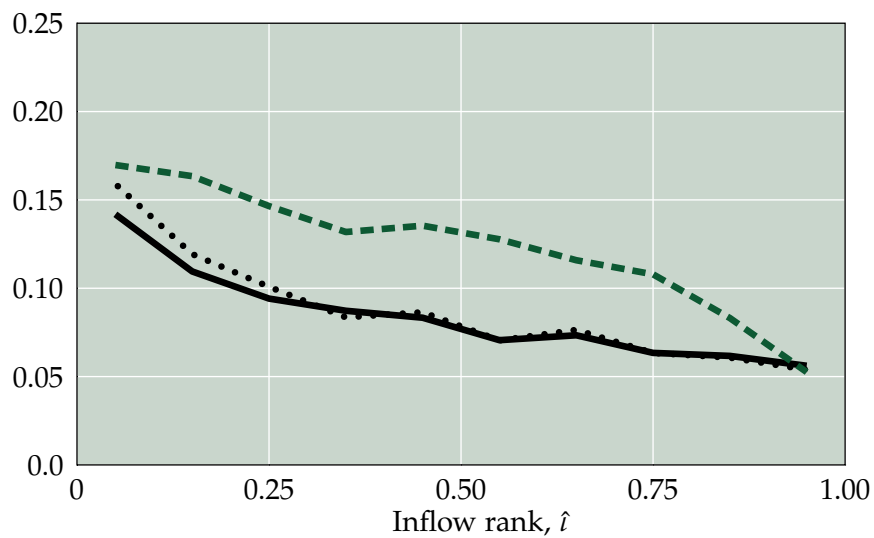
Note: Data in dashed line. Preferred model estimate in dotted line. Overfitted model estimate in solid line.

Kaplan-Meier estimates of employment hazard functions, split in competing risk job-to-job and job-to-unemployment transition hazards. The the overfitting has minimal impact on the fit to the Kaplan-Meier job-to-job transition hazard rates. However, the fit to the Kaplan-Meier job-to-unemployment transition hazard function is significantly worsened. In particular, the model estimate based on the augmented weighting matrix is underestimates the hazard rate at job durations up to two years. For longer job durations, the fit of the overfitted model is similar to our preferred estimate.

Figure B.3, showing job-to-job transition hazard rates as a function of the inflow rank  $\hat{t}_i$ , further confirms that the job-to-job transition process is largely unaffected by the overfitting of  $Corr(\bar{t}_i^u, w_i)$ . The model estimate obtained using the augmented weighting matrix basically coincides with the hazard rate profile predicted by our preferred estimate, except perhaps at the least productive firms, where underestimation of the hazard rate is slightly more pronounced in the with overfitting than without.

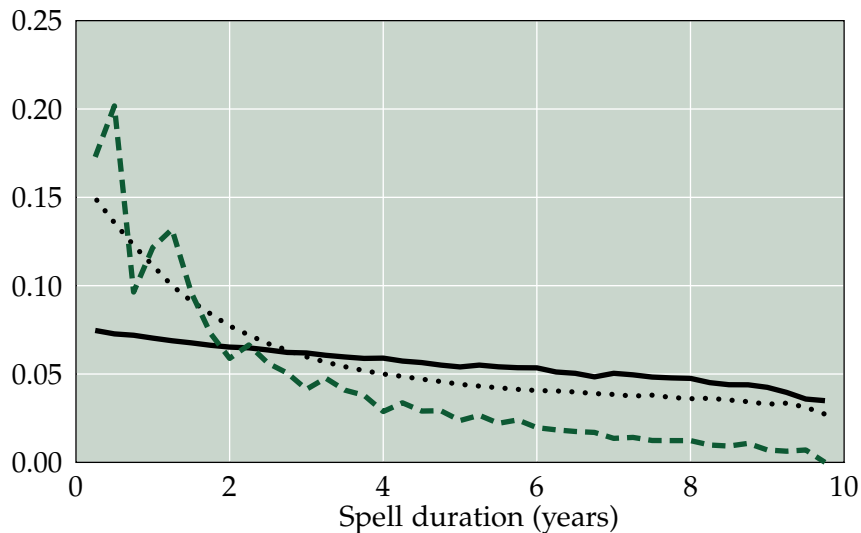
Figure B.4, plotting empirical and simulated unemployment-to-job transition hazard functions, confirms the overfitted model's problems in reproducing observed unemployment duration data. Even if we did not use the unemployment-to-job transition hazard functions as a

Figure B.3: Model Fit with Overfitting—Quarterly Inflow Rank Conditional Job-to-job Transition Hazard Rates



Note: The green dashed line shows job-to-job transition hazard rates estimated on real data (Gaussian non-parametric regression with bandwidth 0.02). The solid black line shows simulated job-to-job transition hazard rates for the estimated model equilibrium (preferred estimates). The black dotted line shows simulated job-to-job transition hazard rates for the estimated model equilibrium with overfitting.

Figure B.4: Model Fit with Overfitting—Quarterly Kaplan-Meier Unemployment-to-job Hazard.



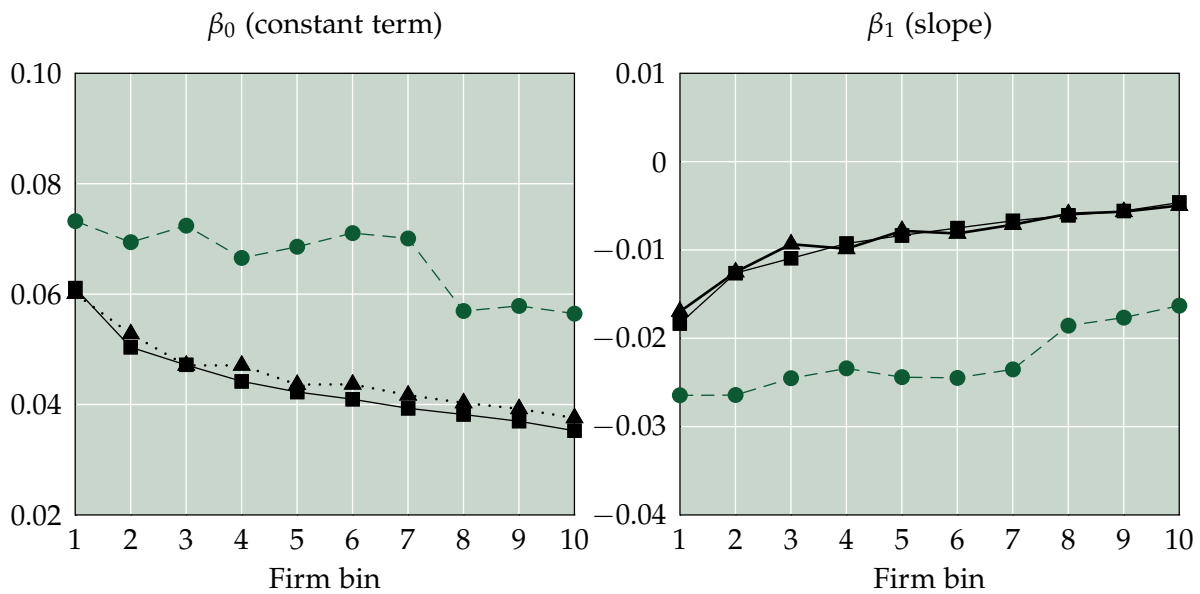
Note: Data in green dashed line. Preferred model estimate in black dotted line. Overfitted model fit in black solid line.

target in the estimation, our preferred estimate does a notably good job in fitting these. In contrast, Figure B.4 reveals that the overfitted model does a poor job at predicting the strong duration dependence at relatively short unemployment durations (less than 2 years), where the model severely underestimates the unemployment-to-job transition hazard rate. At longer durations, the overfitted model overestimates the hazard rate, even more than our preferred model estimate.

Overfitting  $\text{Corr}(\bar{T}_i^u, w_i)$  was done to force a near-perfect fit to a particular auxiliary statistic that directly help pin down production function complementarities in the data. Our estimation also include another set of auxiliary statistics that directly informs on labor market sorting, namely the regression coefficients  $\beta_{0k}$  and  $\beta_{1k}$  from the regression (5.3). Figure B.5 plots the fit of the augmented estimation to this set of moments (along with the real data counterpart and the fit of our preferred estimation). We note that our preferred and overfitted model estimate virtually coincide.

Overall, overfitting  $\text{Corr}(\bar{T}_i^u, w_i)$  in the estimation (naturally) leads to a perfect fit to this auxiliary statistic. The improved fit obtained at the expense of a worsened fit to the distribution of unemployment spell durations. Other moments are not affected in any substantial way.

Figure B.5: Model Fit with Overfitting—Inflow Rank Conditional Job-to-job Transition Hazard Functions



Note: Left panel: Green dots represent the constant term  $\beta_0$  estimated on real data, black squares represent the constant term  $\beta_0$  estimated on simulated data (preferred estimate), black triangles represent the constant term  $\beta_0$  estimated on simulated data (overfitted estimate). Right panel: Green dots represent the slope term  $\beta_1$  estimated on real data, black squares represent the slope term  $\beta_1$  estimated on simulated data (preferred estimate), black triangles represent the constant term  $\beta_0$  estimated on simulated data (overfitted estimate).

Table 3: Log Wage Variance Decomposition with Overfitting—The AKM Approach

	Data		Sim.			
	Value	Percent of $Var(\ln w_{it})$	Value	Percent of $Var(\ln w_{it})$	Value	Percent of $Var(\ln w_{it})$
$Var(\ln w_{it})$	0.097	100%	0.087	100%	0.082	100%
$Var(\chi_i)$	0.070	71%	0.056	64%	0.047	57%
$Var(\varphi_{J(i,t)})$	0.014	14%	0.011	13%	0.015	18%
$Var(\varepsilon_{it})$	0.015	15%	0.015	17%	0.016	20%
$2Cov(\chi_i, \varphi_{J(i,t)})$	-0.002	0%	0.005	6%	0.004	5%

Note: Data refers to empirical results. Sim. refers to results obtained on data simulated from the estimated structural model (preferred and overfitted).

### B.1.3 Log Wage Variance Decompositions

Finally, we check the sensitivity of our log wage variance decompositions to overfitting  $Corr(\bar{r}_i^u, w_i)$ .

**AKM Regressions** Table 3 presents log wage variance decompositions using the AKM approach. The table contains the decomposition on data, for our preferred estimate (both also reported in Table 4), and for the overfitted estimates. The overfitted model generates slightly less log wage variation than our preferred estimate, which underestimates the variation observed in the data. However, the relative importance of the four components in the AKM log wage variance decomposition is not substantially affected by the overfitting. According the overfitted AKM log wage decomposition, 57% of log wage variation is comprised of worker heterogeneity, 18% of firm heterogeneity, 5% is wage sorting, and 20% is comprised of residual log wage variation. The corresponding AKM shares for our preferred estimate are 64%, 13%, 6% and 17%, respectively.

**Accounting for Labor Market Sorting** Next, we consider our main log-wage variance decomposition (??), which decompose log wage variance into components coming from firm heterogeneity, worker heterogeneity, labor market frictions and labor market sorting. Table 4 contains this decomposition. To facilitate comparison we also present the decomposition for our main estimation (otherwise to be found in Table ??). First, we note again that the overfitted model estimate generates less wage dispersion, with total simulated log wage variance at 0.071 relative to 0.079 in the main estimation. However, the log wage projections in the two estimated models comprise the same share of the simulated variance, about 93%.

Overfitting has very little impact on the log wage variance decomposition. The main source

Table 4: Log-Wage Variance Decomposition with Overfitting—Accounting for Labor Market Sorting

	Preferred		Overfitted	
	Value	Percent of $Var(\ln w_{it})$	Value	Percent of $Var(\ln w_{it})$
Worker effect	0.037	51%	0.031	47%
Firm effect	0.008	11%	0.011	17%
Friction effect	0.017	23%	0.016	24%
Sorting effect	0.011	15%	0.008	12%
Total predicted variance	0.073	100%	0.066	100%
Total simulated variance	0.079	108%	0.071	108%

of variation is still worker heterogeneity (47% in the augmented estimation versus 51% in our preferred estimation), followed by labor market frictions (23% versus 24%). In the main estimation the labor market sorting generates a slightly higher share of wage variation than firm heterogeneity. With overfitted estimates, the patterns is reversed, and firm heterogeneity account for a slightly larger share of log wage variation than labor market sorting. Still, the actual shares are comparable, with firm heterogeneity accounting for 11% in the main estimation and 17% with overfitted estimates. The shares for labor market sorting are 15% (preferred) and 12% (overfitted).

In conclusion, overfitting  $Corr(\bar{t}_i^u, w_i)$  leads to relatively minor changes in the structural parameter estimates (as expected as both our preferred estimates and the overfitted estimates are obtained from consistent estimators), and a near-perfect fit to  $Corr(\bar{t}_i^u, w_i)$  at the expense of a deteriorated fit to the distribution of unemployment durations. The fit to remaining moments is largely unaffected. Our log wage variance decompositions are robust with respect to the overfitting. That is, the fact that our preferred estimates does not exactly reproduce  $Corr(\bar{t}_i^u, w_i)$  as observed in the data, does not in any substantial way impact the conclusions we draw from the model regarding the degree of sorting in the labor market, and its implication for the observed wage distribution.

## B.2 Stratification by Education

In this appendix we present estimated models, fit analysis, and log wage variance decompositions on data stratified in high- and low-educated workers. The conclusions if the main text are robust to stratification by education.

### B.2.1 Stratified Data

We split our analysis data as described in Section 4 into high and low educated workers. A high educated worker has 15 or more years of education. This amounts to at least an undergraduate college degree or equivalent. We denote the remaining workers low educated workers, even if they may have e.g. high school or vocational qualifications. Our treatment of the stratified data is identical to the treatment of the full sample; see section 4 for details. Table 5 contains basic descriptive statistics for the two strata, as well as the pooled data.

### B.2.2 Structural Parameter Estimates

Table 6 presents the estimates of the structural parameters for high and low educated workers, and for comparison purposes, also the estimates obtained on the pooled data (also reported in Table 2 in the main text). The structural parameters are precisely estimated.

Our estimates for the job destruction process again clearly identifies two latent types workers with very different labor market attachment for each education group: high and low layoff workers. The high layoff workers have extremely high job destruction rates, estimated to be 2.087 and 1.949 per year for high and low education workers, respectively. Expected duration in employment between two unemployment spells is thus only about 6 months for the high layoff types, independent of length of education. The low layoff type workers in the two education groups face annual job destruction rates of 0.152 for high education workers, and 0.075 for low education workers. It is perhaps surprising that the low layoff type workers with low education face a lower job destruction than the low layoff types with high education. Notice however, that the share of high layoff types is much higher among low education workers, at 0.137, than among high education workers, where the share of high layoff types is 0.041. The resulting average annual job destruction rate in the populations of high and low education workers are 0.0231 (high education workers) and 0.332 (low education workers). Hence, on average, education insures against labor market risk as measured by the likelihood of job destruction.

The search cost for low education workers is slightly more elastic with respect to search intensity than for of high education workers, with elasticities estimated at 1.10 and 1.07 for low and high education workers, respectively. These elasticities are very similar to that estimated on pooled data. The recruitment cost function is highly convex for both groups, although slightly more so for high education workers.

The exogenous level of search  $\underline{s}$  and the relative efficiency of off-the-job to on-the-job search  $\kappa$  are also similar across education groups, and similar to the estimates obtained on the pooled



Table 5: Stratification by Education—Analysis Data Summary Statistics

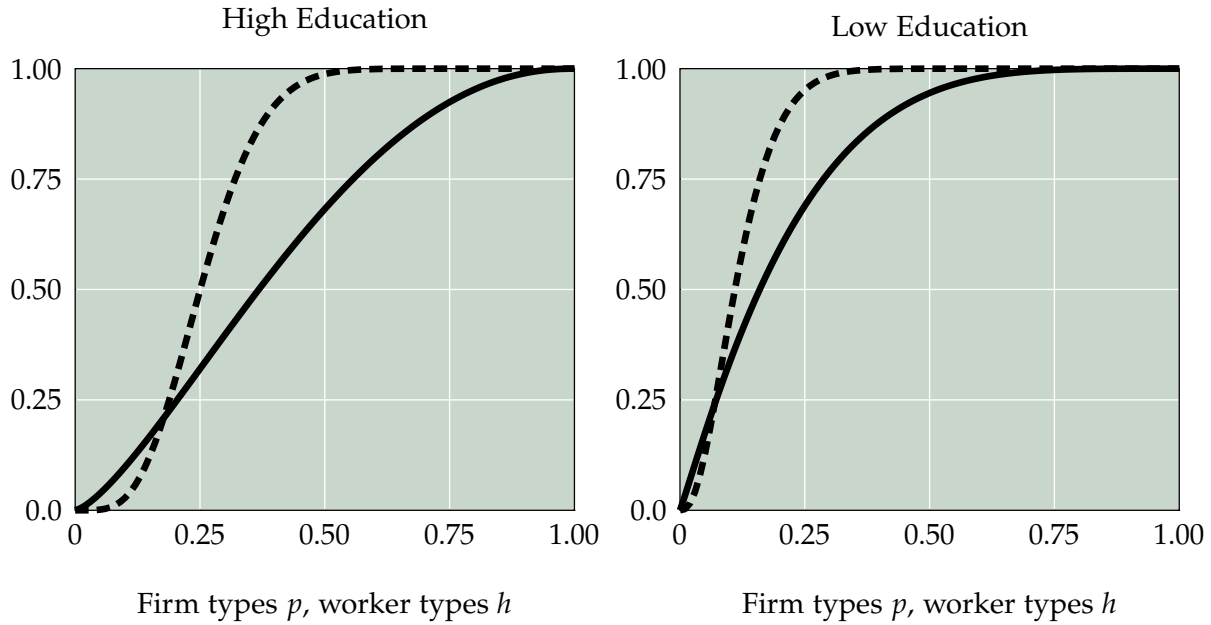
	Pooled data			High Education			Low Education		
	All years	1994	2003	All years	1994	2003	All years	1994	2003
Number of observations	6,815,884	658,465	703,707	759,511	62,172	86,156	6,056,373	596,293	617,551
Number of individuals	782,951	552,869	588,643	97,197	53,729	75,208	685,754	499,140	513,435
Number of job spells	1,698,990	490,309	511,604	191,662	49,617	67,559	1,507,328	440,691	444,045
Number of unemployment spells	608,065	168,155	192,102	61,278	12,554	18,596	546,787	155,601	173,505
Number of firms	117,847	53,537	58,210	27,783	10,664	12,671	113,686	51,789	55,744
Number of firm-years	559,920	53,537	58,249	118,646	10,664	12,671	539,619	51,789	55,744

Table 6: Structural Parameter Estimates—Stratification by Education

	Pooled data	High Education	Low Education
Annual job destruction rate, low type, $\delta_L$	0.063 (0.0001)	0.152 ( )	0.075 (0.0001)
Annual job destruction rate, high type, $\delta_H$	1.905 (0.0001)	2.087 ( )	1.949 (0.0001)
Job destruction type distribution, $\zeta_L = \Pr(\delta = \delta_L)$	0.858 (0.0001)	0.959 ( )	0.863 (0.0001)
Search cost function $c(s) = \frac{(c_0 s)^{1+1/c_1}}{1+1/c_1}$			
$c_0$	54.420 (0.0002)	62.772 ( )	35.385 (0.0002)
$c_1$	12.911 (0.0002)	14.791 ( )	10.170 (0.0002)
Recruitment cost function $c_v(v) = \frac{v^{1+1/c_1^v}}{1+1/c_1^v}$			
$c_{1v}$	0.012 (0.0001)	0.012 ( )	0.019 (0.0001)
Exogenous search, $\underline{s}$	0.034 (0.0001)	0.045 ( )	0.041 (0.0001)
Annual reallocation rate, $\delta_0$	0.106 (0.0001)	0.113 ( )	0.096 (0.0001)
Off-the-job to on-the-job relative search efficiency, $\kappa$	0.845 (0.0001)	0.764 ( )	0.834 (0.0001)
Vacancy distribution CDF on $p \in [0, 1]$ , $\Gamma(p) = \text{Beta}(\beta_0^\Gamma, \beta_1^\Gamma)$			
$\beta_0^\Gamma$ (scale)	1.188 (0.0001)	1.415 ( )	1.070 (0.0001)
$\beta_1^\Gamma$ (shape)	3.151 (0.0001)	2.214 ( )	4.342 (0.0001)
Worker skill distribution CDF on $h \in [0, 1]$ , $\Psi(h) = \text{Beta}(\beta_0^\psi, \beta_1^\psi)$			
$\beta_0^\psi$ (scale)	2.638 (0.0970)	5.149 ( )	2.728 (0.1016)
$\beta_1^\psi$ (shape)	16.022 (0.0971)	14.760 ( )	19.779 (0.1017)
Match production function, $f(h, p) = f_0 \left( \alpha(h + \underline{h})^\rho + (1 - \alpha)(p + \underline{p})^\rho \right)^{\frac{1}{\rho}}$			
$f_0$	931.169 (0.0001)	934.804 ( )	969.588 (0.0001)
$\alpha$	0.311 (0.0001)	0.434 ( )	0.471 (0.0001)
$\rho$	-2.045 (0.0001)	-1.845 ( )	-2.147 (0.0001)
Workers' bargaining power, $\beta$	0.177 (0.0155)	0.140 ( )	0.203 (0.0125)
Std. deviation, wage measurement error, $\sigma_w$	0.094 (0.0015)	0.102 ( )	0.110 (0.0018)

Note: Standard errors in parentheses.

Figure B.6: Stratified Estimation—Firm- and Worker Heterogeneity Distributions



Note: The black solid line shows the estimated vacancy heterogeneity distribution  $\Gamma(p)$ . The dashed line shows the estimated population worker heterogeneity distribution  $\Psi(h)$ .

data. The reallocation rate at 0.113 per year is slightly higher for high education workers than for low education workers where the estimated reallocation rate is 0.096. From a modeling point of view, reallocation shocks are treated as job destruction shocks without the unemployment experience, or with very short unemployment spells that are not recorded in the data. Hence, the presence of reallocation spells allow us to capture a(n extreme form of) structural duration dependence in the job finding rate among unemployment workers.

The estimated distributions of firm (i.e. vacancy) and worker heterogeneity  $\Gamma(p)$  and  $\Psi(h)$  are rendered graphically in Figure B.6 for high education (left panel) and low education (right panel) workers. We attach no comment to the estimated distributions, except noting that the estimated vacancy distribution  $\Gamma(p)$  exhibits more variation than the worker skill distribution  $\Psi(h)$  in both education groups. This was also the case for the distribution estimated on pooled data.

Given our focus on labor market sorting, the estimated match production function, measuring the degree of complementarity between worker and firm heterogeneity,  $h$  and  $p$ , in the production, is salient. Our CES specification production function specification is characterized by three parameters,  $f_0$  measuring the scale of production,  $\alpha \in [0,1]$  measuring the relative weight of labor input in the production, and  $\rho$  measuring the modularity of the production function. The estimates of  $f_0$  and  $\alpha$  are not of central interest, but are comparable across the two education

group. The estimated  $\rho$ -parameters indicate that the production function is supermodular in both education groups, inducing a matching pattern between worker and firms characterized by positive sorting. For high educated workers we find  $\rho = -1.845$ , and the resulting steady state equilibrium match distribution has  $\text{Corr}(h, p) = 0.11$ . For low educated workers, we find  $\rho = -2.147$ , implying a steady state equilibrium match distribution with  $\text{Corr}(h, p) = 0.15$ . In comparison, our estimate of  $\rho$  using all workers, resulted in  $\text{Corr}(h, p) = 0.12$ . Overall, we do not find evidence that labor market sorting differ substantially by education levels, at least when comparing the labor market of high educated workers to rest, here denoted low educated workers.

With respect to workers' bargaining power  $\beta$ , we estimate that to be 0.140 and 0.203 for high education and low education workers, respectively. This is not a true reflection of workers' bargaining power since they can use outside offers to gain larger shares of match output. Taking this into account we find that high education workers, on average in a steady state cross section, obtain XXX% of match output. The corresponding number of low education workers is XXX%. Finally, we note that the measurement error process in wages appear similar across education groups.

### B.2.3 Model Fit

The education group specific models' fit to the auxiliary statistics are reported in Table 7 and Figures B.7, B.8, and B.9. The moments used in the estimation is described in Section 5, and are identical to the set of moments used for the main estimation presented in the paper.

Starting with the moments reported in Table 7, the fit is overall good for both high and low education worker, perhaps slightly better among low educated workers. In particular, the estimated model for high education workers underestimates the employment effect in wages (here, measured as the difference between average wages in a cross section, and the average wages of newly hired workers), and overestimates starting wages for workers hired into top ranked firms. For high education workers, the estimated model accounts for about 40% of observed within-job wage growth, with the corresponding share for low education workers being 50%. This is not surprising as our model omits human capital accumulation, likely to be a particularly important source of wage growth among high education workers (see ?). In both groups, we are facing problems fitting the covariance of starting wages and unemployment durations of workers hired into top-ranked firms from unemployment. This was also the case for the main estimation in text, pooling all workers. As also pointed out there, the less than perfect fit is likely a result of this particular moment being very data demanding, and therefore imprecisely estimated. As shown

in Appendix B.1, overfitting the moment deteriorates the model's fit to unemployment durations, but is not likely to substantially change any of the conclusion of the paper. Still, we notice that the slightly stronger production function complementarities observed among low education workers is consistent with these workers having a stronger correlation between unemployment duration and starting wages in top ranked firms.

Figure B.7 plots the estimated models' fit to Kaplan-Meier job-to-job and job-to-unemployment transition hazard functions. In the data, all hazard functions exhibit negative duration dependence, albeit the duration dependence is substantially stronger for low education workers than for high education workers. For job-to-job transitions, the model fitted to data on high education workers does capture negative duration dependence qualitatively, although the simulated hazard function underestimates the job-to-job transition hazard at short job durations, and overestimates them at very long durations exceeding 8 years. Overall, it seems the model underestimates the extent of job-to-job transitions in the data. The fit to the Kaplan-Meier job-to-unemployment transition hazard rate for high education workers is good, especially at short job durations, less than 2 years, where the model captures the decline in the unconditional hazard rate very well. At longer durations, the simulated job-to-unemployment transition hazard rate slightly exceeds the empirical hazard. With respect to the model fitted to data on low education workers, the fit to the Kaplan-Meier job-to-job transition rate is good, and better than was the case for high education workers. Again, the simulated job-to-job transition hazard underpredicts the hazard rate at shorter durations, here less than 5 years, and overpredicts the hazard rate at longer durations. When it comes to the fit to the Kaplan-Meier job-to-unemployment transition hazard rate, we again obtain a good fit at shorter durations. The estimated model for low education workers captures the initial sharp decline in the Kaplan-Meier job-to-unemployment transition hazard almost perfectly. However, for job spells of duration 2 years or more, the simulated job-to-unemployment transition hazard rate exceeds the empirical rate.

In Figure B.8 we directly consider the models ability to fit the relationship between the firm ladder, as estimated by the inflow rank measure  $\hat{t}$ , and job-to-job transitions, by plotting job-to-job transition hazard rates against the inflow rank measure  $\hat{t}$ . For high education workers the empirical hazard rate drops sharply when we move from firms in the first (i.e. bottom) decile of the firm ladder, to firms in the second decile. After this initial drop, the hazard rate continues to fall throughout the firm ladder, but only modestly. The estimated model for high education workers captures perfectly the slow decline in hazard rates on the firm ladder from the second decile and up. The estimated model is, however, unable to capture the sharp decline at the very bottom of the firm ladder. With respect to low education workers, the fit to the inflow rank

Table 7: Model Fit—Stratification by Education

	High Education		Low Education	
	Data	Sim.	Data	Sim.
<b>Labor Market Transitions</b>				
Number of jobs in employment cycle, average	2.267	2.329	2.171	2.135
Number of jobs in employment cycle, std. dev.	1.399	1.308	1.559	1.263
Average share of matches in a cross section ending in EU-transition	0.232	0.230	0.352	0.392
<b>Cross Section Heterogeneity</b>				
Log firm wage, employment weighted average	5.554	5.577	5.219	5.226
Log firm wage, employment weighted std. dev.	0.168	0.161	0.146	0.139
Log firm wage, newly hired workers average	5.415	5.483	5.136	5.141
Log firm wage, newly hired workers std. dev.	0.256	0.244	0.188	0.193
Firm size in FTE, average	4.995	5.508	7.899	8.288
Fraction of active firms to worker population	0.167	0.158	0.098	0.096
Within-job annual log wage growth, average	0.017	0.007	0.008	0.004
Firm effects from auxiliary log wage regression, average	5.541	5.538	5.197	5.197
Firm effects from auxiliary log wage regression, std. dev.	0.187	0.152	0.163	0.131
Worker effects from auxiliary log wage regression, std. dev.	0.223	0.229	0.197	0.200
Residuals from auxiliary log wage regression, std. dev.	0.147	0.147	0.131	0.128
Mean-min wage ratio	2.096	1.979	1.786	1.723
<b>Labor Market Sorting</b>				
Unemployment duration (in weeks) for workers hired into top ranked firms, average	48.830	95.215	59.155	70.787
Unemployment duration (in weeks) for workers hired into top ranked firms, std. dev.	62.778	81.452	71.624	76.912
Starting wage (in DKK) for workers hired into top ranked firms, average	220.848	238.396	177.824	174.930
Starting wage (in DKK) for workers hired into top ranked firms, std. dev.	70.965	61.145	56.533	42.422
Correlation(unemployment duration, starting wage) for workers hired into top ranked firms	-0.056	-0.275	-0.169	-0.376

Note: Standard errors obtained by block-bootstrap in parentheses.

Figure B.7: Model Fit with Stratification by Education—Quarterly Kaplan-Meier Employment Hazards.

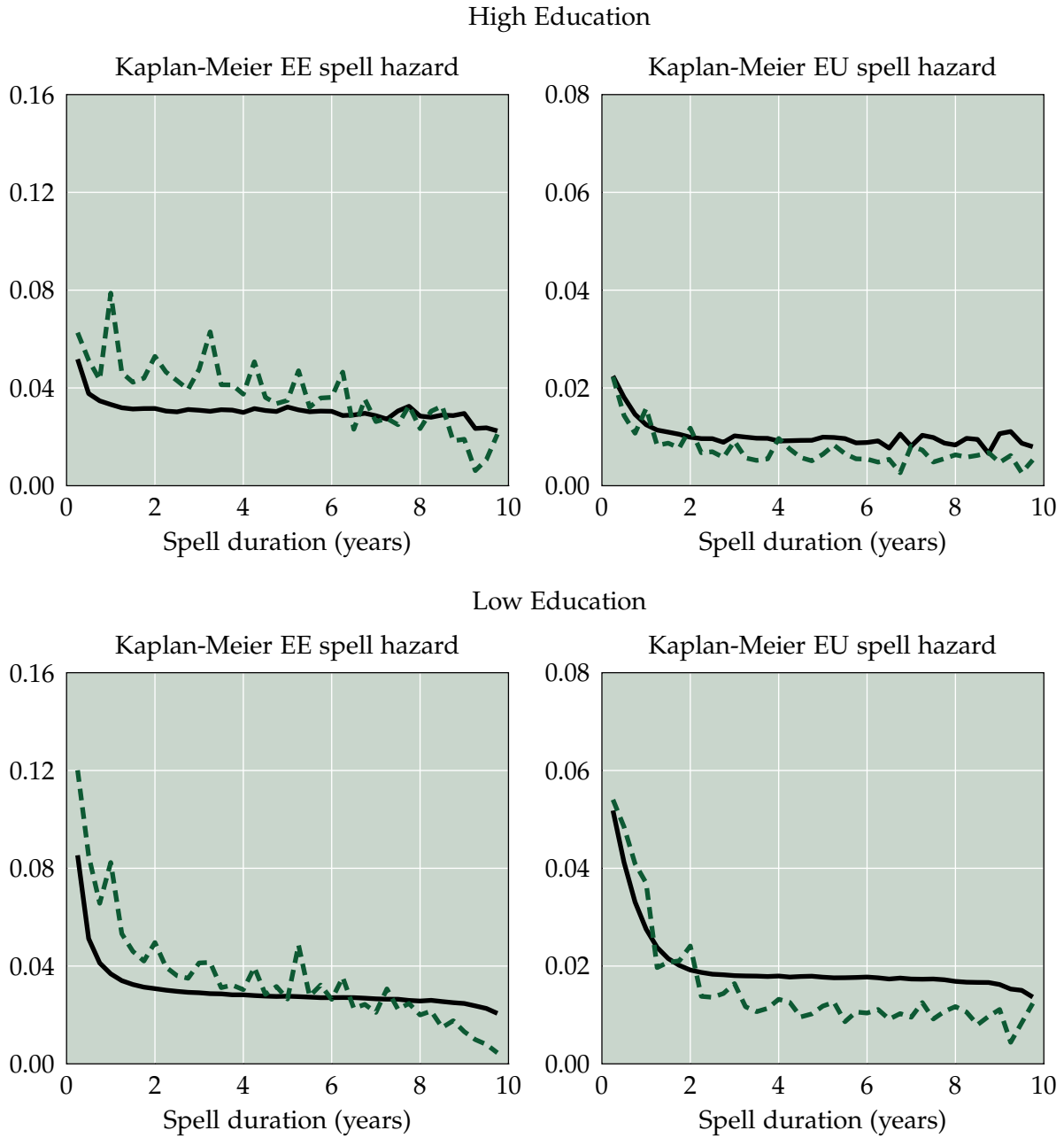
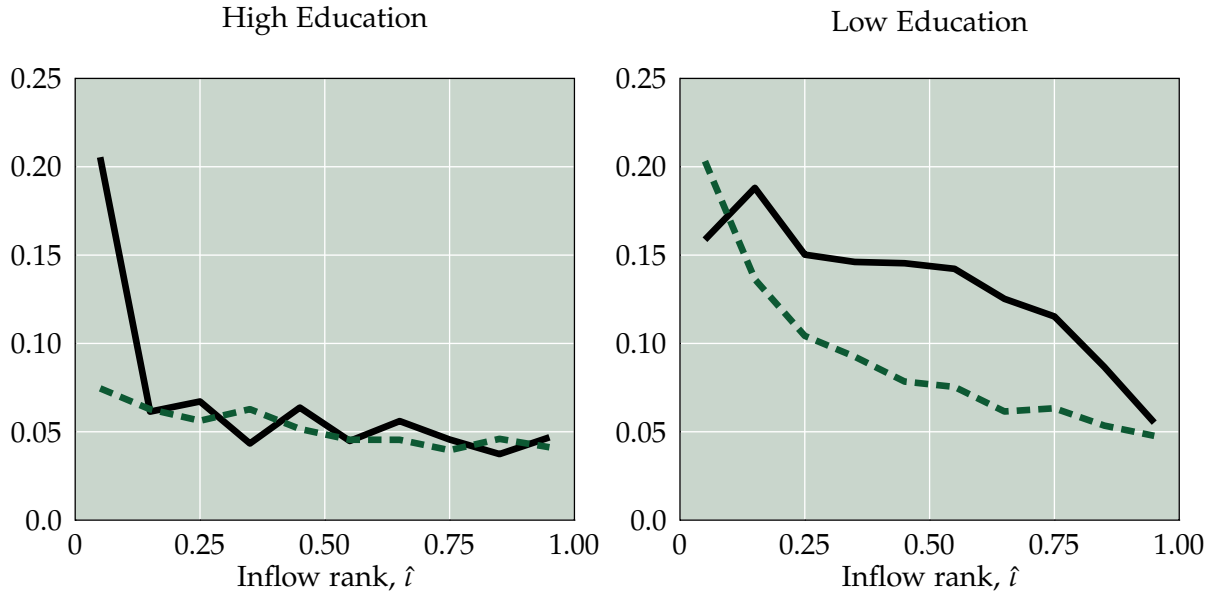


Figure B.8: Model Fit with Stratification by Education—Quarterly Inflow Rank Conditional Job-to-job Transition Hazard Rates



Note: The green dashed line shows job-to-job transition hazard rates estimated on real data (Gaussian non-parametric regression with bandwidth 0.02). The solid line shows simulated job-to-job transition hazard rates for the estimated model equilibrium.

conditional job-to-job transition hazard rates is similar to the fit we obtained when pooling all workers together. The estimated model captures well the hazard rate at the bottom and at the top of the firm ladder, but the model predicts a declining pattern that is convex to the origin, whereas the empirical hazard rates trace out a concave profile. Hence, for low education workers, the model underpredicts the hazard rate out of firms on the middle of the firm ladder.

The coefficients  $\beta_{0k}$  and  $\beta_{1k}$  from the regression (5.3). These coefficients are plotted against firm productivity bin in Figure B.9 for high education workers (top panels) and low education workers (bottom panels). In the main estimation, pooling high and low education workers, we found a decreasing pattern for the constant term  $\beta_0$ , and an increasing-towards-zero pattern for the slope coefficient  $\beta_1$ . As is evident from Figure B.9, the same pattern is found when we stratify the data into high and low education workers, although the declining  $\beta_0$ -profile is less pronounced in the stratified data, especially so in the small strata of high education workers. In this sample the estimated constant term  $\beta_0$  is a bit unstable across the productivity bins, although the overall negative gradient remains visible. The estimated model is able to reproduce some features of the observed patterns for the  $\beta_0$ -coefficients. First, the  $\beta_0$ -profile for high educated



workers declines modestly (as in the data), although the model consistently underpredicts the  $\beta_0$ . For low education workers, the model reproduces well the  $\beta_0$ -coefficient for the least productive firms, but overestimates the decline in the estimated  $\beta_0$ 's as progressively more productive firms are considered.

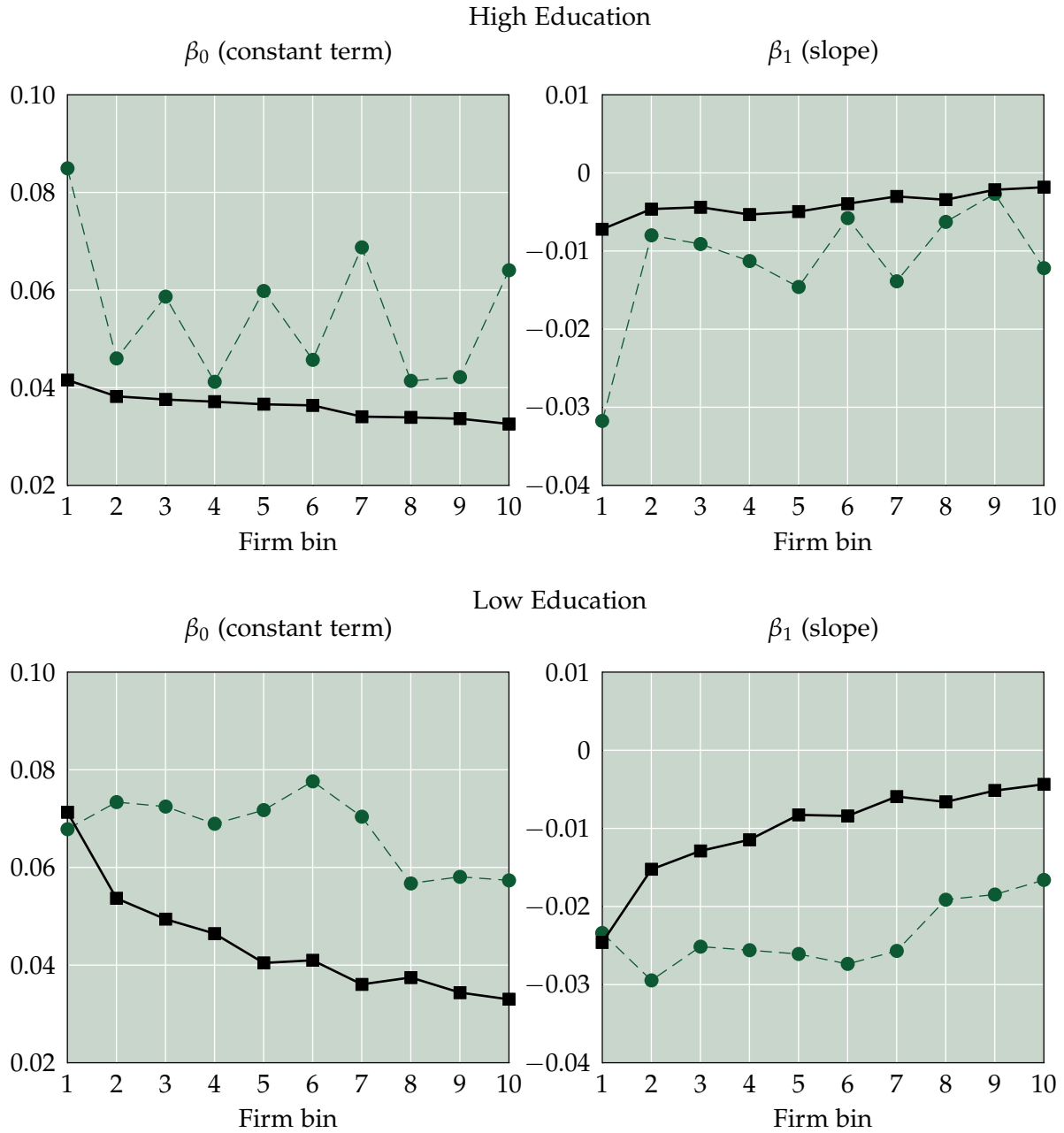
With respect to the slope parameter  $\beta_1$ , measuring the extent to which Kaplan-Meier job-to-job transition hazard functions within productivity bins exhibit negative duration dependence, both high and low educated workers exhibit the an increasing towards zero pattern. For the high education workers, there is a large increase in the estimated  $\beta_1$ -coefficient in the bottom of the productivity ladder, after which the  $\beta_1$ -coefficients increase only modestly, if at all. The estimated model is cannot reproduce the large increase in  $\beta_1$  at the bottom of the firm productivity ladder, but fits reasonable well the modest increase in  $\beta_1$  over the rest of the ladder, albeit with a tendency for slight overestimation. For low education workers, the estimated model hits almost exactly the  $\beta_1$ -coefficient at the least productive firms. However, the model overpredicts the increasing empirical  $\beta_1$ -profile.

#### **B.2.4 Log Wage Variance Decomposition**

We here briefly review the log wage variance decompositions obtained on the stratified samples.

**AKM Regressions** Table 8 presents log wage variance decompositions using the AKM approach on real and simulated data. Looking first at the decomposition for the data, we note that the covariance between worker and firm fixed effects is negative in both strata, as it was in the pooled data. However, the covariance is quantitatively more important in the stratified log wage variance decompositions, in particular among high education workers. Keep in mind that the covariance is likely to be negatively biased in small samples (see e.g. Postel-Vinay and Robin (2006)). Overall, the stratified decompositions are in line with the decomposition obtained on the pooled data. The lion's share of log wage variance results from worker heterogeneity, and this source of dispersion appear to be more important among high education workers than among low education workers. Variance in the estimated firm fixed effects is also important in both groups, as is residual wage dispersion. Comparing the empirical decompositions to their simulated counterparts, we see that the estimated models generate the observed structure of wages reasonably well. In general, the estimated models underpredict the total amount of log wage variance. This underestimation occurs despite the fact that the estimated models are unable to capture the negative covariance between worker and firm fixed effects that appears in the data, and is driven by underestimation of the variance of both worker- and firm fixed effects.

Figure B.9: Model Fit with Stratification by Education: Inflow Rank Conditional Job-to-job Transition Hazard Functions



Note: Left panels: Green dots represent the constant term  $\beta_0$  estimated on real data, black squares represent the constant term  $\beta_0$  estimated on simulated data. Right panels: Green dots represent the slope term  $\beta_1$  estimated on real data, black squares represent the slope term  $\beta_1$  estimated on simulated data.

Table 8: Log Wage Variance Decomposition with Stratification by Education—The AKM approach

	High Education				Low Education			
	Data		Sim.		Data		Sim.	
	Value	Percent of $Var(\ln w_{it})$	Value	Percent of $Var(\ln w_{it})$	Value	Percent of $Var(\ln w_{it})$	Value	Percent of $Var(\ln w_{it})$
$Var(\ln w_{it})$	0.101	100%	0.096	100%	0.082	100%	0.073	100%
$Var(\chi_i)$	0.082	81%	0.064	67%	0.056	67%	0.047	64%
$Var(\varphi_{J(i,t)})$	0.018	18%	0.011	11%	0.015	18%	0.007	10%
$Var(\varepsilon_{it})$	0.016	16%	0.019	20%	0.015	18%	0.014	19%
$2Cov(\chi_i, \varphi_{J(i,t)})$	-0.015	-15%	0.002	2%	-0.005	-5%	0.005	7%

Table 9: Log Wage Variance Decomposition with Stratification by Education—Accounting for Labor Market Sorting

	High Education		Low Education	
	Value	Percent of $Var(\ln w_{it})$	Value	Percent of $Var(\ln w_{it})$
Worker effect	0.039	48%	0.035	60%
Firm effect	0.008	10%	0.005	8%
Friction effect	0.026	32%	0.009	15%
Sorting effect	0.008	10%	0.010	17%
Total predicted variance	0.081	100%	0.058	100%
Total simulated variance	0.087	107%	0.061	105%

**Accounting for Labor Market Sorting** Finally, we consider the stratified versions of our structural log wage variance decompositions which features wage dispersion due to worker effects, firm effects, labor market frictions, and labor market sorting. The decompositions, which we can only carry out on simulated data, are reported in Table 9. Table 9 only reports the variance decomposition of the low layoff type workers, which account for almost all observed wage variation. 99% of the simulated log wage variation is within layoff type workers, and the low layoff type workers account for 99% of employed workers. Comparing the decompositions in Table 9 to that obtained on the pooled data and reported in Table 5, we note that worker effects always come out as the largest contributor to wage dispersion. In the pooled data and among high education workers, labor market frictions are the second highest contributor. Among low education workers, labor market sorting is slightly more important than the labor market frictions in generating wage dispersion. In the pooled data, firm effects come out as the least important source of dispersion, a pattern we also find among low education workers. Among high education workers, firm effects and labor market sorting account for equal shares of log wage variance. Quantitatively, each of the four dimensions of wage dispersion contributes significant amounts of dispersion, with the shares ranging between 48% and 10% for high education workers, and 8% and 60% for low education workers. Overall, the log wage variance decompositions for high and low education workers reported in Table 9 suggests that our findings and conclusions in the main text are robust with respect to stratification on education.

## References

- 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.
- Dey, Matthew S. and Christopher J. Flinn (2005). An equilibrium model of health insurance provision and wage determination. *Econometrica* 73, no. 2: 571–627.
- Postel-Vinay, Fabien and Jean-Marc Robin (2006). Microeconomic search-matching models and matched employer-employee data. *Advances in Econometrics, Eight World Congress*.
- Rubinstein, Ariel (1982). Perfect equilibrium in a bargaining model. *Econometrica* 50, no. 1: 97–109.