Taking competitive search to data

Anna Godøy and Espen R. Moen

February 15, 2011

Abstract

In search theory, an important distinction can be drawn between models with directed search and with random /undirected search. In the present paper we first present a simple model of competitive on-the-job search, where firms' productivity differences emerge endogenously through the firms' investment choices. We believe this model gives some new insights in its own right. Then we discuss testable differences between our model and the BM model. In the empirical part of the paper, we use Norwegian data to discriminate between the theories. To overcome problems with noise in the data, the estimation is done in two steps, fitting a finite mixture model to classify data into two groups. Results are consistent with the competitive search model.

Keywords: competitive search, random search, directed search, register data, observable differences, empirical tests

JEL Classification Numbers: J30, J60

1 Introduction

In search theory, an important distinction can be drawn between models with directed search and with random /undirected search. With directed search, wages are observed by workers prior to their search decision. With random search, workers learn the wages a firm posts after being matched with the firm. In the Burdett-Mortensen (BM) model the firms post wages. Other models of random search assume wage bargaining. The two modes of wage determination have different implications for the role of wages in allocating resources. With directed search, wages influence which firms workers direct their search. With random search, wages only influence the *ex post* decision of whether or not to accept the match. As a result, directed search gives rise to an efficient allocation of resources in a wide variety of environments, while random search typically does not.

In the present paper we first present a simple model of competitive on-the-job search, where firms' productivity differences emerge endogenously through the firms' investment choices. This gives a purified version of the competitive search equilibrium with on-the-job search relative to that of Garibaldi and Moen (2010*b*, 2010*a*). We believe this model gives some new insights in its own right. Then we discuss testable differences between our model and the BM model. In the empirical part of the paper, we use Norwegian data to discriminate between the theories. Microdata has been used extensively in empirical work on equilibrium search and matching models (see, e.g., Cahuc et al. 2006, Christensen et al. 2005, Lise et al. 2009). We find that noise in the data may complicate our empirical strategy if a majority of workers direct their search. To overcome this, the estimation is done in two steps, first estimating a finite mixture model to classify data into two groups. This estimation yields results consistent with the competitive search hypothesis.

2 The model

The paper is related to Garibaldi and Moen (2010*b*, 2010*a*), but with some important differences, as the productivity of the firms in our model is endogenous.

- All agents are risk neutral, infinitely lived, and discount the future at rate r.
- Workers are identical, search equally well on and off the job, and receives y_0 when unemployed. Their number (measure) is exogenous and normalized to 1.
- Before they enter the market, firms choose a technology from a menu $y = f(K)$, where y is output, K is investment, and $f(K)$ is strictly increasing and concave.
- Firms constantly post one vacancy, and hence hire continuously (can easily be extended to v vacancies with convex maintenance cost $c(v)$).
- Firms die at rate δ. In addition, workers separates from firms at an exogenous rate s.
- No internal coordination problems in firms -workers search so as to maximize joint income. This may be because workers and firms can contract upon the worker's on-the job search behaviour directly. Alternatively, the worker may pay a quit fee to the firm if leaving or buy the job up-front after being matched. (see Moen & Rosén (2004) for more examples).

The assumption that firms contract efficiently implies that the remuneration of the worker does not influence her on-the-job search behaviour.

The search market endogenously separates into submarkets, consisting of a set of workers and firms with vacancies searching for each-other. In each submarket, the flow of matches is determined by a constant-returns-to scale matching function $x(u, v)$, where u and v are the measures of workers and firms in that submarket, respectively. Let $\theta = v/u$, and define $p(\theta) = x(u, v)/u =$

$x(1,\theta)$ and $q(\theta) = x(u,v)/v = x(1/\theta,1)$. Finally. let $\eta = |q'(\theta)\theta/q|$ denote the absolute value of the elasticity of η with respect to θ. It is convenient to assume that $\eta(\theta)$ is non-decreasing in θ.

Firms advertise contracts and workers search for the different contracts. Suppose a firm offers a wage and on-the-job search possibilities so that the workers' expected lifetime income is W. We then say that the firm advertises an NPV wage (or just wage) W. Suppose a vector $W_1, ..., W_k ...$ of wages are advertised. The firms that advertise a given wage and the workers that apply to those firms form a submarket. Let $\theta_1, ..., \theta_k ...$ denote the associated vector of labor market tightness.

As will be clear below, firms will diversify and choose different investment levels, and hence enter the market with different productivities. Suppose the different values of y forms a countable set $\{y_1, y_2, ...\}$, where $y_i < y_{i+1}$. Let $M_i = M(y_i)$ denote the joint expected discounted income flow of a worker and a job in a firm of type i, where the gains from on-the-job search is included. Since on-the-job search is efficient, it follows that M_i is given by

$$rM_i = y_i + (s+\delta)(M_0 - M_i) + \max_k p(\theta_k)[W_k - M_i] \tag{1}$$

The first term is the flow production value created on the job. The second term captures the expected capital loss due to job separation, which happens at rate $s+\delta$, and reduces the joint income to M_0 (since the firm then earns zero on this match). The last term shows the expected joint gain from on-the-job search. Since the current wage is a pure transfer from the employer to the worker, it does not appear in the expression. From (1) it follows that the optimal search behaviour of a worker depends on her current position, as this influences M_i.

The indifference curve of a worker of type i shows combinations of θ and W that gives a joint income equal to M_i. We can represent this as $\theta_i = f_i(W; M)$.[1] It follows that f_i is defined implicitly by the equation

$$rM_i = y_i + (s+\delta)(M_0 - M_i) + p(g_i(W, M))[W - M_i] \tag{2}$$

where M_i is the equilibrium joint income in firm i. It follows that for $M_i < W_i$

$$g_i(W; M) = p^{-1}\left(\frac{(r+s+\delta)M_i - y_i - (s+\delta)M_0}{W - M_i}\right) \tag{3}$$

The indifference curve is defined for all W, not only the values advertised in equilibrium. Define

$$g(W; M) = \min_{i \in \{0,1,...,n\}} g_i(W; M) \tag{4}$$

The function $g(W; M)$ is thus the lower envelope of the set of functions $g_i(W; M)$. In equilibrium, $g(W; M)$ shows the relationship between the wage advertised and the labor market tightness in a submarket. Suppose that for a given W,

[1] Strictly speaking, f_i only depends on M_i and M_0, but we write it as a function of the vector M for convenience

the minimum in (4) is obtained for worker type i'. This worker type will then flow into the market up to the point where $\theta = g_{i'}(W;M)$. At this low labor market tightness, no other worker types want to enter this submarket. The labor market tightness is thus given by $g_{i'}(W;M)$, and only workers of type i' enter the market.

Consider then a firm with productivity y_j. At any point in time, a firm of type j maximizes the value of search given by[2]

$$\pi(y_j.W) = -c + q(\theta)[M_j - W]. \tag{5}$$

where W_j is the NPV wages paid by the firm. The first part is the flow cost of posting vacancies, while the second part is the gain from search. The firm maximizes profit with respect to W given that $\theta = g(W;M)$. Denote the associated maximum profit flow by $\pi_j^*(y_j)$. The expected profit of a firm entering the market as a type j firm is thus

$$\Pi_j = \frac{\pi_j^*(y_j)}{r+\delta} \tag{6}$$

At the entry stage, the firm chooses the investment level K, hence we can write $\Pi(K) = \frac{\pi^*(y(K))}{r+\delta}$. Firms choose K so as to maximize $\Pi(K) - K$.

Let θ_{ij} denote the labor market tightness in a labor market where workers in firms of type i searches for firms of type j.

Definition 1. *A competitive search equilibrium is a vector $K_1, K_2....$ of investment levels, a vector of joint incomes M_0,M_1, ..., a matrix of wages W_{ij}, $i,j \in \{1,2,...\}$, and a matrix θ_{ij}, $i,j \in \{1,2,...\}$ such that*

1) M_i is given by (1) with $y_i = f(K_i)$

2) W_{ij} maximizes $\pi_j(y(K_j), W)$ for some j

3) $\Pi(K) - K \leq 0$ for all K with equality for all K chosen in equilibrium

4) Aggregate consistency: inflow of workers equal to outflows in all submarkets

3 Characterizing equilibrium

We will now show that the equilibrium of the model can be characterized in a simple way. Our first observation is that the vector of investment levels $K_1, K_2, ...$ cannot be finite. Suppose there is a highest value K_n. Consider a firm that is investing $K_n + 1$. Since $f'(K) > 0$ it follows that this firm can offer a wage slightly above $y(K_n)$, fill its vacancies infinitely quickly, and hence obtain an infinite profit.

The next result is key for characterizing equilibrium

[2] At any point in time, the firm decides on the number of vacancies to be posted and the wages attached to them. This only influences profits through future hirings, and is independent of the stock of existing workers.

Proposition 1. *The equilibrium is a pure job ladder, where workers working in firms of type j with capital K_j search for jobs in firms of type $j+1$ with capital K_{j+1} only.*

(not proved yet). Consider a firm of type j. To see why, consider a firm that attracts workers employed in firms of type j. The investment level and wage must solve the dual problem

$$\max_{W,\theta,K} M_j = \max_{W,\theta,K} y_j + p(\theta)(W - M_j) \qquad \text{S.T.} \;\; \Pi(K) = K \tag{7}$$

The solution of this problem depends on j, hence a worker of type j will never want to search for other firms than those solving (7) of worker type j. In addition it is easy to show that the value of K that solves (7) is increasing in worker type.

Taking the derivative of (1) with respect to y gives (due to the envelope theorem)

$$M'(y) = \frac{1}{r + s + \delta + p_{j+1}}$$

First order condition for K requires that $\Pi'(K) = 1$, or

$$q_j \frac{f'(K_j)}{r + s + \delta + p_{j+1}} = r + \delta$$

The first order condition for wages is given by

$$W_j = M_{j-1} + \eta(M_j - M_{j-1}) \tag{8}$$

which uniquely defines W_j (assuming that η is non-decreasing in θ). It follows that the zero profit constraint writes

$$q(\theta_j)(1 - \eta)M(y(K_j) = r + \delta$$

For given $M()$, this uniquely pins down θ_j.

Proposition 2. *1. The equilibrium can be characterized as a vector of investments K_1^*, K_2^*,...,joint incomes M_0^*, M_1^*, ..., and labor market tightness θ_0^*, θ_1^*, ... satisfying the following conditions*

a) Optimal investments

$$q_j(\theta_i^*) \frac{f'(K_j)}{r + s + \delta + p_{j+1}} = r + \delta$$

b) Joint incomes

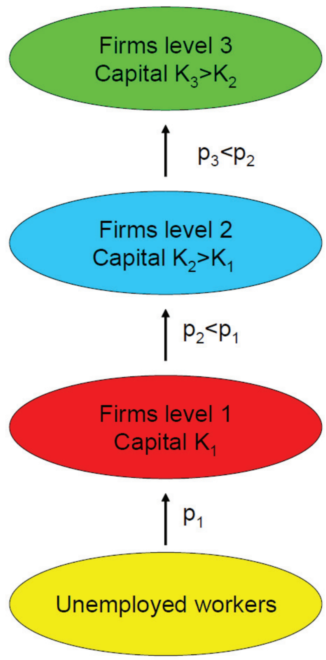
$$M(y(K_j^*)) = \frac{y(K_j^*) + \eta(M(y_{j+1}) - M(y_j)) + (s + \delta)M^0}{r + \delta + s}$$

c) Zero profits

$$q(\theta_j)(1 - \eta)M(y(K_j)) = r + \delta$$

2. $K_1^* < K_2^* < K_3^*$.... and $\theta_1 > \theta_2 > \theta_3$......

Proof: To be done, but should not be too difficult



4 Empirical analysis

In this section we will briefly discuss testable differences in predictions between our model and some other important models of on-the-job search.

The Burdett-Mortensen (BM) model. In the BM model (Burdett & Mortensen 1998), a firm's output is proportional to the labor force,and firms post wages prior to being matched with workers. Firms and workers match randomly, hence the distribution of wages after successful on-the-job search is equal to the wage distribution over vacancies truncated at previous wage w^o ($w \geq w_o$). Hence, if the wage distribution over vacancies is denoted by $H^v(w)$, it follows that

$$H(w|w^o) = \frac{H^v(w) - H^v(w^0)}{1 - H^w(w^0)}$$

The support of the distribution is $[w^o, w^s]$, where w^s is the supremum of the support of advertised wages. Define $H(w|w \geq \overline{w}, w^0)$ as the distribution function of new wages w, contingent on $w \geq \overline{w}$, as a function of the old wage w^0. Then for any $w \geq \overline{w} \geq w^0$,

$$\begin{aligned} H(w|w \quad \geq \quad \overline{w}, w^0) = \frac{H(w|w^0) - H(\overline{w}|w^0)}{1 - H(\overline{w}|w^0)} \\ = \frac{H^w(w) - H(\overline{w})}{1 - H^w(\overline{w})} \end{aligned}$$

independently of w_i^o.

Similarly, the distribution of prior wages $H^{old}(w^0|w^1)$ (where w^1 is still the new wage) is equal to the distribution of wages over employees (including unemployment benefit) truncated at w^1. By using the same argument as above it follows that $H^{old}(w|w^1, w \leq \overline{w})$ is independent of w^1 for all $w^1 > w$.

The Postel-Vinay and Robin (PR) wage setting procedure. Postel-Vinay & Robin (2002) assume that after successful on-the-job search, the incumbent firm and the new firm compete for the worker in a Bertrand fashion. Furthermore, firms compete in NPV wages, hence a worker takes into account that expected future wages (after encountering another job offer) will be higher the higher is the productivity of the employer. The latter is referred to as the option value of the job.

If we we consider productivity instead of wages, the results for the BM model carries over to this model. If we let D denote the distribution of productivities of the new firm after successful on-the-job search, it follows that D is equal to the productivity distribution of the vacancies truncated at the productivity level of the previous employer. Hence, if the productivity distribution over vacancies is given by $D^p(y)$, it follows that

$$D_p(y^1|y^0) = \frac{D^p(y) - D^p(y^0)}{1 - D^p(y^0)}$$

Consider then the distribution of wages $D_w(w|w_o)$. There is no one-to-one correspondence between wages and productivity in a given job. However, wages and productivity are positively correlated. Hence there will be a positive correspondence between wages in a previous job and wages in the new job, even though the distribution of underlying productivities does not show such a correspondence.

Due to the fact that the option value is increasing in productivity, the model predicts that, contingent the productivity of the employer, there is a *negative* relationship between wages in the new job and the productivity of the new employer.

Consider then the distribution of wages in a given firm. On average, a high-productivity firm will pay higher NPV wages when attracting workers, since they will be willing to bid higher and be able to attract workers previously employed in more productive firms. On the other hand, since the option value of staying in a high-productivity firm is higher than the option value of staying in a low-productivty firm, hence contingent on the productivity of the previous employer, the more productive workers pay less.

Competitive on-the-job search. Our model is not a model of wages, but rather of NPV wages. However, assume that the wage that the worker obtains in a firm is constant, and that the workers' search behaviour is contracted upon directly. As the value of job search is lower the higher is the worker in the hierarchy, it follows easily that $w = w(W)$, $w\prime(W) > 0$.In the competitive search equilibrium it follows that there is a one-to-one correspondence between wages and productivity and between wages before and after a job change.

5 Data

The employer-employee register, containing data on earnings and duration of jobs, is used to construct a dataset on job-to-job transitions. Separate spells with the same employer are counted as a single spell, that is, recall unemployment is ignored. The sample consists of job-to-job transitions in 2004. This data can be matched with personal characteristics including education level. Individuals who have changed their education level are excluded from the sample, as are all people aged under 25 or over 55 in the year of transition. Observations where earnings or wage lies in the top or bottom one percent of the population distribution are also excluded from the sample. Wages are deflated to 2003 levels using the CPI. A job-to-job transition is defined as an individual worker changing employers (firm and establishment), where there was at most 30 days break between jobs.

The employer-employee register contains data on annual earnings for each job. In order to compare earnings before and after the transition in a meaningful way using data from a single year, we would need precise data on what fraction of the year the person has been employed in each job as well as the number of hours worked. However, data on hours worked and quit dates in the employment register are recorded with error. We take this into account in two ways.

First, the sample of annual earnings includes only persons who either worked full time in both jobs or the same part time fraction in both jobs, being employed in the old job all year 2003 and in the new job all year 2005. Then the old-job earnings 2003 can be compared with the new-job earnings in 2005. This procedure excludes 40% of job-to-job transitions, leaving a sample of 44279 transitions. We are concerned that excluding these observations, keeping only the most stable part of the workforce, may make results less representative.

Second, for some jobs, data is available for hourly wages as well as hourly earnings. The data source for hourly wages is Statistics Norway's "Wage Statistics": For public sector employees, data on wages and hours are collected using existing register data. For private sector workers, the data source is questionnaires on individual workers filled out by all firms included in the annual sampling. Large firms are oversampled in the survey; in total, hourly wage data is available for 50-65% of workers in each industry each year. Requiring data on hourly wages to be recorded for both jobs excludes 75% percent of job-to-job transitions, leaving 17760 observations. Due to the sampling structure of the "Wage Statistics" data set, individuals in the wage sample are more likely to be working in large firms compared to the earnings sample, but we will exclude fewer people who worked in the job for a shorter period.

Some descriptive statistics for each sample are presented in table 1. Somewhat unexpectedly, average earnings in the new job are lower than average earnings in the old job. 52.2% of transitions in the earnings sample are associated with a wage cut. This may in part be due to problems with incomplete data on jobs in the employer-employee register, making it difficult to see whether a person has actually worked a full year in the new job. If so, this is another reason why we should use both sources of data.

Table 1: Descriptive statistics

(a) Annual earnings

	mean	sd
Earnings in old job 2003	347069.7	131618.6
Earnings in new job 2005	346451.0	132673.4
Age in year of transition	40.16	8.218
Male	69.3%	
Basic education only	7.01%	
Secondary school degree	58.1%	
Higher education	33.7%	
Education unknown	1.20%	
Observations	42438	

(b) Hourly wages

	mean	sd
Wage in old job 2003	204.0	67.64
Wage in new job 2005	206.1	67.42
Age in year of transition	40.48	8.201
Male	66.0%	
Basic education only	6.84%	
Secondary school degree	60.6%	
Higher education	31.4%	
Education unknown	1.13%	
Observations	17760	

Table 2: Regression on subsample $w_1 > \bar{w} > w_0$

(a) Annual earnings

	(1)	(2)	(3)
$\bar{w}$ is percentile	25^{th}	50^{th}	75^{th}
Earnings old job	-0.0784***	-0.0584***	-0.0376***
	(-14.74)	(-11.79)	(-6.29)
Constant	13.63***	13.57***	13.54***
	(210.27)	(219.83)	(178.39)
Observations	3376	3685	2418
Adjusted R^2	0.060	0.036	0.016

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(b) Hourly wages

	(1)	(2)	(3)
$\bar{w}$ is percentile	25^{th}	50^{th}	75^{th}
Wage old job	-0.169***	-0.0847***	-0.0304
	(-6.78)	(-3.53)	(-1.40)
Constant	6.036***	5.802***	5.736***
	(48.95)	(47.50)	(50.13)
Observations	1321	1129	937
Adjusted R^2	0.033	0.010	0.001

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

6 Empirical strategy

Considering the subsample of the population where the wage in the new job is above some cutoff $\bar{w}$ and the wage in the old job is below $\bar{w}$: $w_1 > \bar{w} > w_0$, suppose we were to estimate a simple regression equation

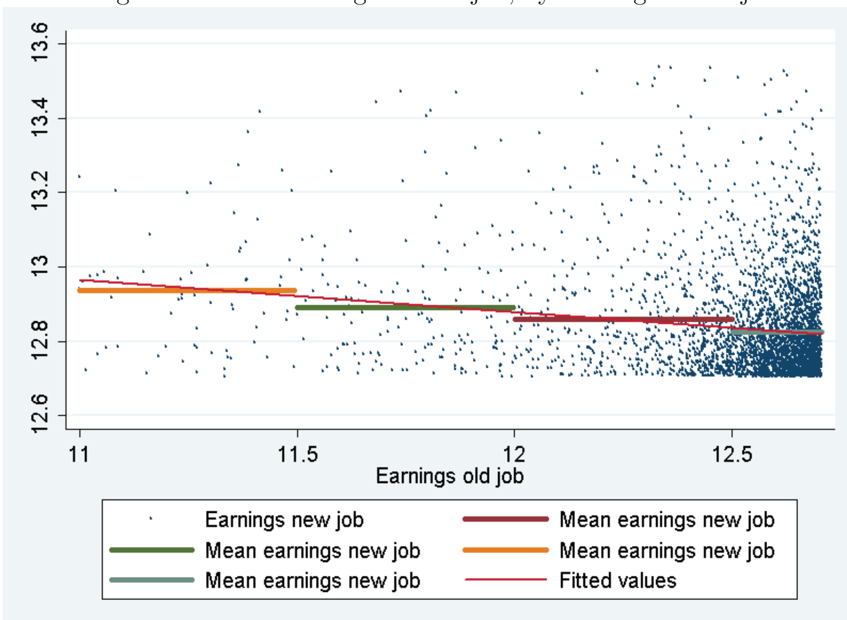
$$w_{1i} = \alpha + \beta w_{0i} + \varepsilon_i \tag{9}$$

on this subsample. From the discussion in section 4 it follows that, on this subsample, CSE would predict $\beta > 0$ while the BM model (taken literally) would imply $\beta = 0$.

Equation (9) is estimated setting $\bar{w}$ equal to the 25^{th}, 50^{th} and 75^{th} percentile of the sample distribution. Results are shown in table 2.

In five of the six estimations, β is estimated to be negative and significant. It is not clear how this result could be explained by either model of on-the-job-search. Rather, the explanation could be noise in the data leading to negative estimates of β even if a large majority do job-to-job transitions within a search

Figure 1: Mean earnings in new job, by earnings in old job



framework. When we look at the data, it appears that while most persons earn about the same wage in the new and the old job, a few job-to-job transitions are characterized by very large (positive and negative) change in wages. If some small fraction of the sample experience job-to-job transitions due to some other process, eg. as a draw from some distribution of wages, we may be unable to distinguish between different models of on-the-job search by estimating (9) on the full set of observations where $w_1 > \bar{w} > w_0$.

Figure 1 illustrates the point. The figure shows an excerpt of observations of old and new earnings where $w_1 > \bar{w} > w_0$. The density of observations is highest in the lower right hand corner of the graph, representing observations with a small earnings increase. For lower values of old earnings, there are much fewer observations. Mean wage in new job is calculated separately for different segments of wage in the old job, mean wage in new job is consistently higher for groups where wage in the old job was low. It would appear that the negative slope of the fitted line is largely driven by the few dispersed observations of low old-job wages.

In general, if directed search is important, regular job-to-job movements imply relatively small upward adjustments of the wage, and the observations of regular job movers will concentrate in lower right-hand side corner in the figure. Hence for old wages close to $\bar{w}$, ordinary job transitions dominate, and the new wage is close to $\bar{w}$. For lower initial wages, few ordinary job movers will obtain a new wage above $\bar{w}$. Hence for this initial wage, the noisy movers will dominate, and contingent on being above $\bar{w}$ these obtain relatively high wages after the

Table 3: Step 1, estimated finite mixture model

(a) Earnings

Cluster 0	π_0	0.2090
	μ	12.6548
	σ_0	0.5342
Cluster 1	π_1	0.7910
	a	0.9515
	b	0.9251
	σ_1	0.1336

(b) Hourly wage

Cluster 0	π_0	0.1600
	μ	5.2820
	σ_0	0.2954
Cluster 1	π_1	0.8400
	a	0.3595
	b	0.9316
	σ_1	0.0947

transition. As a result, the average wage after the transition may be higher for low initial wages than for initial wages close to $\bar{w}$.

We use a finite mixture model strategy to control for this kind of noise (Leisch 2008). A similar approach has been used to deal with background noise in mixtures of regression models. In these models, a few outliers can complicate the identification of mixture components. Including an empty regression model component can be useful as a simple way of controlling for these outliers.

The estimation is done in two steps: First, a finite mixture model is estimated in order to identify which observations are noise. The second step estimates (9) separately on the "noise" and the "not-noise" observations.

For a fraction π_1 of the population, (log) wage in the new job is assumed to be a linear function of wage in the old job: $w_{1i} = a + bw_{0i} + \varepsilon_i$, where $\varepsilon_i \sim N(0, \sigma_1)$ is gaussian noise. Both models of on-the-job search predict $b > 0$. For a fraction $\pi_0 = 1 - \pi_1$, the wage in the new job is assumed to be normally distributed $w_{0i} \sim N(\mu_0, \sigma_0)$, independent of the wage in the old job. That is, the model can be written in terms of the likelihood contribution of individual i:

$$f(w_{1i}|w_{0i}, \psi) = \pi_0 \times \phi(\frac{w_{1i} - \mu_0}{\sigma_0}) + (1 - \pi_0) \times \phi(\frac{w_{1i} - a - bw_{0i}}{\sigma_1}) \qquad (10)$$

where ϕ denotes the density of the standard normal distribution and $\psi = \{\pi_0, \pi_1, \mu, a, b, \sigma_0, \sigma_1\}$ is the vector of model parameters. The model is estimated by maximum likelihood using the iterative EM algorithm (Dempster et al. 1977). Results are given in table 3.

Having obtained maximum likelihood estimates of $\hat{\psi}$, the posterior class probabilities are be calculated for each observation as follows:

$$\hat{p}_{0i} = P(0|w_{0i}, w_{1i}, \hat{\psi}) = \frac{\hat{\pi}_0 \phi(\frac{w_{1i} - \hat{\mu}_0}{\hat{\sigma}_0})}{\hat{\pi}_0 \phi(\frac{w_{1i} - \hat{\mu}_0}{\hat{\sigma}_0}) + \hat{\pi}_1 \phi(\frac{w_{1i} - (\hat{a} + \hat{b} w_0 i)}{\hat{\sigma}_1})}$$

$$\hat{p}_{1i} = P(1|w_{0i}, w_{1i}, \hat{\psi}) = \frac{\hat{\pi}_1 \phi(\frac{w_{1i} - (\hat{a} + \hat{b} w_0 i)}{\hat{\sigma}_1})}{\hat{\pi}_0 \phi(\frac{w_{1i} - \hat{\mu}_0}{\hat{\sigma}_0}) + \hat{\pi}_1 \phi(\frac{w_{1i} - (\hat{a} + \hat{b} w_0 i)}{\hat{\sigma}_1})}$$

and each observation can be assigned to the cluster with the highest probability.

Figure 2 illustrates this first step. The sample is separated into two clusters, where cluster 0 contains those observations where $\hat{p}_{0i} > \hat{p}_{1i}$.

Having assigned each observation to the cluster with the highest posterior class probability, the regression equation (9) is estimated on the subsample where $w_1 > \bar{w} > w_0$, separately on cluster 0 and cluster 1. The estimated models are reproduced below.

This approach yields estimates of β that are significant and positive for both hourly wage data and annual earnings data. Tables 4 and 5 show estimates of (9). The cutoffs $\bar{w}$ are the same as in table 2, but now the estimation is done separately on observations assigned to each cluster.

Table 4: Step two, annual earnings

(a) Cluster 0

	(1)	(2)	(3)
$\bar{w}$ is percentile	25^{th}	50^{th}	75^{th}
Earnings old job	0.0203^{**}	0.0404^{***}	0.0477^{***}
	(2.68)	(5.31)	(5.58)
Constant	12.55^{***}	12.46^{***}	12.56^{***}
	(139.78)	(135.10)	(117.99)
Observations	1462	1243	820
Adjusted R^2	0.004	0.021	0.035

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(b) Cluster 1

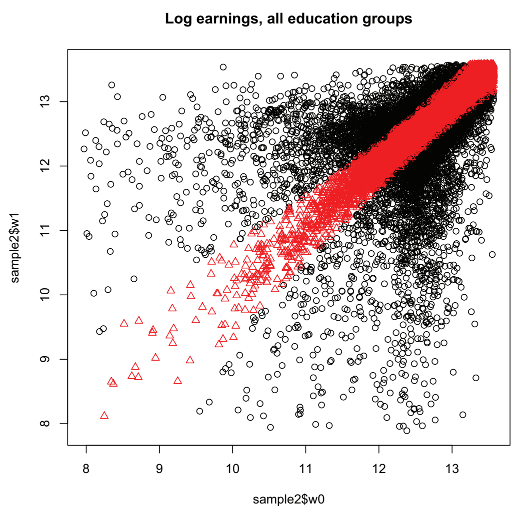
	(1)	(2)	(3)
$\bar{w}$ is percentile	25^{th}	50^{th}	75^{th}
Earnings old job	0.231^{***}	0.233^{***}	0.232^{***}
	(10.15)	(12.31)	(10.71)
Constant	9.725^{***}	9.848^{***}	10.04^{***}
	(34.43)	(41.30)	(36.06)
Observations	1914	2442	1598
Adjusted R^2	0.051	0.058	0.066

t statistics in parentheses

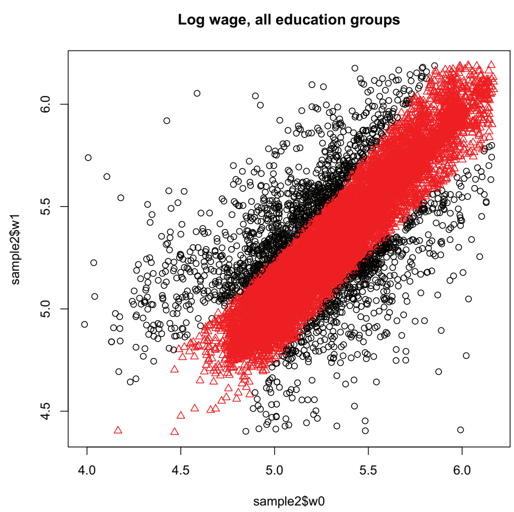
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Comparing table 2 with tables 4 and 5, the results from the two step estimation are very different from the earlier simpler approach. Panels 4(b) and 5(b) show β estimated on cluster 1. The effect of old wages on new wages on the subsample where $w_1 > \bar{w} > w_0$ is positive and significant; the result is robust to choice of cutoff $\bar{w}$.

Figure 2: Step 1 - cluster membership assigned by maximum posterior class probability



(a) Annual earnings



(b) Hourly wages

Table 5: Step 2, hourly wages

(a) Cluster 0

	(1)	(2)	(3)
$\bar{w}$ is percentile	25^{th}	50^{th}	75^{th}
Wage old job	0.210***	0.212***	0.189***
	(5.83)	(6.65)	(6.46)
Constant	4.299***	4.394***	4.654***
	(24.66)	(27.63)	(30.71)
Observations	512	546	479
Adjusted R^2	0.061	0.073	0.078

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(b) Cluster 1

	(1)	(2)	(3)
$\bar{w}$ is percentile	25^{th}	50^{th}	75^{th}
Wage old job	0.211***	0.280***	0.388***
	(7.06)	(7.01)	(9.24)
Constant	4.074***	3.848***	3.428***
	(27.31)	(18.62)	(15.18)
Observations	809	583	458
Adjusted R^2	0.057	0.076	0.156

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figures 3 and 4 illustrate the two step approach. The upper panels show observations in each cluster and the upward sloping regression lines when estimation is done separately by each cluster. The lower panels show the regression fitted on the pooled sample, where observations with large increase in wages appear to be driving the downward sloping regression line.

It is worth noting that β is also estimated to be positive for cluster 0, the ”noise” observations. For earnings, the effect is statistically significant but smaller for these. For wages, the estimated β is about the same for both clusters. It thus appears that there is a relationship between the wage before and after a job change and that this effect is there whether the wage change is large or small.

Recalling the discussion in section 4, this effect is consistent with the CSE model of on-the-job search. The effect can also be explained by heterogeneity among workers. Persistent differences in ability or human capital would generate a similar tight correspondence between earnings before and after a job change.

We have estimated the model by education level; the effects do not change much as a result of this. Setting $\bar{w}$ equal to the median of new earnings or wages w_1 in each sample and estimating (9) on the full set of observations where $w_1 > \bar{w} > w_0$ yields estimates of β that are either negative and significant or not significantly different from zero (table 6). Having fitted a finite mixture model with an empty regression component in order to separate the sample into two clusters, running the estimation on the same sample where $w_1 > \bar{w} > w_0$ for each cluster separately, estimated β is positive and significant for most groups (tables 7 and 8). The magnitude of the effects are broadly similar to the estimates obtained when all education levels where pooled together.

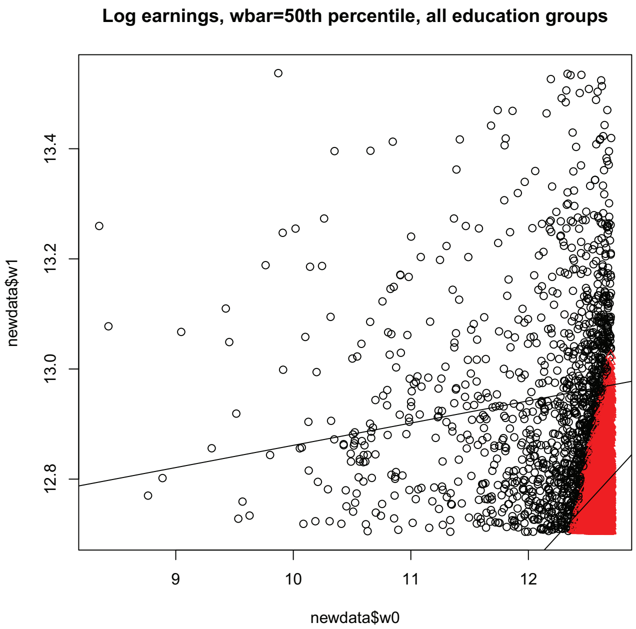
7 Conclusion

In the theory model of this paper, we have derived a model of on-the-job search where productivity differences between firms emerge endogenously as a response to search frictions and directed search. Furthermore, the equilibrium of the model can be described as a pure job ladder, where workers advance the ladder one step at the time.

We then show that the competitive search model implies that wages before on-the-job search influences wages after successful on-the-job search. Furthermore, we also argue that this is not the case for the wage distributions after on-the-job search predicted by the Burdett-Mortensen model. Hence competitive search models and random search models give different predictions at this point.

We have used Norwegian data to discriminate between the two models. A mixture model approach is used to control for noise in the data. Our preliminary findings suggest that there is a positive relationship between wages before and after the job shift, which tend to favour competitive search. However, we want to point out that this may also be due to unobservable fixed worker effects.

Figure 3: Twostep approach, annual earnings, $\bar{w}$ is median



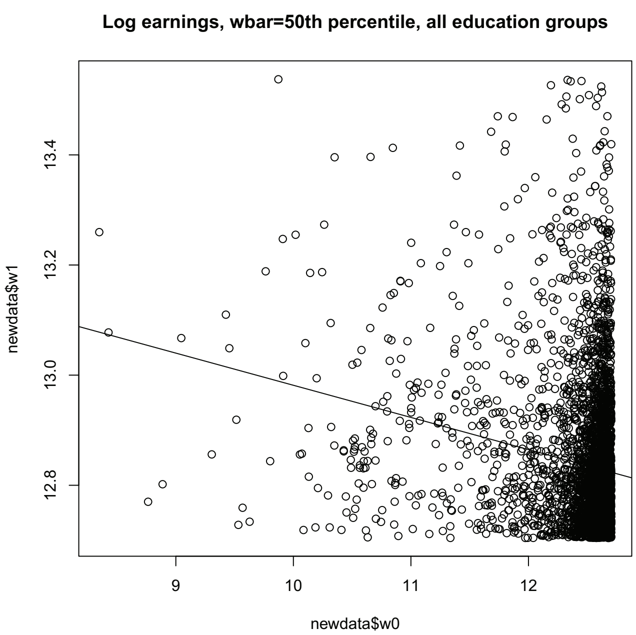
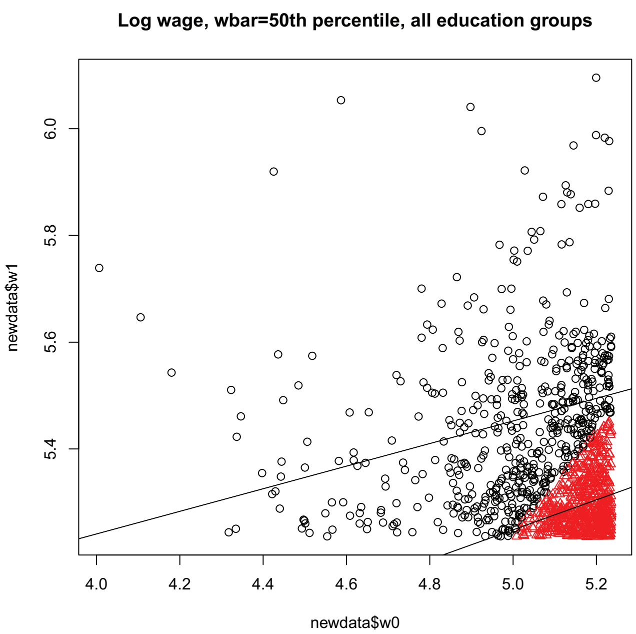


Figure 4: Twostep approach, hourly wage, $\bar{w}$ is median



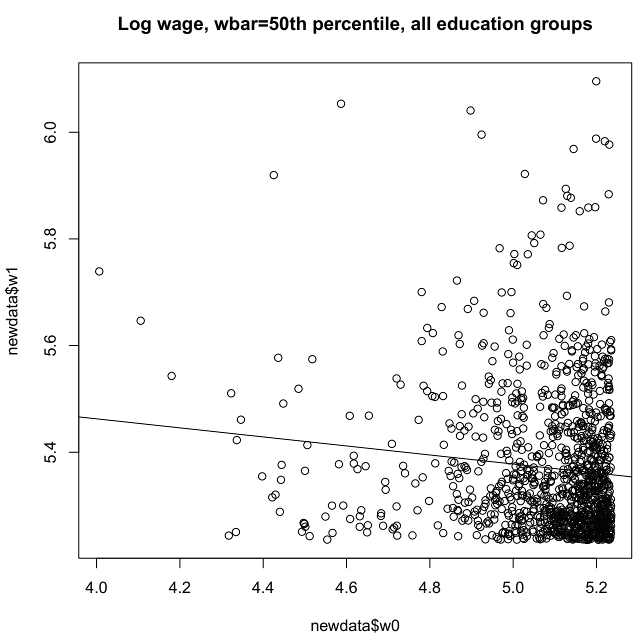


Table 6: Regression on subsample $w_1 > \bar{w} > w_0$, $\bar{w}$ is median

(a) Annual earnings

	(1) Basic education	(2) Secondary school	(3) Higher education
Earnings old job	-0.0765***	-0.0565***	-0.0631***
	(-3.82)	(-9.49)	(-7.35)
Constant	13.63***	13.47***	13.80***
	(55.44)	(182.72)	(126.99)
Observations	251	2323	1321
Adjusted R^2	0.052	0.037	0.039

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(b) Hourly wages

	(1) Basic education	(2) Secondary school	(3) Higher education
Wage old job	-0.161	-0.0752*	0.0120
	(-1.39)	(-2.44)	(0.37)
Constant	6.036***	5.673***	5.509***
	(10.52)	(36.48)	(31.96)
Observations	78	745	487
Adjusted R^2	0.012	0.007	-0.002

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 7: Earnings. Regression on subsample $w_1 > \bar{w} > w_0$, $\bar{w}$ is median

(a) Cluster 0

	(1) Basic education	(2) Secondary school	(3) Higher education
Earnings old job	0.0179	0.0386***	0.0351**
	(0.59)	(4.38)	(2.79)
Constant	12.58***	12.41***	12.69***
	(34.62)	(116.57)	(81.69)
Observations	99	854	382
Adjusted R^2	-0.007	0.021	0.017

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(b) Cluster 1

	(1) Basic education	(2) Secondary school	(3) Higher education
Earnings old job	0.194*	0.213***	0.246***
	(2.38)	(8.76)	(7.79)
Constant	10.21***	10.04***	9.801***
	(10.03)	(32.84)	(24.24)
Observations	152	1469	939
Adjusted R^2	0.030	0.049	0.060

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 8: Hourly wage. Regression on subsample $w_1 > \bar{w} > w_0$, $\bar{w}$ is median

(a) Cluster 0

	(1)	(2)	(3)
	Basic education	Secondary school	Higher education
Wage old job	0.0907	0.229***	0.310***
	(0.50)	(5.53)	(6.84)
Constant	4.859***	4.232***	4.064***
	(5.51)	(20.54)	(17.39)
Observations	42	358	182
Adjusted R^2	-0.019	0.076	0.202

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

(b) Cluster 1

	(1)	(2)	(3)
	Basic education	Secondary school	Higher education
Wage old job	0.413**	0.254***	0.289***
	(2.98)	(5.79)	(5.91)
Constant	3.068***	3.924***	3.970***
	(4.37)	(17.48)	(15.14)
Observations	36	387	305
Adjusted R^2	0.184	0.078	0.100

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

References

Burdett, K. & Mortensen, D. T. (1998), 'Wage differentials, employer size and unemployment', *International Economic Review* **39**(2), 257–273.

Cahuc, P., Postel-Vinay, F. & Robin, J. (2006), 'Wage bargaining with on-the-job search: Theory and evidence', *Econometrica* **74**(2), 323–364.

Christensen, B., Lentz, R., Mortensen, D., Neumann, G. & Werwatz, A. (2005), 'On-the-job search and the wage distribution', *Journal of Labor Economics* **23**(1).

Dempster, A., Laird, N. & Rubin, D. (1977), 'Maximum likelihood from incomplete data via the em algorithm', *Journal of the Royal Statistical Society. Series B (Methodological)* **39**(1), 1–38.

Garibaldi, P. & Moen, E. R. (2010*a*), Competitive on-the-job search.

Garibaldi, P. & Moen, E. R. (2010*b*), 'Job to job movements in a simple search model', *American Economic Review* **100**(2), 343–374.

Leisch, F. (2008), 'Modelling background noise in finite mixtures of generalized linear regression models', *COMPSTAT 2008* pp. 385–396.

Lise, J., Meghir, C. & Robin, J. (2009), Matching, sorting and wages.

Moen, E. R. & Rosén, s. (2004), 'Does poaching distort training?', *Review of Economic Studies* **71**(4), 1143–1162.

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