Job Sampling and Sorting*

Daniel le Maire[†]
University of Copenhagen

Christian Scheuer
Copenhagen Business School

July 23, 2013

Abstract

We propose a search model with on-the-job search where heterogeneous workers and firms match. Workers can select a sample size of firms to apply to, and since more productive workers sample more firms, assortative matching arises when the production function is strictly supermodular and multiplicatively separable. The model delivers a log-linear wage equation with additively separable worker and firm effects. We simulate the theoretical model and estimate the two-way fixed-effects model as Abowd, Kramarz, and Margolis (1999). By adding a match-effect, we obtain an estimated negative correlation between worker and firm effects, although there is positive assortative matching (positive correlation) in the theoretical model. By applying Woodcock's (2011) mixed effects estimator which also takes the match effect into account, we obtain a positive, albeit slightly negatively biased correlation. Subsequently, both estimators are used on Danish matched employer-employee data and we find evidence for positive assortative matching.

Keywords: Assortative matching, labor market search, linked employer-employee data, person and firm effects.

JEL Classification: C23, C78, J31, J62, J64.

*We are very grateful for helpful comments from Dale Mortensen, Karsten Albæk, Martin Browning, Fane Groes, Bo Honore, John Kennes, Bertel Schjerning, Peter Sørensen, Chris Taber and seminar participants of the LMDG 2008 conference in Sandbjerg. All remaining errors are ours.

[†]Corresponding author: Department of Economics, University of Copenhagen, Øster Farimagsgade 5, building 26, 1353 Copenhagen K, Denmark, daniel.le.maire@econ.ku.dk, Telephone: +4535323063.

1 Introduction

In data from many different countries[1], the commonly-used two-way-fixed effect estimator invented by Abowd, Kramarz, and Margolis (1999) (henceforth, AKM) reveals that high-wage workers are not systematically employed in high-paying firms. This finding is often cited as evidence against positive assortative matching between workers and firms. We present an on-the-job search model in which positive assortative matching obtains, but because of omitted variable bias due to omission of a match effect, it is not reflected in the correlation between earnings components associated with workers and firms. Therefore, this paper offers a theoretical and empirical resolution to a controversial stylized feature of employer-employee matched data.

Frictionless models of assignment along the lines of Becker's (1973) marriage model imply positive assortative matching as long as the production function is strictly supermodular implying complementarity of the production function. Shimer and Smith (2000) show that for assortative matching to arise in a search model, the sufficient condition is that the production function is strictly log supermodular, which is a stronger notion of complementarity. As noted by Atakan (2006), the reason is that higher gains to search for more productive workers are offset by higher costs of rejecting an offer.

Previous studies on assortative matching have examined the sets of matches that are acceptable by both workers and firms and have found perfect segregation implying that the labor market is segmented into multiple non-overlapping markets.[2] However, with on-the-job search, unemployed workers accept any offer greater than or equal to the benefit level and subsequently climb the wage ladder by searching on-the-job. Instead of the matching sets approach, we compare the worker distribution across different firm and match productivities conditional on worker type and prove that the equilibrium distribution of more productive workers stochastically dominates that of less productive workers.

We obtain assortative matching with the strictly supermodular and multiplicatively separable Cobb-Douglas production function by letting workers choose how many job offers they want to sample, as in the seminal contribution by Stigler (1961). Assortative matching arises since more productive workers sample more jobs and, hence, they will, on average, end up in better jobs.

Recent independent papers also propose search models with sorting and reconsider the correlation estimated from the AKM estimation. Similar to us they show that positive assortative matching does not necessarily imply a positive correlation between worker and firm effects.

[1]See Abowd, Creecy, and Kramarz (2002) and Abowd, Kramarz, Lengermann, and Perez-Duarte (2004) for results for both the US and France, Gruetter and Lalive (2004) for results for Austria, Piekkola (2006) for Finland, Andrews, Gill, Schank, and Upward (2008) for results for Germany, and Barth and Dale-Olsen (2003) for results for Norway.

[2]See for example Atakan (2006), Becker (1973), Burdett and Coles (1999), Chade (2001) and Smith (2006).

Eeckhout and Kircher (2011) propose a directed search model which requires less complementarity than the search case, but more than the frictionless case. They find that with use of wage data alone, one cannot distinguish a model that features positive sorting from a model of negative sorting, but that wages can provide information about the strength of sorting. de Melo (2009) develops a search model along the lines of Shimer and Smith (2000). The implication of the models of both Eeckhout and Kircher and de Melo is that high productivity firms have better outside options when matching with low productivity workers and, hence, the bargained wage is non-monotonic in firm productivity.

Lise, Meghir, and Robin (2008) propose an equilibrium search model with on-the-job search where wage contracts are renegotiated upon the arrival of outside offers and productivity shocks building on Postel-Vinay and Robin (2002) and Postel-Vinay and Turon (2010). Using the AKM estimation on simulated data from their search model, the estimated correlation between worker and firm effects is negatively biased. Furthermore, depending on the production function, the estimated correlation can be negative, even though the economy features positive assortative matching.

Our theoretical model is also related to Lentz (2010) and Bagger and Lentz (2012) model of assortative matching in an on-the-job search environment. By allowing that workers select an optimal search intensity rather than the optimal number of jobs to apply for, assortative matching arises with only a strictly supermodular production function. Consequently, our requirements are less general than in the model of Bagger and Lentz, since we need the production function to be strictly supermodular as well as multiplicatively separable.

Compared to the recent papers on sorting we have chosen another wage setting mechanism, piece rate contracts on current output, because this assumption together with the assumptions on the production function imply that our model delivers a log linear wage equation with additively separable worker and firm productivities. In this way we align our theoretical model completely to the original AKM model similarly to Abowd, Kramarz, Lengermann, and Perez-Duarte (2004), such that wages are monotonically increasing in firm productivity.

However, there are no a priory reasons to believe that worker productivity and firm productivity should capture all variation in wages. It is very likely that complementarities between specific workers and firms could exist, or that human capital is accumulated according to the quality of the match between the worker and the firm involved. Both suggest a role for a heterogeneous match specific component. Moreover, several contributions have argued that the quality of the match between workers and firms in itself influences the earnings variation. A prominent example is the search model in Jovanovic (1979), where the flow production is a match quality plus a stochastic term. In principle, a match productivity is not needed in order to have a match component of the wage. The match effect could be due to differences in bargaining strength, for instance, as a result of labor market tightness at the time of contract negotiation.

In the proposed model framework we let the match effect be a random productivity component which enters the production function together with the worker and firm productivity. Our model is, therefore, related to Woodcock (2010), who proposes a search model with match specific learning. However, whereas there is no learning in our model, there are no complementarities in production or on-the-job search in Woodcock's model.

As the log linear wage equation is built into an equilibrium search model with assortative matching, it gives us the opportunity to simulate data from our theoretical model and perform the AKM estimation on this data. When we simulate our model without the match effect and estimate the AKM model and apply the bias correction suggested by Andrews, Gill, Schank, and Upward (2008), the estimated correlation equals the true correlation. In contrast, simulating our model with the match effect and estimating the AKM model, we obtain that even though data is generated from a model with positive assortative matching, the estimated correlation is negative. Furthermore, the share of wages attributed to the worker effect is upward-biased, whereas the share attributed to the firm effect is downward-biased.

The biases which arise are due to omitted variable bias. Worker and firm effects dummies will, by definition, be correlated with match effect dummies, and in presence of a match effect a fixed effects approach ignoring match dummies will lead to biased estimated worker and firm effects. For example, a worker who, on average, is employed in jobs with high match effects will seem to have a higher worker effect than the true. Although a high match effect will also tend to increase the AKM firm effect, the typical firm enters many more matches than the typical worker, so high and low realizations of match effects will tend to be averaged out for firms. The implication is that the match effect variation will upward-bias the variance of the estimated worker effect and that the correlation between the estimated worker and firm effects will be attenuated as a consequence of the spuriously inflated variance of the estimated worker effect. In addition to all this, the theoretical model implies a negative correlation between firm and match productivities, since for the worker to accept a job in a low productivity firm, the match productivity needs to be higher. Since the estimated worker effect is inflated with the match effect variation, and since the firm effect is negatively correlated with the match effect, the correlation between the estimated worker and firm effects may become negative.

Abowd, McKinney, and Schmutte (2010) derive a test for whether worker mobility is affected by match effects. Our simulations show that this test works very well for detecting this type of endogenous job-to-job mobility which can lead to the negative correlation between the estimated worker and firm effects.

Woodcock (2011) suggests an estimation procedure for an empirical model with person, firm and match effects. Using this estimation procedure, Woodcock finds that the estimated negative correlation of the AKM model on US data is, in fact, positive when taking the match heterogeneity into account. We also apply this estimator both to our simulated and empirical data. From the simulations, we learn that the model performs well in attributing the different parts of variation in the dependent variable to the firm, person and match effects, but also this estimator consequently underestimates the true correlation between workers and firms. In our theoretical model framework, workers search for a better joint firm and match productivity.

We apply both the AKM estimation and Woodcock's hybrid mixed effects estimation to a panel of Danish employer-employee data. Using the mixed effects estimation, we find that the match effect has empirical relevance, since it accounts for 14 percent of the variation in log wages. Furthermore, employing a match effects endogenous mobility test by Abowd, McKinney, and Schmutte (2010) clearly shows that worker mobility is in fact based on the match effect similar to the theory model. With the mixed effects model we estimate (what consequently is a downward biased) a correlation of worker and firm effects of 0.16, suggesting that the Danish labor market is characterized by positive assortative matching.

The paper is organized as follows. In section 2, we present our search model where more productive workers sort themselves into more productive matches. In section 3, we briefly consider the AKM model and the recent alternative estimation procedure outlined in Woodcock (2011). In section 4, we simulate our theoretical model and estimate both the AKM and Woodcock models on this model-generated data, while in section 5 we perform estimations on Danish register data. In section 6, we conclude.

2 Theoretical model

Consider an economy with a large number of firms M and a large number of workers N that participate in the labor market. All agents are rational, forward-looking, risk-neutral and infinitely lived. Workers have the opportunity to search both when unemployed and employed.

Heterogeneity exists on both sides of the market as well as in the match between a worker and a firm. Denote the worker productivity by p_w, the firm productivity by p_f and the match effect by p_m. Both the worker and the firm know their own productivity term as well as the decomposition in a given match and, hence, we rule out uncertainty and learning about any of the productivity terms.

We assume that we have a Cobb-Douglas production function $f(p_w, p_f, p_m) = p_w^{\alpha_1} p_f^{\alpha_2} p_m^{\alpha_3}$ which is strictly increasing in p_w, p_f and increasing in p_m. As we want to consider the possibility of the match productivity to play no role, we allow that $\alpha_3 = 0$. The Cobb-Douglas function has two important properties. First, it is strictly supermodular. Second, it is multiplicatively separable.[3] The latter assumption combined with the assumed wage setting mechanism, piece-

[3]We do not need to assume the Cobb-Douglas form, and all implications of the model are true for all functions admitting strictly supermodularity and multiplicative separability.

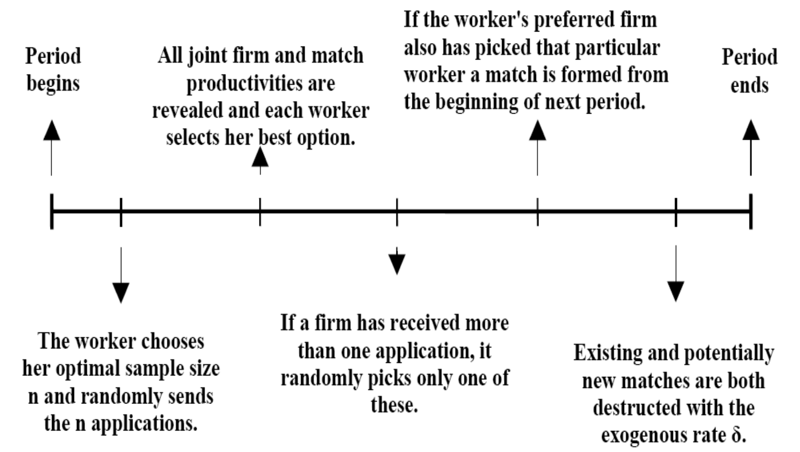
rate contracts, deliver log-additively separable wages.

The distribution of worker productivity p_w is given by the cumulative distribution function $H(p_w)$ with density $h(p_w)$ whereas the distribution of the firm and match productivity $p_{fm} \equiv p_f^{\frac{\alpha_2}{\alpha_2+\alpha_3}} p_m^{\frac{\alpha_3}{\alpha_2+\alpha_3}}$ is given by $\Gamma(p_{fm})$ with corresponding density $\gamma(p_{fm})$.

2.1 Micro foundation of the matching function

The model is set in discrete time and provides the micro foundation of the usual endogenous search intensity models. The timing of events is illustrated in Figure 1.

FIGURE 1: ONE PERIOD SEQUENCE OF EVENTS



The job searcher, whether unemployed or employed, is given the opportunity to apply for n different jobs in each time period. Searching is costly, and the more jobs applied to, the larger costs. We assume that this flow cost function is strictly convex in the number of jobs applied to, $c'(n) > 0$ and $c''(n) > 0$, and that $c(0) = c'(0) = 0$. The more jobs applied to, the more likely the worker is to draw a high joint firm and match productivity p_{fm}. For the worker to get to know the joint firm and match productivity of a given job match, she needs to apply for the job in the particular firm. Hence, she applies to all jobs sampled, although she is only willing to match with the job which turns out to be most productive (if above a reservation productivity level).

Each firm consists of one or more autonomous units. It is possible that a unit gets more than one application in each discrete time interval. Since the time interval is short, each firm unit is only capable of hiring one worker, and one application is chosen randomly among the applicants. The assumption of random selection is made to keep the model as simple as possible and to focus solely on the implications of workers' search decisions. However, one can think of the units' random choice of workers as a framework where the firm is only able to use sequential search in continuous time, and where it is random which application is the first to arrive.[4] The capacity constraint on firm hiring makes jobs scarce similar to the capacity constraint in the de Melo (2009) model. Without this constraint, we would still achieve assortative matching, but at the expense of no coordination frictions and, hence, unemployment would only last for one period.

The implication of this search environment is that the expected number of job applications received is the same for all units, and that the acceptance rate workers face is independent of worker productivity. If worker and firm both choose each other, a match will be formed from the beginning of next period. Finally, in the end of each period a fraction δ of existing and newly formed matches are exogenously destructed.

There is no traditional job arrival rate in this search context, but by abusing the usual notation, let the worker's chance of getting her chosen firm be denoted λ. Letting $\bar{n}$ denote the average number of jobs applied to by the N workers and $\bar{m}$ the average number of units per firm, the share of firm units which gets at least one job application is given by $1-\exp\left(-\frac{N\bar{n}}{M\bar{m}}\right)$. Hence, we can express the probability of getting the chosen firm as

$$\lambda = (1-\delta)\left(1-\exp\left(-\frac{N\bar{n}}{M\bar{m}}\right)\right)\frac{M\bar{m}}{N\bar{n}}$$

where the term $(1-\delta)$ accounts for the fact that a share δ of new matches is destroyed before they come to exist.

2.2 Asset equations

As unemployed, the worker receives benefits which depend on her own productivity $bp_w^{\alpha 1}$. This takes account of the fact that benefits typically depend on previous income as employed. Furthermore, the exact specification used here will simplify the analysis.[5] With the usual notation we denote the value of unemployment $U(\cdot)$ and the value of employment $W(\cdot,\cdot)$. Letting r denote the discount rate, the Bellman equation for an unemployed worker is

$$(1+r)\,U(p_w) = bp_w^{\alpha_1} - \max_n \left\{ c(n) + \lambda \int_{\underline{p}_{fm}}^{\bar{p}_{fm}} \max\left\{W\left(p_w, p'_{fm}\right), U(p_w)\right\} dQ\left(p'_{fm}|n\right) + (1-\lambda)\,U(p_w) \right\} \tag{1}$$

[4] If allowing firms to select their preferred job applicant, it is no longer necessarily the case that all workers will prefer the highest joint firm and match productivity. Rather, less productive workers may only have a very low chance of becoming employed in high productive firms and, thereby, prefer to join low productivity firms. Therefore, this will only tend to increase the degree of assortative matching.

[5] A similar assumption is used in Postel-Vinay and Robin (2002).

where $\bar{p}_{fm}$ and $\underline{p}_{fm}$ denote respectively the upper and lower bounds of the joint firm and match productivity distribution and $Q(p_{fm}|n) = \Gamma(p_{fm})^n$ denotes the distribution of the joint firm and match productivity conditional on the number of sampled jobs, n.

Letting p_{fm}^r be the worker's reservation productivity and using integration by parts we can express this Bellman equation as

$$rU(p_w) = bp_w^{\alpha_1} - \max_n \left\{ c(n) + \lambda \int_{p_{fm}^r}^{\bar{p}_{fm}} W'_{p_{fm}}(p_w, p'_{fm}) \left(1 - \Gamma(p'_{fm})^n\right) dp'_{fm} \right\} \tag{2}$$

With probability δ the job is destructed and the worker returns to unemployment. Let β denote the worker's share of output. The Bellman equation for an employed worker is

$$(r+\delta) W(p_w, p_{fm}) = \beta p_w^{\alpha_1} p_{fm}^{\alpha_2+\alpha_3} - \max_n \left\{ c(n) + \lambda \int_{p_{fm}}^{\bar{p}_{fm}} W'_{p_{fm}}(p_w, p'_{fm}) \left(1 - \Gamma(p'_{fm})^n\right) dp'_{fm} \right\} + \delta U(p_w) \tag{3}$$

Both workers and firms have minimum values of productivities that they are willing to match with. For an employed worker, the reservation productivity is the current firm and match productivity, p_{fm}, whereas for an unemployed worker, the reservation value is the firm and match productivity for which the worker is indifferent between being employed and unemployed. The reservation productivity for an unemployed worker can be derived as

$$\beta p_w^{\alpha_1} \left(p_{fm}^r\right)^{\alpha_2+\alpha_3} - \max_n \left\{ c(n) + \lambda \int_{p_{fm}^r}^{\bar{p}_{fm}} W'_{p_{fm}}(p_w, p'_{fm}) \left(1 - \Gamma(p'_{fm})^n\right) dp'_{fm} \right\}$$
$$= bp_w^{\alpha_1} - \max_n \left\{ c(n) + \lambda \int_{p_{fm}^r}^{\bar{p}_{fm}} W'_{p_{fm}}(p_w, p'_{fm}) \left(1 - \Gamma(p'_{fm})^n\right) dp'_{fm} \right\}$$
$$\Leftrightarrow$$
$$p_{fm}^r = \left(\frac{b}{\beta}\right)^{\frac{1}{\alpha_2+\alpha_3}} \tag{4}$$

Thus, the reservation productivity p_{fm}^r is identical for all unemployed workers[6] such that the reservation wage is increasing in the worker productivity, p_w.

Differentiating the Bellman equation for an unemployed worker and substituting out $W'_{p_f}(p_w, p_f)$ from equation (3) give us the first-order condition for the sample size n for an unemployed worker

[6] We would not have an identical reservation productivity if all workers received the same amount of unemployment benefits. In that case, more productive workers would be more impatient and have a lower reservation productivity than less productive workers, since more productive workers have higher opportunity costs of staying unemployed and, therefore, would be more eager to get a job.This would not have been the case if all workers received the same amount of unemployment benefits. In that case, more productive workers would be more impatient and have a lower reservation productivity than less productive workers, since more productive workers have higher opportunity costs of staying unemployed and, therefore, would be more eager to get a job.A similar result appears in the general model of Burdett and Coles (1999).

$$c'(n) = \lambda \int_{\left(\frac{b}{\beta}\right)^{\frac{1}{\alpha_2+\alpha_3}}}^{\bar{p}_{fm}} \frac{\beta(\alpha_2+\alpha_3) p_w^{\alpha_1} \left(p'_{fm}\right)^{\alpha_2+\alpha_3-1} \left[-\Gamma\left(p'_{fm}\right)^n \ln\left(\Gamma\left(p'_{fm}\right)\right)\right]}{r+\delta+\lambda\left(1-\Gamma\left(p'_{fm}\right)^n\right)} dp'_{fm} \tag{5}$$

Since $[-\Gamma(p_{fm})^n \ln(\Gamma(p_{fm}))] \geq 0$ for all p_{fm}, the right hand side is increasing in p_w for $\alpha_1, (\alpha_2+\alpha_3) > 0$. Since $c''(n) > 0$ and the r.h.s. is decreasing in n for $\Gamma(p_{fm}) \in]0,1[$, more productive workers sample more jobs, that is $n'_{p_w}(p_w, p_{fm}) > 0$. The restriction that $\alpha_1, (\alpha_2+\alpha_3) > 0$ corresponds to the production function having a positive cross derivative, $f''_{p_w,p_{fm}}(p_w, p_{fm}) > 0$, which again exactly is the requirement of complementarity (or supermodularity) in the production function for Becker's (1973) model implying assortative matching.

The number of jobs applied to, n, only takes on integer values, and the n maximizing (5) is most likely not an integer. However, since the l.h.s. is increasing in n, while the r.h.s. is decreasing in n, the optimal integer value of n is one of the two integers adjacent to n, unless the optimal value of (5) itself is an integer.

Employed workers' first-order condition for n is completely analogous to equation (5) and is given by

$$c'(n) = \lambda \int_{p_{fm}}^{\bar{p}_{fm}} \frac{\beta(\alpha_2+\alpha_3) p_w^{\alpha_1} \left(p'_{fm}\right)^{\alpha_2+\alpha_3-1} \left[-\Gamma\left(p'_{fm}\right)^n \ln\left(\Gamma\left(p'_{fm}\right)\right)\right]}{r+\delta+\lambda\left(1-\Gamma\left(p'_{fm}\right)^n\right)} dp'_{fm} \tag{6}$$

Hence, we obtain the following proposition:

Proposition 1 *When the production function $f(p_w, p_{fm})$ admits supermodularity such that $\alpha_1, (\alpha_2+\alpha_3) > 0$, more productive workers sample more jobs conditional on their current joint firm and match productivity, that is $n'_{p_w}(p_w, p_{fm}) > 0$.*

Proof. See equation (5) and (6). ■

Since workers are employed at $p_{fm} \geq \left(\frac{b}{\beta}\right)^{\frac{1}{\alpha_2+\alpha_3}}$, they have smaller expected gains of searching than unemployed workers with the same productivity, and consequently, they search less. Therefore, it is not necessarily the case that more productive employed workers search more than less productive employed workers, since, on average, they will be employed in more productive matches already in the first match after being unemployed. Unemployed workers choose the firm with the highest productivity among the sampled productivities given $p_{fm} \geq p^r_{fm}$, whereas employed workers only accept to join a firm of productivity p_{fm} if their current productivity is lower, and if p_{fm} is the highest productivity sampled.

The firm picks the worker at random, given that the worker's productivity together with the drawn match effect is above a reservation threshold. We will express the reservation threshold in terms of the match effect, but make it dependent on worker productivity, that is $p^r_m(p_w)$. The firm's reservation productivity is decreasing in p_w, such that the highest reservation match effect is at the lower support of the worker distribution. To make the model tractable, we assume that workers are accepted at any firm. More formally,

Assumption 1 $p_m^r\left(\underline{p}_w\right) \leq \underline{p}_m$ for all $p_f \in \left[\underline{p}_f, \bar{p}_f\right]$, where $\underline{p}_m$ is the lowest possible match effect.

2.3 Equilibrium

Since firms choose workers randomly, we can express the inflow to and outflow from unemployment, u, as in usual search models. In steady-state the inflow and outflow must balance

$$\delta(1-u) = \lambda u \tag{7}$$

Letting the optimal number of jobs applied to by a worker with productivity p_w employed in a match with productivity p_{fm} be denoted as $n(p_w, p_{fm})$, the average number of jobs applied to in the economy is given by

$$\bar{n} = u\int_{\underline{p}_w}^{\bar{p}_w} n\left(p'_w, \left(\frac{b}{\beta}\right)^{\frac{1}{\alpha_2+\alpha_3}}\right) h(p'_w)\, dp'_w + (1-u)\int_{\underline{p}_w}^{\bar{p}_w}\int_{\underline{p}_{fm}}^{\bar{p}_{fm}} n\left(p'_w, p'_{fm}\right) d\Gamma(p_{fm})\, dH(p'_w)$$

The mass $G(p_w, p_{fm})$ is the share of employed workers with worker productivity less or equal to p_w working at a joint firm and match productivity less or equal to p_{fm}. The flow into $G(p_w, p_{fm})$ must equal the outflow in steady-state. The outflow, on the l.h.s. below, consists of two terms, the exogenous destruction which happens at rate δ and the endogenous job quits. When considering outflow from $G(p_w, p_{fm})$ in the form of quits, we - by definition - only consider workers with productivity p_w or less, who leave a job with productivity equal to or below p_{fm} and get a job with productivity above p_{fm}. Since workers only change to more productive matches, inflow into $G(p_w, p_{fm})$ only comes from unemployment where the number of jobs sampled is $n\left(p_w, p_{fm}^r\right)$. The steady-state condition is given by

$$\begin{aligned}
&\delta(1-u)\,G(p_w, p_{fm}) + (1-u)\,\lambda\int_{\underline{p}_w}^{p_w}\int_{\underline{p}_{fm}}^{p_{fm}}\left(\Gamma(\bar{p}_{fm})^{n\left(p'_w, p'_{fm}\right)} - \Gamma(p_{fm})^{n\left(p'_w, p'_{fm}\right)}\right) g\left(p'_w, p'_{fm}\right) dp'_{fm}dp'_w \\
= \;& u\lambda\int_{\underline{p}_w}^{p_w}\left(\Gamma(p_{fm})^{n\left(p'_w, p^r_{fm}\right)} - \Gamma\left(\underline{p}_{fm}\right)^{n\left(p'_w, p^r_{fm}\right)}\right) h(p'_w)\, dp'_w
\end{aligned} \tag{8}$$

Rearranging and using the equilibrium equation for unemployment gives us

$$\int_{\underline{p}_w}^{p_w}\int_{\underline{p}_{fm}}^{p_{fm}}\left[\delta + \lambda\left(1 - \Gamma(p_{fm})^{n\left(p'_w, p'_{fm}\right)}\right)\right] g\left(p'_w, p'_{fm}\right) dp'_{fm}dp'_w = \delta\int_{\underline{p}_w}^{p_w}\Gamma(p_{fm})^{n\left(p'_w, p^r_{fm}\right)} h(p'_w)\, dp'_w$$

Differentiating this with respect to p_w gives

$$\int_{\underline{p}_{fm}}^{p_{fm}} \left[\delta + \lambda \left(1 - \Gamma\left(p_{fm}\right)^{n\left(p_w, p'_{fm}\right)}\right)\right] g\left(p_w, p'_{fm}\right) dp'_{fm} = \delta \Gamma\left(p_{fm}\right)^{n\left(p_w, p^r_{fm}\right)} h\left(p_w\right) \quad (9)$$

Evaluating this expression in $\bar{p}_{fm}$ obviously gives $\int_{\underline{p}_{fm}}^{\bar{p}_{fm}} g\left(p_w, p'_{fm}\right) dp'_{fm} = h\left(p_w\right)$. Combining this with equation (9), we obtain

$$\int_{\underline{p}_{fm}}^{p_{fm}} \tilde{g}\left(p_w, p'_{fm}\right) dp'_{fm} = \frac{\delta}{\delta + \lambda} \Gamma\left(p_{fm}\right)^{n\left(p_w, p^r_{fm}\right)} + \frac{\lambda}{\delta + \lambda} \int_{\underline{p}_{fm}}^{p_{fm}} \Gamma\left(p_{fm}\right)^{n\left(p_w, p'_{fm}\right)} \tilde{g}\left(p_w, p'_{fm}\right) dp'_{fm} \quad (10)$$

where $\int_{\underline{p}_{fm}}^{p_{fm}} \tilde{g}\left(p_w, p'_{fm}\right) dp'_{fm} \equiv \int_{\underline{p}_{fm}}^{p_{fm}} g\left(p_w, p'_{fm}\right) dp_{fm} \left(\int_{\underline{p}_{fm}}^{\bar{p}_{fm}} g\left(p_w, p'_{fm}\right) dp'_{fm}\right)^{-1}$ is the cumulative distribution function of firm productivities conditional on worker skills.

We want to examine whether the allocation of workers in the economy reflects positive sorting. While previous studies on assortative matching have examined the sets of matches that are acceptable by both workers and firms, this approach is not useful with on-the-job search. Instead, we compare worker distributions over different firm and match productivities conditional on worker type similarly to Lentz (2010). For this purpose, we use the following definition:

Definition 1 *Consider two workers A and B, where $p_w^A > p_w^B$. Positive assortative matching implies that*

$$G^B\left(p_{fm}\right) \equiv \int_{\underline{p}_{fm}}^{p_{fm}} \tilde{g}\left(p_w^B, p'_{fm}\right) dp'_{fm} \geq \int_{\underline{p}_{fm}}^{p_{fm}} \tilde{g}\left(p_w^A, p'_{fm}\right) dp'_{fm} \equiv G^A\left(p_{fm}\right)$$

for $p_{fm} \in \left]\underline{p}_{fm}; \bar{p}_{fm}\right[$.

In the present model set-up, it is not immediately clear on the grounds of equation (10) whether or not the model framework implies assortative matching. Before we more formally prove this, we can gain some intuition by assuming that the number of jobs sampled is independent of the current firm and match productivity, that is $n\left(p_w, p_{fm}\right) = n\left(p_w\right)$. With this assumption, it is clear that we have

$$\int_{\underline{p}_{fm}}^{p_{fm}} \tilde{g}\left(p_w, p_{fm} | n\left(p_w\right) = n\left(p_w, p_{fm}\right)\right) dp'_{fm} = \frac{\delta \Gamma\left(p_{fm}\right)^{n\left(p_w\right)}}{\delta + \lambda\left(1 - \Gamma\left(p_{fm}\right)^{n\left(p_w\right)}\right)} \quad (11)$$

which for a constant number of jobs sampled across worker type, that is $n\left(p_w\right) = 1$, gives us the usual distribution of realized matches (see e.g. Postel-Vinay and Robin (2002)). If equation (11) is differentiated with respect to p_w, we see that the r.h.s. becomes negative implying assortative matching. However, assuming that $n\left(p_w, p_{fm}\right) = n\left(p_w\right)$ will overvalue the degree of assortative matching since it does not take into account that workers in better matches search less.

A formal proposition for assortative matching is given below. The proof uses a discretized version of equation (10) and is in Appendix A. This proposition only encompass $n = 0, 1, 2$, since allowing for higher n makes the math very cumbersome.

Proposition 2 *When the production function $f(p_w, p_{fm})$ admits supermodularity and multiplicatively separability such that $\alpha_1, (\alpha_2 + \alpha_3) > 0$, and workers are allowed to sample $n = 0, 1, 2$ jobs, the model features assortative matching.*

Proof. See Appendix A. ■

Furthermore, the steady-state equilibrium is a tuple $\left\{p^r_{fm}, n(p_w, p_{fm}), n\left(p_w, p^r_{fm}\right), G(p_w, p_{fm}), u\right\}$ satisfying equations (4), (5), (6), (7) and (8), which leads us to the following proposition.

Proposition 3 *There exists a unique steady-state.*

Proof. See Appendix B. ■

3 Econometric methodology

3.1 The two-way fixed effects model

The assumptions of a multiplicatively separable production function and piece rate contracts imply that our theoretical model gives a log-linear wage equation, where the log of worker productivity and the log of firm productivity are additively separable. This property is very convenient, since it aligns perfectly with the leading empirical model for estimating the degree of assortative matching in the labor market developed by Abowd, Kramarz, and Margolis (1999) (henceforth AKM). Before we exploit this direct link between the theoretical and empirical model, we will give a brief introduction to the AKM model, which takes the following form

$$y_{it} = x_{it}\beta + \theta_i + \psi_{J(i,t)} + \varepsilon_{it} \qquad i = 1, ..., N; \quad t = 1, .., T_i \tag{12}$$

where y_{it} is the log of the hourly wage rate for individual i in period t, x_{it} is a $1 \times k$ vector of time-varying explanatory variables that may both relate to the individual and the firm, β is the parameter vector, θ_i is the unobserved person effect, $\psi_{J(i,t)}$ is the unobserved firm effect, and ε_{it} is the error term with $E\left(\varepsilon_{it}|x_i, \theta_i, \psi_{J(i,.)}\right) = 0$. The function $J(i,t)$ associates an employer, indexed by J, with an individual i at time t.

Since most empirical applications relate to data with more individuals than firms ($N > J$), we begin by making a within-individual transformation that sweeps away θ_i. Expressing the model in matrix form, using $\sim$ to denote within transformed variables and letting D be the matrix of firm dummies gives us the following model to estimate

$$\tilde{Y} = \tilde{X}\beta + \tilde{D}\Psi + \tilde{\varepsilon} \quad (13)$$

where Ψ is a vector of firm effects for every firm in the sample.[7] The parameter estimates can be found by solving

$$\begin{bmatrix} \hat{\beta} \\ \hat{\Psi} \end{bmatrix} = \begin{bmatrix} \tilde{X}'\tilde{X} & \tilde{X}'\tilde{D} \\ \tilde{D}'\tilde{X} & \tilde{D}'\tilde{D} \end{bmatrix}^{-1} \begin{bmatrix} \tilde{X}'\tilde{Y} \\ \tilde{D}'\tilde{Y} \end{bmatrix} \quad (14)$$

The problem of estimating the resulting model is that the cross-product matrix is potentially very high-dimensional due to $\tilde{D}'\tilde{D}$ containing a dummy for each firm. However, with use of sparse matrix algebra we can estimate the $\left[\hat{\beta}, \hat{\Psi}\right]$ and afterwards recover $\hat{\theta}$.

One key aspect of this estimator worth noticing is that the firm coefficients are only identified by workers moving between firms in the sample period. Hence, looking only at a worker employed in the same firm in all years, there is no way for the econometrician to disentangle the person and firm fixed element of log wages.

In fixed effects models, the variances of the fixed effects have a positive bias. Furthermore, Postel-Vinay and Robin (2006) note that due to the additive structure, the covariance of the estimated worker and firm effects will be biased since an over-estimate of the one fixed effect will lead to an under-estimate of the other. Andrews, Gill, Schank, and Upward (2008) develop the formula to correct for these biases. We use their method of correcting the estimates under the assumption that the explanatory variables are uncorrelated with the worker and firm effects.[8]

3.2 Woodcock's hybrid mixed effects estimation

Woodcock (2011) argues that the presence of a match effect will bias the correlation between worker and firm effects unless the match effect is completely orthogonal. In other words, if mobility depends on the match effect, the estimates and the resulting correlation will be biased. Estimating an AKM model which also includes a match effect as a fixed effect is impossible and, therefore, Woodcock suggests using the mixed effects model. Instead of estimating all individual effects, Woodcock's estimation procedure consists of estimating the variances of the worker, firm, match effects and error term and subsequently predicting all individual effects. Woodcock needs to assume that the three random effects are uncorrelated with each other,

[7]Identification of firm effects is in principle only possible within a group, where a group is defined by the movement of workers between firms. For a thorough discussion see Abowd, Creecy, and Kramarz (2002). For expositional simplicity we assume that we already have identified the groups, dropped one firm dummy for each group, and normalized the mean in each group to zero to allow for cross-group comparison. All while redefined $\tilde{D}$ accordingly.

[8]This assumption is made since without it we need to invert a $N \times N$ matrix to solve for the biases, and this is not feasible with our current computational power. Note also that unlike in the empirical estimation we have no explanatory variables when we simulate data from our theoretical model. Hence, in that case the assumption is met by definition.

when estimating the variances. However, there is no such restriction on the predicted individual effects. Woodcock's model is

$$y_{it} = x_{it}\beta + \theta_i + \psi_j + \phi_{ij} + \varepsilon_{it} \qquad i = 1, ..., N;\ t = 1, .., T_i;\ j = 1, ..., J$$

where ϕ_{ij} is the match effect. It is easy to estimate β even if the worker effect θ_i, the firm effect ψ_j, and the match effect ϕ_{ij} are fixed effects, since β is the within-match estimator. However, separately identifying the worker, firm, and match effects in a fixed effects context is impossible since we from, say, M matches cannot estimate $M + N + J$ effects.

Woodcock suggests the following 3-step estimation procedure: First, estimate $\hat{\beta}$ as the within-match estimator in the first stage and compute $\left(y_{it} - x_{it}\hat{\beta}\right)$. In the second stage, the variance of the random effects $\left(\sigma_\theta^2, \sigma_\psi^2, \sigma_\phi^2\right)$ and the error variance σ_ε^2 are estimated by Restricted Maximum Likelihood (REML) on $\left(y_{it} - x_{it}\hat{\beta}\right)$ under the assumption that the effects are uncorrelated. These REML estimates are computed with the use of the average information algorithm of Gilmour, Thompson, and Cullis (1995), which exploits the sparsity of the matrix design. In the third and final stage, Woodcock makes the Best Linear Unbiased Predictor (BLUP) of the random effects, which are not subject to the uncorrelatedness, and estimate the correlations between the various terms.

4 Simulation study

4.1 Solving and simulating the search model

Since our theoretical model delivers a log-linear wage equation, we can solve and simulate data from the theoretical model and estimate the AKM model as well as Woodcock's mixed effect model on this simulated data.

To simulate our model, we need a number of functional form assumptions. In the following we assume that worker productivity, p_w, firm productivity, p_f, and match productivity, p_m, all are log-normal distributed; $p_w \sim LN\left(\mu_w, \sigma_w^2\right)$, $p_f \sim LN\left(\mu_f, \sigma_f^2\right)$, and $p_m \sim LN\left(\mu_m, \sigma_m^2\right)$. Besides this, we specify the search cost function as $c(n) = c_1 n^{c_2}$.

Before simulating data from our model, we approximate the function $n\left(p_w, p_{fm}\right)$. To do this, we solve equations (5) and (6) for n using quadrature methods to approximate the integrals. The function $n\left(p_w, p_{fm}\right)$ is step-wise increasing in p_w and decreasing in p_{fm}, as illustrated in Figure 2, so we just need to know precisely the values of $\left(p_w, p_{fm}\right)$ in which n changes in equilibrium. We use an algorithm that determines the number of points, where n changes in the $\left(p_w, p_{fm}\right)$ space with a precision, such that the maximum deviation in p_{fm} and p_w from their true values is 0.01.[9] Next, we use cubic splines to approximate a curve for each change in

[9] First, we set $p_{fm} = 0$ and determine for which values of p_w n changes. Then for $\underline{p}_w$, each of these points

n in the (p_w, p_{fm}) space. These cubic splines are then used to draw $n(p_w, p_{fm})$ in the actual simulation.

Notice that since the production function is given by $f(p_w, p_{fm}) = p_w^{\alpha_1} p_f^{\alpha_2} p_m^{\alpha_3}$, our assumption of log-normality in all inputs implies that the joint firm and match productivity $p_{fm} = p_f^{\frac{\alpha_2}{\alpha_2+\alpha_3}} p_m^{\frac{\alpha_3}{\alpha_2+\alpha_3}}$ is also log-normal distributed $p_{fm} \sim LN\left(\mu_{fm}, \sigma_{fm}^2\right)$ with $\mu_{fm} = \frac{\alpha_2}{\alpha_2+\alpha_3}\mu_f + \frac{\alpha_3}{\alpha_2+\alpha_3}\mu_m$ and $\sigma_{fm} = \sqrt{(\frac{\alpha_2}{\alpha_2+\alpha_3}\sigma_f)^2 + (\frac{\alpha_3}{\alpha_2+\alpha_3}\sigma_m)^2}$. This again implies that when we simulate from a model without a match effect, we can simply let $\alpha_3 = 0$, which implies that $\mu_{fm} = \mu_f$ and $\sigma_{fm} = \sigma_f$ such that $p_{fm} = p_f$ and $f(p_w, p_{fm}) = f(p_w, p_f)$.

To facilitate comparison between simulations with and without match effects, we specify μ_m and σ_m, such that $\mu_{fm} = \mu_w$ and $\sigma_{fm} = \sigma_w$ for simulations with match effects, and such that $\mu_f = \mu_w$ and $\sigma_f = \sigma_w$ for simulations without match effects.

We wish to simulate an economy inhabited with firms whose size is log-normal distributed. Given μ_f and σ_f, we define the minimum and maximum p_f as the values corresponding to the 0.05 percentile and the 99.5 percentile. In between the minimum and maximum values, we let each firm productivity be equally spaced, whereas each firm's share of the job offers is the log-normal density. Hence, the firm productivity is only approximately log-normal, but will, as the number of firms increases, converge to the log-normal distribution.

We also need to choose values for our parameters. The parameter for the job destruction rate δ is 7 percent, whereas the rate at which the worker gets a job λ in equilibrium becomes 63 percent. These values give us an equilibrium unemployment of 10 percent. We have set the parameters of the log normal distributions such that $\alpha_3 = 0.25$, implying that the match effect constitutes roughly 25 percent of the explained variation in log wages. Obviously, the influence of the match effect that we find below will be smaller (larger), if the match effect constitutes a smaller (larger) share of the wage than the assumed 25 percent. The rest of the parameters in the two simulations are given in Table 1.

Table 1: Parameter values

Match effect included:												
α_1	α_2	α_3	$\mu_w = \mu_{fm}$	$\sigma_w = \sigma_{fm}$	μ_f	σ_f	μ_m	σ_m	β	c_1	c_2	r
0.40	0.35	0.25	5.00	0.60	5.00	0.84	5.00	0.84	0.50	5.50	1.20	0.05
Match effect not included:												
α_1	α_2	α_3	$\mu_w = \mu_{fm}$	$\sigma_w = \sigma_{fm}$	μ_f	σ_f	μ_m	σ_m	β	c_1	c_2	r
0.40	0.60	0.00	5.00	0.60	5.00	0.60	.	.	0.50	5.50	1.20	0.05

in p_w, $\bar{p}_w$, and two additional values evenly distributed among each of these points we determine all values of p_{fm} where n changes.

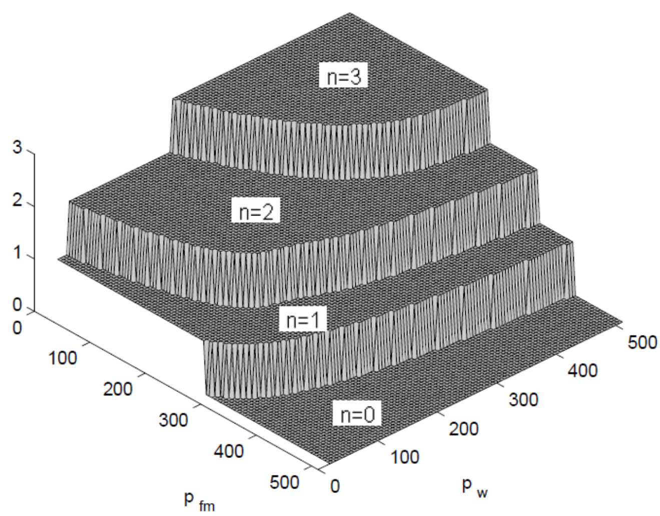
To complete the link between the theoretical and empirical models, we add a normal distributed error term to the log linear wage equation. We set the variance of the error term, such that it contributes with $5 - 10$ percent of the total wage variation.

Given our initial guess of a correlation between worker and the joint firm and match effect, the model runs for 35 periods before the worker allocation is completely stable and, hence, we discard the first 34 simulated periods.

With the assumed parameter values, the maximum number of jobs that any worker samples is 3. In Figure 2, a surface plot of the n-function is shown. Unemployed workers with productivity above approximately 203, corresponding to 30 percent of the workers, sample 3 jobs and almost all of the remaining 70 percent of unemployed workers sample 2 jobs. The figure also shows that as workers climb the firm and match productivity ladder, they reduce the number of jobs they sample and when they reach sufficiently far, they stop searching.

Since the workers in the economy without a match effect and in the economy with a match effect draw productivities from the same distribution, the n-function in the two scenarios is identical.

FIGURE 2: THE OPTIMAL JOB SAMPLE SIZE



In Table 2, we show the results from the AKM estimation on the simulated data. The left panel has results from the simulation study without a match effect ($\alpha_3 = 0$), the middle panel with a match effect ($\alpha_3 = 0.25$), while the right panel is an intermediate case. With the latter case, we want to study the implication of having a match effect in the wages which workers do not take into account when searching and accepting jobs. Hence, this match effect can be seen as a compensating wage differential.[10] However, at the same time, we want this simulated economy to be as close to the two other cases. Therefore, we assume the existence of another match effect which only affects the job mobility pattern, but does not enter wages. In other words, the two match effects in the intermediate case are uncorrelated, and the one only affects job mobility, whereas the other only affects wages.

[10] Alternatively, a time-constant match effect, which does not affect job mobility, arises in directed search models where wages are determined by auctions, cf. Kennes and le Maire (2013).

Table 2: Monte-Carlo estimations

	Match effect not included			Match effect included			Intermediate case (two separate match effects)		
No. of Monte Carlo replications		100			100			100	
No. time periods		6			6			6	
No. of observations		67,311			67,459			67,446	
No. of persons		12,500			12,500			12,500	
No. of firms		469			494			494	
Average of wages		124.2			129.6			107.7	
Variance of wages		40.7			46.6			43.7	
Average no. of obs per firm		143			137			137	
Proportion of variance explained (averages):	True	AKM	Mixed	True	AKM	Mixed	True	AKM	Mixed
Firm effect	0.383	0.384	0.392	0.232	0.189	0.198	0.256	0.261	0.272
Person effect	0.554	0.565	0.544	0.493	0.705	0.479	0.382	0.622	0.363
Match effect	0.000	.	0.000	0.203	.	0.252	0.277	.	0.280
Error term	0.063	0.051	0.063	0.072	0.106	0.072	0.085	0.118	0.085
Theoretical correlations:	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean
True corr$(\alpha_1 \ln(p_w), \alpha_2 \ln(p_f))$	0.118	0.158	0.135	0.063	0.103	0.084	0.064	0.101	0.085
True corr$(\alpha_1 \ln(p_w), \alpha_3 \ln(p_m))$	0.000	0.000	0.000	0.057	0.099	0.078	-0.020	0.021	-0.001
True corr$(\alpha_2 \ln(p_f), \alpha_3 \ln(p_m))$	0.000	0.000	0.000	-0.317	-0.281	-0.296	-0.021	0.019	-0.001
Estimated correlations:	Min	Max	Mean	Min	Max	Mean	Min	Max	Mean
AKM, estimated corr$(\hat{\theta}_i, \hat{\psi}_{J(i,t)})$	0.110	0.149	0.128	-0.068	-0.014	-0.038	-0.028	0.021	-0.002
AKM, corrected corr$(\hat{\theta}_i, \hat{\psi}_{J(i,t)})$	0.117	0.157	0.135	-0.056	0.000	-0.025	-0.013	0.038	0.013
Mixed effects corr$(\hat{\theta}_i, \hat{\psi}_{J(i,t)})$	0.106	0.142	0.123	0.025	0.056	0.042	0.044	0.074	0.060
Endogenous mobility test:									
Average χ^2 statistic		48.61			229.96			51.55	
Average p-value		0.594			0.000			0.511	
Number significant at 5% level		1			100			5	
Number significant at 1% level		0			100			0	

For each of the three cases, we have simulated 100 data sets, on which we perform the AKM estimation and Woodcock's hybrid mixed effects estimation. There are only small differences in the overall labor market between the three cases. The minor differences in the mean and variance of wages arise, since, whereas we assume a discrete distribution of firm productivities which is cut off at the 0.05 and 99.5 percentiles, the match effect is allowed to be continuous. Therefore, the joint firm and match distribution will have a larger dispersion than the firm productivity in the case without match effect. Furthermore, in the middle panel workers can get a better job within a firm by drawing an offer from the exact same firm, but with a higher match effect. This might also have a small positive effect on both the mean wage and the dispersion of wages.

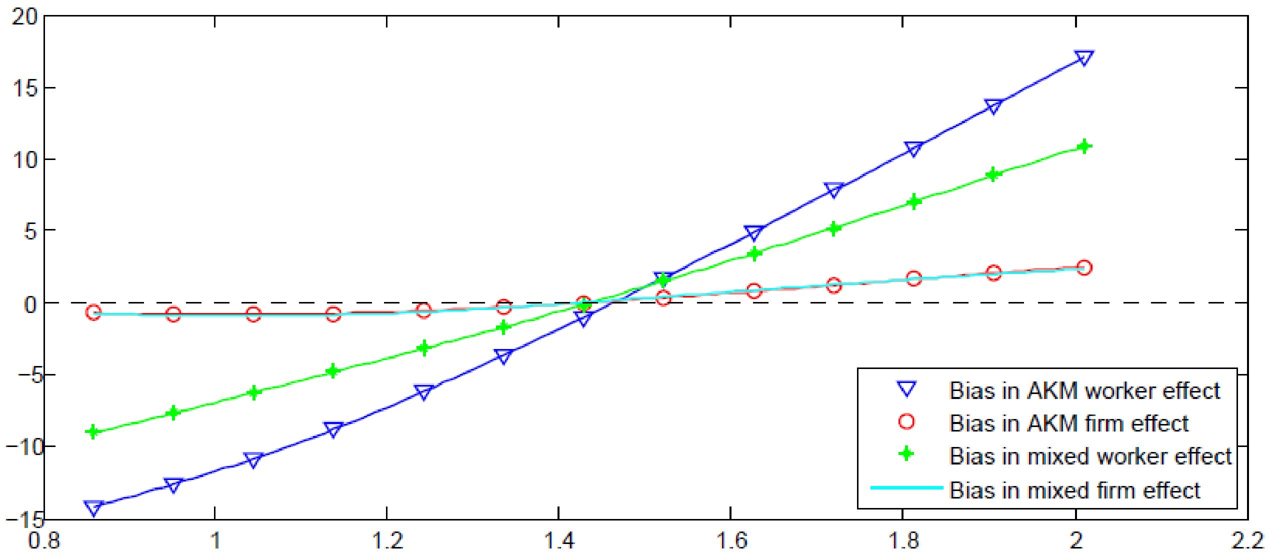
4.2 Variance decomposition on worker, firm and match effects

In absence of a match effect in the data generating process, variances of fixed effects are positively biased. The left panel of Table 2 confirms this, as the proportions of the variance of log wages attributed to estimated worker and firm effects are slightly larger than the true shares.

In presence of an omitted match effect, Woodcock (2011) shows that the sign of the bias of the variances is theoretically indeterminate. In our case, however, the middle panel of Table 2 shows that the variance of the AKM worker fixed effects is positively biased, whereas the variance of the AKM firm fixed effects is negatively biased. The AKM estimation attributes 71 percent of the variance of log wages to the worker effect, even though the true worker effect only accounts for 49 percent. At the same time, the firm effect only explains 19 percent of the variation according to the AKM estimation compared to a true contribution of 23 percent.

These biases arise, since by definition worker and firm dummies are correlated with the omitted match effects dummies. Furthermore, it is primarily the variance of the estimated worker effect, which is biased, because the worker only matches with few firms, unlike the typical firm which employs several workers. Therefore, firm averages of the true match productivities will only display small differences compared to worker averages of true match effects. Hence, it will mainly be the variance of the worker effect, which is inflated by the match effects variation. To see this, Figure 3 shows a kernel regression of the percentage estimation error for worker and firm effects as a function of the true match effect. Whereas the percentage estimation error for firm effects, on average, is close to zero over all support of the true match effect, there is a sizeable error for the estimated worker effects.

Figure 3: Percentage estimation error and the match effect



As all three productivities increase the wage, we would ceteris paribus have expected that also the variance of firm effects would have a positive bias. In fact, Figure 3 seems to suggest that this is the case. However, this is not the case due to the theoretical model framework. Since workers seek to maximize the joint firm and match productivity, they will tend to only choose a job with low firm (match) productivity in combination with high match (firm) productivity, such that the theoretical firm and match productivities become negatively correlated. Indeed, the average correlation between firm and match productivities in the theoretical model is -0.296 in the match effects case in the middle panel of Table 2. In contrast to this, this correlation is zero in the intermediate case with the two uncorrelated match effects (right panel). The implication is that both worker and firm effects have positively biased variances, but again that mainly the variance of the former is inflated by the match effect.

In absence of match effects, the estimated worker and firm shares are closer to the true for the AKM model than Woodcock's hybrid mixed effects model, but the differences are minor. In presence of a match effect in the data generating process, Figure 3 shows that the mixed effects estimator - as is the case for the AKM estimator - cannot completely distinguish between worker and match effects. For low (high) match effects, the AKM worker effect is downward (upward) biased up to roughly 15 percent, whereas the mixed worker effect is downward (upward) biased up to roughly 10 percent. However, Table 2 shows that overall the mixed effects estimator performs well in attributing the correct proportions to worker, firm and match effects.

4.3 The correlation between worker and firm effects

For the case without match effects (left panel), the estimated correlation between worker and firm effects from the AKM estimation is very close to the true correlation of, on average, 0.135. With the bias correction by Andrews, Gill, Schank, and Upward (2008), the estimated correlation equals the true. For both cases with a match component in the wages, the estimated and corrected correlations between AKM worker and firm effects are negatively biased. For the right panel, the explanation is that the estimated worker effect is inflated by the match effect, which is uncorrelated with the firm effect and, thereby, the correlation between estimated worker and firm effects is biased towards zero. The fact that the estimated correlation, on average, is slightly negative (-0.002) follows from the fact that the two fixed effects (and the error term) must add up, so underestimating the one implies overestimating the other. When correcting for this bias following Andrews, Gill, Schank, and Upward (2008), the average correlation is positive (0.013).

When not only there is a match component of the wages, but workers also change jobs based on this component, we see that the estimated and corrected correlations become even negative: The true correlation is 0.084, whereas the estimated and corrected correlations are -0.038 and -0.025. The intuition is that the spurious variation in the worker effect due to the omitted match productivity is negatively correlated with the firm productivity. If the correlation between match and firm productivities is sufficiently large and negative, it may dominate the positive correlation between worker and firm effects, as long as the variance of the estimated worker effect is inflated by the match effect. Therefore, existing studies finding a negative correlation between worker and firm effects may actually reflect a labor market with positive assortative matching.

The correlation between worker and firm effects from Woodcock's hybrid mixed effects estimation is positive for each of the simulations for the three cases considered. There is a negative bias in the mixed effects correlation for each of the three cases considered, but importantly the estimated correlation is less affected by presence of match effects. Only in the case of no match effect, the AKM correlation is closer to the true correlation.

4.4 Endogenous mobility

When match effects influence job mobility decisions, we have found that the corrected correlation between AKM worker and firm effects can be negative, even though the true correlation is positive. Abowd, McKinney, and Schmutte (2010) derive a test to assess exactly, whether the match effect plays a role for the job-to-job mobility. The test is based on match effects estimated as the average AKM residuals for each worker-firm combination and determines, whether this estimated match effect is correlated with the firm productivity of a subsequent job match.

The test is implemented as an independence test, and Abowd, McKinney, and Schmutte (2010) suggest to divide worker effects, firm effects, subsequent firm effects and match effects from the AKM residuals into deciles. Due to the small samples used in each simulation, we only use 3 different cells for each variable.[11] In total, this gives 81 different cells. In Table 2, we report the p-values for the χ^2 test. The test works extremely well for detecting whether workers change jobs based on match effects: In all 100 simulations where the data generating process has a match effect that affects the mobility pattern, the test strongly rejects the null of independence. At the same time, only in one simulation the test rejects the independence, when there is no match effect in the data generating process. Finally, for the case where the match effect plays a role of a compensating wage differential, the test rejects in 5 percent of the times at the 5 percent significance level.

4.5 Sample length

Andrews, Gill, Schank, and Upward (2008) argue that the negative correlation between worker and firm effects estimated using real data to a large extent is due to limited mobility. Although omitted match effect bias and limited mobility bias are two distinct sources of bias, they might both vanish, if we consider firms with more worker mobility or have a longer panel. To examine this, we extend the sample length.

Table 3 shows that more estimation periods imply that the gap between the true and the corrected estimate narrows. The intuition is that as we observe the worker in more years, especially the worker effect becomes more precisely estimated. Hence, the variance of the estimated worker effect is to a lesser extent inflated with the variation of the match effect, which is negatively correlated with the firm effect. With 10 periods, the estimated correlation from the AKM estimation turns positive, but even with 50 periods the bias still persists. Furthermore, we see that the correlation from the mixed effects estimation in all simulations is in between the true correlation and the correlation from the AKM estimation.

[11] The samples used in the test are particularly small as we only consider job changes for individuals with a two-year spell at the time of the job change. We restrict the sample this way to avoid that the average of the residuals depends on the duration of the spell which in turn may depend on worker and firm fixed effects.

Table 3: Estimated correlation and the sample length

No.of time periods	No. of obs.	Ave. obs. per firm	True correlation	AKM corrected	Mixed effects
6	67,459	137	0.084	−0.025	0.042
10	112,438	227	0.085	0.005	0.042
20	224,944	453	0.084	0.031	0.050
30	337,433	679	0.085	0.046	0.059
50	562,395	1,129	0.084	0.058	0.067

5 Empirical analysis

We have access to a register data set for the entire Danish population above 15 years. The data set is an unbalanced panel data set for 1999-2008, and the variables originate from two databases maintained by Statistics Denmark and include detailed information on a wide range of variables from administrative registers.

First, the IDA database (the integrated database for labor market research) contains detailed information on individual socioeconomic characteristics on an annual basis.[12] The dependent variable, the log of hourly wages, originates from the IDA database. The hourly wage rate is calculated as the sum of total labor income divided by the total number of hours worked in any given year. The measure for total labor income is highly reliable, since it is third-party reported to the tax authorities. The annual hours are imputed from the knowledge of each individual's payment to the mandatory pension fund, ATP, as this payment is grouped into four different intervals according to the hours worked. This is a rather crude measure of hours worked, but in what follows we look only at full time workers, which helps alleviate measurement errors arising due to the grouped nature of the hours variable. As with all nominal variables used in the analysis, the hourly wage rate is deflated with the GDP deflator.

Due to the nature of the fixed effect estimator, which precludes the use of time-invariant explanatory variables, the only individual worker variables included are a second order polynomial in the actual labor market experience and dummies for whether the individual lives in a rural area or in a large city.

The FIDA (the firm integrated database for labor market research) includes variables for individual firms. Each individual and firm has a unique identifier and all employed individuals are linked to a firm identifier at the last week of November. The firm data set contains annual information on value-added and the capital stock, but these two variables are in many cases imputed by Statistics Denmark, as only a rotating sample of Danish firms are required to

[12]For more details on the IDA dataset see for example Abowd and Kramarz (1999).

provide accounting data. Firms are sampled according to their size. For example, each year 10 percent of firms with $5-9$ employees are sampled, whereas all firms with 50 employees or more are sampled each year.

The IDA dataset has 45.625.345 person-year observations. Restricting attention to persons aged $25-60$ years leaves us with $26,551,453$ observations. Next, only including persons working full time in the last week of November with a positive wage leaves us with $17,664,461$ observations. Discarding firms with zero and one employees, public sector firms, and missing values of value-added and capital stock leaves us with $8,768,538$ observations. This data set consists of $1,545,892$ persons employed by $172,946$ firms. In the first column of Table 4, descriptive statistics are provided for this sample.

In order to arrive at the main estimation sample, we discard all firms with less than 40 observations in total over the entire 10-year panel. The resulting data set has $7,434,187$ observations on $1,354,460$ persons employed in $29,348$ firms as shown in the second column of Table 4. Finally, we also construct a sample where we discard observations with imputed values for the firms' value-added and capital stock. This deletes roughly 40 percent of the observations, but due to the rotating sampling scheme the share of firms and workers deleted are 4 and 17 percent, respectively. The third column of Table 4 contains descriptive statistics for this sample. In general, the sample means for the three samples are very similar.

Table 4: Sample characteristics

	All observations	Firms with more than 40 obs.	Firms with more than 40 obs. and no imputed values
No. of observations	8,768,538	7,434,187	4,463,519
No. of persons	1,545,892	1,354,460	1,124,499
No. of firms	172,946	29,348	28,179
No. of worker-firm matches	2,913,098	2,312,212	1,613,417
Sample means:			
Log hourly wage rate	5.361	5.383	5.372
	(0.362)	(0.353)	(0.352)
Experience	17.849	18.131	17.813
	(9.367)	(9.386)	(9.185)
Woman	0.317	0.316	0.313
	(0.465)	(0.465)	(0.464)
Copenhagen	0.09	0.088	0.087
	(0.286)	(0.283)	(0.282)
Large city	0.131	0.133	0.133
	(0.338)	(0.34)	(0.34)
Rural	0.668	0.668	0.667
	(0.471)	(0.471)	(0.471)
Value-added per worker (million DKK)	0.524	0.540	0.514
	(1.684)	(1.694)	(1.009)
Capital per worker (billion DKK)	0.002	0.002	0.002
	(0.029)	(0.028)	(0.035)
Log of firm size (no. of workers)	4.629	5.157	5.149
	(2.239)	(1.984)	(1.99)

Table 5: Parameter estimates

	Firms with more than 40 obs.		Firms with more than 40 obs. and no imputed values	
	AKM	Mixed effects	AKM	Mixed effects
Experience	0.0205	0.0181	0.0186	0.0099
	(0.00011)	(0.00016)	(0.00013)	(0.00018)
$(\text{Experience})^2$	-0.0053	-0.0050	-0.0053	-0.0050
	(0.00001)	(0.00002)	(0.00002)	(0.00002)
Large cities	-0.0130	0.0001	-0.0114	-0.0001
	(0.00081)	(0.00095)	(0.00094)	(0.00109)
Rural area	0.0035	0.0108	0.0053	0.0111
	(0.00057)	(0.00065)	(0.00066)	(00075)
Value added per employee (mill. DKK)	0.0063	0.0063	0.0053	0.0053
	(0.00014)	(0.00013)	(0.00015)	(0.00014)
Capital stock per employee (bill. DKK)	0.0077	0.0070	0.0034	0.0022
	(0.00246)	(0.00224)	(0.00255)	(0.00261)
Male ratio	-0.0106	-0.0141	-0.0018	-0.0126
	(0.00136)	(0.00131)	(0.00220)	(0.00216)
Log of firm size	0.0232	0.0292	0.0249	0.0294
	(0.00024)	(0.00025)	(0.00032)	(0.00032)

Note: The omitted category is the greater Copenhagen area. Standard errors are in parenthesis.

Table 5 presents the parameter estimates from the AKM and Woodcock's hybrid mixed effects estimation. The left panel shows estimates using the sample with all observations for firms with more than 40 observations, whereas the sample used for the right panel discards all observations with imputed firm values. The four different combinations of estimators and samples give very similar estimates. Most estimated coefficients are highly significant and have expected signs. The estimated returns to experience are lower when taking account of match effects. Woodcock (2011) and Sørensen and Vejlin (2010) find a similar pattern in data for the US and Denmark, respectively, and document that part of the individual's wage growth is achieved by job-to-job changes to higher job matches. This would seem to support that workers do, in fact, take match effects into account when searching, as we will return to below. Both the value added and capital stock are found to be significant, when we consider the sample with partly imputed values. However, when estimating on the sample with no imputed values, the capital stock becomes insignificant. Since our main focus is assortative matching, we proceed to look at the estimated correlation between worker and firm fixed effects.

Table 6: AKM estimator, standard errors and correlation matrix (obs per firm >40)

	Standard deviation	Correlations								
		Y	$X_w\hat{\beta}_w$	$X_f\hat{\beta}_f$	$\hat{T}_{dum}$	$\hat{\psi}_{J(i,t)}$	$\hat{\theta}_i$	$X_f\hat{\beta}_f+$ $\hat{\psi}_{J(i,t)}$	$X_w\hat{\beta}_w+$ $\hat{\theta}_i$	$\hat{\varepsilon_{it}}$
Y	0.353	1.000	0.060	0.072	0.050	0.299	0.873	0.325	0.882	0.374
$X_w\hat{\beta}_w$	0.048	0.060	1.000	−0.036	−0.039	0.009	−0.081	−0.009	0.076	−0.000
$X_f\hat{\beta}_f$	0.048	0.072	−0.036	1.000	0.016	−0.193	−0.011	0.314	−0.017	−0.000
$\hat{T}_{dum}$	0.032	0.050	−0.039	0.016	1.000	−0.013	−0.039	−0.005	−0.045	0.000
$\hat{\psi}_{J(i,t)}$	0.093 (0.089)	0.299	0.009	−0.193	−0.013	1.000	0.071 (0.093)	0.871	0.073	−0.000
$\hat{\theta}_i$	0.307 (0.299)	0.873	−0.081	−0.011	−0.039	0.071 (0.093)	1.000	0.063	0.988	0.000
$X_f\hat{\beta}_f + \hat{\psi}_{J(i,t)}$	0.096	0.325	−0.009	0.314	−0.005	0.871	0.063	1.000	0.062 (0.082)	−0.000
$X_w\hat{\beta}_w + \hat{\theta}_i$	0.307	0.882	0.076	−0.017	−0.045	0.073	0.988	0.062 (0.082)	1.000	−0.000
$\hat{\varepsilon_{it}}$	0.132	0.374	−0.000	−0.000	0.000	−0.000	0.000	−0.000	−0.000	1.000

Note: Lower case parenthesis denotes the correlation corrected for statistical bias. X_w includes experience, (experience)^2, large cities and rural areas. X_f includes value added per employee, capital stock per employee, male ratio and ln(employees).

In Table 6, we have shown the correlation matrix from the AKM model used on the sample with partly imputed firm values. In contrast to most empirical studies, we find a positive correlation between the estimated worker and firm components. The estimated correlation between unobserved worker and firm effects is 0.071, while the bias corrected correlation, shown in parenthesis, is 0.093. Taking account of both observed and unobserved effects, we estimate the correlation between the overall worker effect and the overall firm effect to be 0.062, while the corrected is 0.082.[13]

[13] To compute the corrected correlation, we simply subtract the bias in $cov\left(\hat{\theta}_i, \hat{\psi}_{J(i,t)}\right)$ from $cov\left(X_w\hat{\beta}_w + \hat{\theta}_i, X_f\hat{\beta}_f + \hat{\psi}_{J(i,t)}\right)$, and subtract the biases in $var\left(\hat{\theta}_i\right)$ and $var\left(\hat{\psi}_{J(i,t)}\right)$ from respectively $var\left(X_w\hat{\beta}_w + \hat{\theta}_i\right)$ and $var\left(X_f\hat{\beta}_f + \hat{\psi}_{J(i,t)}\right)$.

Table 7: Mixed effects estimator, standard errors and correlation matrix (obs per firm >40)

	Standard deviation	Correlations									
		Y	$X_w\hat{\beta}_w$	$X_f\hat{\beta}_f$	$\hat{T}_{dum}$	$\hat{\psi}_{J(i,t)}$	$\hat{\theta}_i$	$\hat{\phi}_{ij}$	$X_f\hat{\beta}_f+$ $\hat{\psi}_{J(i,t)}$	$X_w\hat{\beta}_w+$ $\hat{\theta}_i$	$\hat{\varepsilon_{it}}$
Y	0.353	1.000	−0.084	0.070	0.043	0.455	0.803	0.617	0.480	0.827	0.408
$X_w\hat{\beta}_w$	0.088	−0.084	1.000	−0.013	−0.065	−0.043	−0.355	−0.205	−0.048	0.004	−0.031
$X_f\hat{\beta}_f$	0.060	0.070	−0.013	1.000	0.012	−0.193	−0.024	−0.028	0.245	−0.031	−0.002
$\hat{T}_{dum}$	0.056	0.043	−0.065	0.012	1.000	−0.036	−0.093	−0.075	−0.030	−0.124	−0.019
$\hat{\psi}_{J(i,t)}$	0.135	0.455	−0.043	−0.193	−0.036	1.000	0.163	0.032	0.904	0.157	0.004
$\hat{\theta}_i$	0.245	0.803	−0.355	−0.024	−0.093	0.163	1.000	0.539	0.150	0.933	0.087
$\hat{\phi}_{ij}$	0.082	0.617	−0.205	−0.028	−0.075	0.032	0.539	1.000	0.019	0.498	0.207
$X_f\hat{\beta}_f + \hat{\psi}_{J(i,t)}$	0.137	0.480	−0.048	0.245	−0.030	0.904	0.150	0.019	1.000	0.142	0.003
$X_w\hat{\beta}_w + \hat{\theta}_i$	0.229	0.827	0.004	−0.031	−0.124	0.157	0.933	0.498	0.142	1.000	0.081
$\hat{\varepsilon_{it}}$	0.109	0.408	−0.031	−0.002	−0.019	0.004	0.087	0.207	0.003	0.081	1.000

Note: Lower case parenthesis denotes the correlation corrected for statistical bias. X_w includes experience, (experience)^2, large cities and rural areas. X_f includes value added per employee, capital stock per employee, male ratio and log of firm size.

Table 7 presents the correlation matrix when estimating using Woodcock's hybrid mixed effects model. We obtain a correlation of 0.119 on the sample, where the minimum number of observations per firm is 50. In Table 8, we compute the mean influence of each term on the variance of the log of the wage rate. While the AKM estimation suggests that roughly 76 percent of this variance can be attributed to the individual heterogeneity, the mixed effects estimation only suggests that about 56 percent is due to worker heterogeneity. This repeats the finding from the simulated data and is due to the high positive correlation of 0.54 between the match effect and the worker effect. From the mixed effects estimation, we find that the match effect explains 14 percent of the variation in the log wages, which is of the same magnitude as Woodcock finds using US data. Sørensen and Vejlin (2010) also decompose the wage dispersion in worker, firm and match effects using Danish register data. They make use of a longer sample for 1980-2003, but at the expense of using work places rather than firms. With a longer panel much more of the variation is attributed to experience accumulation and much less to the worker (fixed and random) effects. The match effect explains a slightly lower part of the variation, but it seems to be due to their dense sampling approach.

Table 8: Variance decomposition		
	AKM	Mixed effects
$X_w\hat{\beta}_w$	0.008	(0.021)
$X_f\hat{\beta}_f$	0.010	0.012
$\hat{T}_{dum}$	0.005	0.007
$\hat{\psi}_{J(i,t)}$	0.078	0.174
$\hat{\theta}_i$	0.759	0.558
$\hat{\phi}_{ij}$	.	0.144
$\hat{\varepsilon_{it}}$	0.140	0.126
Match effect test:		
χ^2 statistic	672.667	
P-value	0.0000	

Note: The match effects test is computed for workers changing jobs from a job with a duration of two years. The worker effects, firm effects before the job change, the firm effects after the job change and match effects from the AKM residuals before the job change were divided into quintiles.

In Table 8, we decompose the variance of log wages in different estimated components for the sample with all observations for firms with more than 40 observations. This is done for both the AKM estimation and the hybrid mixed effects estimation. In both estimations, around $86-87$ percent of the variation is explained, and only a few percent is explained by the explanatory variables. However, the decomposition varies considerably between the AKM estimation and the mixed effects estimation.

Four key features from the simulation study are also found in Table 8. First, a much larger part of the variation is explained by the worker effects, 76 vs. 56 percent, in the AKM estimation compared with the mixed effects estimation. Second, a much lower share of the variance is due to firm effects in the AKM estimation, 8 vs. 17 percent. Third, the match effect contributes to a non-negligible part of the variation. The mixed effects estimation attributes 14 percent of the variation to the match effect. Fourth, we strongly reject exogenous mobility using the match effects test of Abowd, McKinney, and Schmutte (2010), so workers' job-to-job mobility is partly based on match effect.

In Table 9, we examine the effect of shortening the sample in order to mimic the simulation exercise in Table 3. We see that it is mainly the AKM correlation between worker and firm effects that increases when increasing the length of the sample. In fact, the estimated AKM correlation is negative, -0.012, whereas the corrected correlation is 0.023. This relationship is completely in accordance with the simulation study in Table 3. However, it seems to be the case that the hybrid mixed effects correlation is more sensitive to the sample length compared to the results in Table 3.

Table 9: Estimated correlation and the sample length

	$T = 10$ $(1999 - 2008)$	$T = 6$ $(1999 - 2004)$
No. of observations	7,434,187	4,463,519
No. of persons	1,354,460	1,124,499
No. of firms	29,348	28,179
No. of matches	2,312,212	1,613,417
$Corr\left(\hat{\psi}_{J(i,t)}, \hat{\theta}_i\right)$:		
AKM, estimated	0.071	-0.012
AKM, corrected	0.093	0.023
Mixed effects	0.163	0.123

6 Conclusion

In this paper, we argue that presence of a match effect in wages implies that the estimated correlation between worker and firm effects in an AKM estimation will be negatively biased. In order to analyze this, we develop a theoretical search model with continuous heterogeneity on both worker and firm sides. Positive assortative matching is achieved in an on-the-job search model with only a strictly supermodular production function by letting the workers choose how many jobs they want to sample. The theoretical model is completely aligned with the AKM model, as it delivers a log-linear wage equation, which is additively separable in the worker effects and the firm effects. Besides this, the theoretical model also provides the opportunity of having an additively separable match effect in the wage equation.

The match effect bias is a special case of the usual omitted variable bias, since by defintion worker and firm dummies are correlated with the omitted match dummies. As we observe the typical worker in fewer matches than the typical firm, it is primarily the worker effects which are biased by the omitted match effect. The implication is that especially the worker effects are inflated by match effect variation, and the bias of the correlation between the estimated AKM worker and firm effects is, therefore, affected by the correlation between the theoretical firm and match effects. If this theoretical correlation is zero, the estimated correlation between the worker and firm effects will have an attenuation bias. However, it is likely that workers take the match effect into consideration when searching and accepting new job opportunities. In this case, a usual reservation wage strategy will imply a negative correlation between firm and match effects, which in turn can create a negative correlation between the estimated AKM worker and firm effects.

Woodcock's hybrid mixed effects estimation works almost as well as the AKM model in absence of a match effect in the datagenerating process, and in presence of a match effect, the variance of estimated worker and firm effects and the correlation between the two are much closer to the true values.

The empirical application using Danish register data gives results which are in line with the simulation results. According to the mixed effects estimation, 14 percent of the variation in log wages can be attributed to match effects, and the endogenous mobility match effect test developed by Abowd, McKinney, and Schmutte (2010) confirms that workers change jobs also on the basis of the match effect. In addition to this, in the empirical application it seems to be the case that the variance of the estimated worker effect from the AKM estimation is upward biased. The worker effect accounts for as much as 76 percent of variation in log wages in the AKM estimation compared to 56 percent in the mixed effects estimation. The simulation study shows a very similar picture with the mixed effects share being closest to the true value. Finally, using the mixed effects model the correlation between the worker and firm effects is estimated to be 0.16 compared to 0.09 in the AKM estimation. As our simulations show, the estimated correlation using Woodcock's hybrid mixed effect estimation is still a downward biased estimate, so we should expect the true correlation to be above 0.16.

Appendix

Appendix A: Proof of Proposition 2

Proof. The number of jobs sampled is necessarily discrete and recognizing this, it is useful to work with segments of p_{fm} for which a given worker with productivity p_w does not change her number of jobs sampled $n(p_w, p_{fm})$. Defining $\hat{g}\left(p_w, p^j_{fm}\right) \equiv \int_{p^{j-1}_{fm}}^{p^j_{fm}} \tilde{g}\left(p_w, p'_{fm}\right) dp'_{fm}$ and $p^0_{fm} \equiv p^r_{fm}$ we can express equation (10) evaluated in p^i_{fm} as

$$\sum_{j=1}^{i} \hat{g}\left(p_w, p^j_{fm}\right) = \frac{\delta}{\delta+\lambda}\Gamma\left(p^i_{fm}\right)^{n\left(p_w, p^0_{fm}\right)} + \frac{\lambda}{\delta+\lambda}\sum_{j=1}^{i}\Gamma\left(p^i_{fm}\right)^{n\left(p_w, p^{j-1}_{fm}\right)}\hat{g}\left(p_w, p^j_{fm}\right) \tag{15}$$

Rearranging to solve for $\hat{g}\left(p_w, p^i_f\right)$ we get

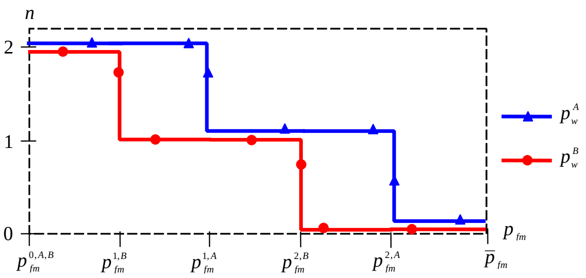
$$\hat{g}\left(p_w, p^i_{fm}\right) = \frac{\frac{\delta}{\delta+\lambda}\Gamma\left(p^i_{fm}\right)^{n\left(p_w, p^0_{fm}\right)}}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^i_{fm}\right)^{n\left(p_w, p^{i-1}_{fm}\right)}} - \{i>1\}\sum_{j=1}^{i-1}\frac{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^i_{fm}\right)^{n\left(p_w, p^{j-1}_{fm}\right)}}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^i_{fm}\right)^{n\left(p_w, p^{i-1}_{fm}\right)}}\hat{g}\left(p_w, p^j_{fm}\right) \tag{16}$$

Next, for some workers of productivity p_w^X we have that

$$\hat{G}^X\left(p_{fm}^I\right) = \sum_{i=1}^{I} \frac{\frac{\delta}{\delta+\lambda}\Gamma\left(p_{fm}^I\right)^{n\left(p_w^X, p_{fm}^{0,X}\right)}}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^I\right)^{n\left(p_w^X, p_{fm}^{I-1,X}\right)}} - \{I>1\}\sum_{i=2}^{I}\sum_{j=1}^{i-1}\left(\frac{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{i,X}\right)^{n\left(p_w^X, p_{fm}^{j-1,X}\right)}}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{i,X}\right)^{n\left(p_w^X, p_{fm}^{i-1,X}\right)}} \cdot \hat{g}\left(p_w^X, p_{fm}^{j,X}\right)\right) \tag{17}$$

Consider two workers A and B, where $p_w^A > p_w^B$, so that we by proposition 1 have that $n\left(p_w^A, p_{fm}\right) \geq n\left(p_w^B, p_{fm}\right)$. Furthermore, assume that $n = 0, 1, 2$. Using Definition 1, we want to compare the allocation of the two worker types to show that $G^B\left(p_{fm}^I\right) \geq G^A\left(p_{fm}^I\right)$. The proof is in two parts. In part A, we show that for the particular example depicted in Figure 6, $G^B\left(p_{fm}^I\right) \geq G^A\left(p_{fm}^I\right)$ for any $p_{fm}^I \in \left[\underline{p}_{fm}, \bar{p}_{fm}\right]$. In part B of the proof, we show that all other possible scenarios are special cases of the example in part A.

FIGURE 6



Part A: The discreteness of the number of jobs sampled implies that n in the interval $[0, 2]$ is a decreasing step-wise function of p_{fm} with five segments. As illustrated in Figure 6, the example considered implies that p_w^A and p_w^B are sufficiently close so that workers A and B are sampling the same number of jobs for some ranges of p_{fm}. Next, we will use that if $n\left(p_w^A, p_{fm}\right) = n\left(p_w^B, p_{fm}\right)$ it must be the case that person B will decrease her n at some firm productivity p_{fm} lower than the firm productivity where A will decrease her n.

By the use of equation (16) and (17), we have that $\hat{G}^B\left(p_{fm}^I\right) - \hat{G}^A\left(p_{fm}^I\right) \geq 0$ no matter which segment p_{fm}^I is situated on in Figure 6. The solutions for each segment are:

$$\left(\hat{G}^B\left(p_{fm}^I\right) - \hat{G}^A\left(p_{fm}^I\right)\right)|_{0 \leq p_{fm}^I < p_{fm}^{1,B}} = 0 \tag{18}$$

$$\left(\hat{G}^B\left(p_{fm}^I\right) - \hat{G}^A\left(p_{fm}^I\right)\right)|_{p_{fm}^{1,B} \leq p_{fm}^I < p_{fm}^{1,A}} =$$

$$\frac{\lambda\delta}{(\delta+\lambda)^2}\Gamma\left(p_{fm}^I\right)\left(\frac{\Gamma\left(p_{fm}^I\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^I\right)^2} - \frac{\Gamma\left(p_{fm}^{1,B}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^{1,B}\right)^2}\right)\frac{1-\Gamma\left(p_{fm}^I\right)}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p_{fm}^I\right)} > 0 \tag{19}$$

$$\left(\hat{G}^B\left(p^I_{fm}\right)-\hat{G}^A\left(p^I_{fm}\right)\right)_{p^{1,A}_{fm}\leq p^I_{fm}<p^{2,B}_{fm}}=$$

$$\frac{\lambda\delta}{(\delta+\lambda)^2}\Gamma\left(p^I_{fm}\right)\left(\frac{\Gamma\left(p^{1,A}_{fm}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^{1,A}_{fm}\right)^2}-\frac{\left(p^{1,B}_{fm}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^{1,B}_{fm}\right)^2}\right)\frac{1-\Gamma\left(p^I_{fm}\right)}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^I_{fm}\right)}>0 \tag{20}$$

$$\left(\hat{G}^B\left(p^I_{fm}\right)-\hat{G}^A\left(p^I_{fm}\right)\right)|_{p^{2,B}_{fm}\leq p^I_{fm}<p^{2,A}_{fm}}=$$

$$\frac{\delta\lambda}{(\delta+\lambda)^2}\frac{1-\Gamma\left(p^I_{fm}\right)}{1-\frac{\lambda}{\delta+\lambda}}\frac{1}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^{1,A}_{fm}\right)^2}\left(\frac{\Gamma\left(p^I_{fm}\right)^2-\Gamma\left(p^{1,A}_{fm}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^I_{fm}\right)}-\frac{\Gamma\left(p^{2,B}_{fm}\right)^2-\Gamma\left(p^{1,A}_{fm}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^{2,B}_{fm}\right)}\right)$$
$$+\frac{\delta\lambda}{(\delta+\lambda)^2}\frac{1-\Gamma\left(p^I_{fm}\right)}{1-\frac{\lambda}{\delta+\lambda}}\left(\frac{\Gamma\left(p^{1,A}_{fm}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^{1,A}_{fm}\right)^2}-\frac{\Gamma\left(p^{1,B}_{fm}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^{1,B}_{fm}\right)^2}\right)\left(\frac{1-\Gamma\left(p^I_{fm}\right)^2}{1-\Gamma\left(p^I_{fm}\right)}-\frac{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^{2,B}_{fm}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^{2,B}_{fm}\right)}\right)$$
$$>0 \tag{21}$$

$$\left(\hat{G}^B\left(p^I_{fm}\right)-\hat{G}^A\left(p^I_{fm}\right)\right)|_{p^{2,A}_{fm}\leq p^I_{fm}<1}=$$

$$\frac{\delta\lambda}{(\delta+\lambda)^2}\frac{1-\Gamma\left(p^I_{fm}\right)}{1-\frac{\lambda}{\delta+\lambda}}\cdot\frac{1}{1-\frac{\lambda}{\delta+\lambda}\left(p^{1,A}_{fm}\right)^2}\left(\frac{\Gamma\left(p^{2,A}_{fm}\right)^2-\Gamma\left(p^{1,A}_{fm}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^{2,A}_{fm}\right)}-\frac{\Gamma\left(p^{2,B}_{fm}\right)^2-\Gamma\left(p^{1,A}_{fm}\right)^2}{\left(1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^{2,B}_{fm}\right)\right)}\right)$$
$$+\frac{\delta\lambda}{(\delta+\lambda)^2}\frac{1-\Gamma\left(p^I_{fm}\right)}{1-\frac{\lambda}{\delta+\lambda}}\left(\frac{\Gamma\left(p^{2,A}_{fm}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^{2,A}_{fm}\right)^2}-\frac{\Gamma\left(p^{1,B}_{fm}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^{1,B}_{fm}\right)^2}\right)\left(\frac{1-\Gamma\left(p^I_{fm}\right)^2}{1-\Gamma\left(p^I_{fm}\right)}-\frac{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^{2,B}_{fm}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^{2,B}_{fm}\right)}\right)$$
$$>0 \tag{22}$$

where the only term not immediately seen to be positive is

$$\left(\frac{1-\Gamma\left(p^I_{fm}\right)^2}{1-\Gamma\left(p^I_{fm}\right)}-\frac{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^{2,B}_{fm}\right)^2}{1-\frac{\lambda}{\delta+\lambda}\Gamma\left(p^{2,B}_{fm}\right)}\right)$$

However, since $\frac{\partial\frac{1-x^2}{1-x}}{\partial x}=1>0$ this term is also clearly positive. Hence, we conclude that

$$\hat{G}^B\left(p^I_{fm}\right)\geq\hat{G}^A\left(p^I_{fm}\right)\quad\forall p^I_{fm}\in\left[\underline{p}_{fm},\bar{p}_{fm}\right]$$

.

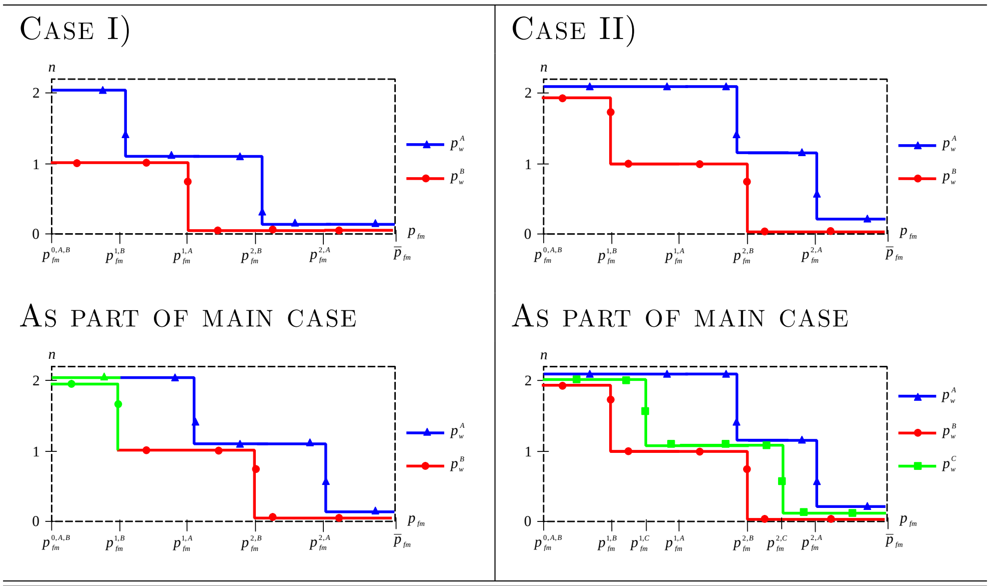
Part B: Since we know that

$$p^A_w>p^B_w,\ n'_{p_w}\left(p_w,p_{fm}\right)>0\quad\Longrightarrow\quad n\left(p^A_w,p_{fm}\right)\geq n\left(p^B_w,p_{fm}\right) \tag{23}$$

it is straightforward to show that every possible scenario satisfying (23) can be included either directly as part of the main case above or be re-stated as combinations of a number of sub-cases, all part of the main case above.

To illustrate this, consider the two examples in Figure 7.

Figure 7



Case I) just parallels the last four segments of the main case from part A since from (18) $\hat{G}^B\left(p^I_{fm}\right) = \hat{G}^A\left(p^I_{fm}\right)$ for $0 \leq p^I_{fm} < p^1_{fm}$. In case II) one can include a worker C with $p^A_w > p^C_w > p^B_w$ (as illustrated with the middle line). By the proof in part A we know that $\hat{G}^B\left(p^I_{fm}\right) \geq \hat{G}^C\left(p^I_{fm}\right)$ and $\hat{G}^C\left(p^I_{fm}\right) \geq \hat{G}^A\left(p^I_{fm}\right)$, whereby we also have that $\hat{G}^B\left(p^I_{fm}\right) \geq \hat{G}^A\left(p^I_{fm}\right)$. By the same line of argument, all cases in the $n \in [0, 1, 2]$ space can be shown to imply assortative matching by using only part A of the proof. ■

Appendix B: Proof of Proposition 3

Proof. Existence of a unique reservation productivity, p^r_w, and a unique unemployment rate, u, follows trivially from equations (4) and (7). The l.h.s. of equations (5) and (6) is increasing in n, while the r.h.s. is decreasing in n, which imply that there is a unique solution for each of the equations. If this value is not an integer, it is possible that the two integers adjacent to the solution of first-order condition imply the same value of the value function. In this case, we assume that the worker chooses the lowest n. Under this assumption there exist a unique $n(p_w, b)$ and $n(p_w, p_{fm})$ for any $p_w \in \left[\underline{p}_w, \bar{p}_w\right]$ and any $p_{fm} \in \left[\underline{p}_f, \bar{p}_{fm}\right]$. Differentiating (10) with respect to p_{fm} gives

$$
\begin{aligned}
\hat{g}\left(p_w, p_{fm}\right) &= \frac{\delta}{\delta+\lambda} n\left(p_w, b\right) \Gamma\left(p_{fm}\right)^{n(p_w,b)-1} + \frac{\lambda}{\delta+\lambda} \Gamma\left(p_{fm}\right)^{n(p_w,p_{fm})} \hat{g}\left(p_w, p_{fm}\right) \\
&\quad + \frac{\lambda}{\delta+\lambda} \int_{\underline{p}_{fm}}^{p_{fm}} n\left(p_w, p'_{fm}\right) \Gamma\left(p_{fm}\right)^{n\left(p_w,p'_{fm}\right)-1} \hat{g}\left(p_w, p'_{fm}\right) dp'_{fm} \\
&\Leftrightarrow \\
\hat{g}\left(p_w, p_{fm}\right) &= \frac{\frac{\delta}{\delta+\lambda} n\left(p_w, b\right) \Gamma\left(p_{fm}\right)^{n(p_w,b)-1}}{1 - \frac{\lambda}{\delta+\lambda} \Gamma\left(p_{fm}\right)^{n(p_w,p_{fm})}} + \frac{\frac{\lambda}{\delta+\lambda} \int_{\underline{p}_{fm}}^{p_{fm}} n\left(p_w, p'_{fm}\right) \Gamma\left(p_{fm}\right)^{n\left(p_w,p'_{fm}\right)-1} \hat{g}\left(p_w, p'_{fm}\right) dp'_{fm}}{1 - \frac{\lambda}{\delta+\lambda} \Gamma\left(p_{fm}\right)^{n(p_w,p_{fm})}}
\end{aligned}
$$

which is an inhomogeneous Volterra equation of second kind with an everywhere continuous and uniformly bounded integral kernel. Given uniqueness of $n\left(p_w, b\right)$ and $n\left(p_w, p_{fm}\right)$ we also have a unique solution for $\hat{g}\left(p_w, p_{fm}\right)$ for any $p_w \in \left[\underline{p}_w, \bar{p}_w\right]$ and any $p_{fm} \in \left[\underline{p}_{fm}, \bar{p}_{fm}\right]$. ■

References

Abowd, J. M., R. Creecy, and F. Kramarz (2002): "Computing person and firm effects using linked longitudinal employer-employee data," Technical Paper 2002-06, U.S. Census Bureau.

Abowd, J. M., F. Kramarz, P. Lengermann, and S. Perez-Duarte (2004): "Are good workers employed by good firms? A test of a simple assortative matching model for France and the United States," working paper.

Abowd, J. M., F. Kramarz, and D. N. Margolis (1999): "High wage workers and high wage firms," *Econometrica*, 67(2), 251 – 333.

Abowd, J. M., K. McKinney, and I. M. Schmutte (2010): "How important is endogenous mobility for measuring employer and employee heterogeneity?," working paper.

Andrews, M., L. Gill, T. Schank, and R. Upward (2008): "High wage workers and low wage firms: Negative assortative matching or limited mobility bias?," *Journal of the Royal Statistical Society Series A*, 171(3), 673–697.

Atakan, A. E. (2006): "Assortative matching with explicit search costs," *Econometrica*, 74(3), 667–680.

Bagger, J., and R. Lentz (2012): "An empirical model of wage dispersion with sorting," working paper.

Barth, E., and H. Dale-Olsen (2003): "Assortative matching in the labour market? Stylised facts about workers and plants," working paper.

Becker, G. S. (1973): "A theory of marriage: Part I," *Journal of Political Economy*, 81(4), 813–846.

Burdett, K., and M. G. Coles (1999): "Long term partnership formation: Marriage and employment," *Economic Journal*, 109, F307–F334.

Chade, H. (2001): "Two-sided search and perfect segregation with fixed search costs," *Mathematical Social Sciences*, 42, 31–51.

de Melo, R. L. (2009): "Sorting in the labor market: Theory and measurement," working paper.

Eeckhout, J., and P. Kircher (2011): "Wage bargaining with on-the-job search: Theory and evidence," *Review of Economic Studies*, 78(3), 872–906.

Gilmour, A. R., R. Thompson, and B. R. Cullis (1995): "Average information REML: An efficient algorithm for variance parameter estimation in linear mixed models," *Biometrics*, 51(4), 1440–1450.

Gruetter, M., and R. Lalive (2004): "The importance of firms in wage determination," IZA Discussion Paper No. 1367.

Jovanovic, B. (1979): "Job matching and the theory of turnover," *Journal of Political Economy*, 87(5), 972–990.

Kennes, J., and D. le Maire (2013): "Job Heterogeneity and Coordination Frictions," working paper.

Lentz, R. (2010): "Sorting by search intensity," *Journal of Economic Theory*, 145(4), 1436–1452.

Lise, J., C. Meghir, and J.-M. Robin (2008): "Matching, Sorting and Wages," working paper.

Piekkola, H. (2006): "Knowledge capital as the source of growth," working paper.

Postel-Vinay, F., and J.-M. Robin (2002): "Equilibrium wage dispersion with worker and employer heterogeneity," *Econometrica*, 70(6), 2295–2350.

——— (2006): "Microeconometric search-matching models and matched employer-employee data," in *Advances in Economics and Econometrics, Theory and Applications, Ninth World Congress*, ed. by R. Blundell, W. Newey, and T. Persson, vol. 2. Cambridge University Press.

Postel-Vinay, F., and H. Turon (2010): "On-the-job search, productivity shocks, and the individual earnings process," *International Economic Review*, 51(3), 599–629.

Shimer, R., and L. Smith (2000): "Assortative Matching and Search," *Econometrica*, 68(2), 343–369.

Smith, L. (2006): "The marriage model with search frictions," *Journal of Political Economy*, 114(6), 1124–1144.

Soerensen, T., and R. Vejlin (2010): "The importance of worker, firm and match fixed effects in the formation of wages," working paper.

Stigler, G. J. (1961): "The economics of information," *Journal of Political Economy*, 69(3), 213–225.

Woodcock, S. D. (2010): "Heterogeneity and learning in the labor market," *The B.E. Journal of Economic Analysis and Policy*, 10(1).

——— (2011): "Match effects," working paper.