

ESTIMATING EQUILIBRIUM SEARCH MODELS FROM FINNISH DATA*

TOMI KYRÄ

*Government Institute for Economic Research, P.O. Box 1279,
00101 Helsinki, Finland; e-mail: tomi.kyra@vatt.fi*

*Empirical equilibrium search models have attracted a growing interest in recent years. Estimation of such models has not been completely successful, however. This has led researchers to develop more sophisticated versions of the models in an attempt to get a better fit to the data. We estimate various proposed specifications of the Burdett-Mortensen model from Finnish panel data. We begin with a pure search model in which search frictions are the only source of wage dispersion. Then we proceed to more complex specifications by introducing measurement error in wages and unobserved employer heterogeneity. Our results show how search frictions vary with education, age and sex, and how these differences are reflected in wage differentials in the Finnish labour market.
(JEL: C41, E24, J31, J64)*

1. Introduction

The equilibrium models of the labour market provide a useful framework for analysing various labour market issues. In such models a policy reform or shock that directly affects some subgroup of firms or workers can change the behaviour of all agents on both sides of the market, leading to changes in the equilibrium wage distribution and unemployment rate. When both sides of the labour market are modelled in a dynamic environment, a number of simplifying assumptions must be adopted to keep the model

tractable. Despite simplifying assumptions, the equilibrium solutions are typically rather complex, involving highly nonlinear functions of structural parameters. A consequence is that the equilibrium models are difficult to estimate, and thereby such models are usually calibrated, not estimated.

In calibration, one collects numbers from various sources for some parameters of the model and solves the model with respect to the remaining parameters. This procedure is quite arbitrary, and therefore the estimation of the structural parameters should be preferred. Another issue is to which extent stylized equilibrium models are able to describe the real-world labour market. If the model has not been tested with the data, one must be gullible to take the predictions of the model seriously. Estimating structural equilibrium models from micro-data

* *This study is based on a chapter of my doctoral thesis. The paper has benefited from helpful comments received at the 56th European Meeting of the Econometric Society, the 1st Nordic Econometric Meeting and various seminars held in Finland. Michael Rosholm provided most of Gauss programs needed in the estimations. I thank three anonymous referees and (co-editor) Heikki Kauppi for their suggestions and comments. All the remaining errors and misinterpretations are my responsibility.*

produces more credible parameter estimates, allows us to assess the fit of the model, and yields valuable information for developing a more rigorous basis for equilibrium labour market analysis.

This study considers the estimation of the equilibrium search model of Burdett and Mortensen (1998).¹ The model generates wage differentials between homogenous workers even when all firms are assumed to be identical. This main prediction of the model is consistent with empirical evidence on the existence of considerable wage differentials which cannot be explained by worker and job characteristics (see Bowlus et al., 1995). Other strong predictions concerning the relationship between wages, job durations and firm sizes follow from this simple model as well. Since many of these predictions are consistent with the stylized features of the labour market, the Burdett-Mortensen model has attracted a growing interest in the empirical literature. The various specifications of the model have been applied to data from the Netherlands (Van den Berg and Ridder, 1998), the US (Bowlus, 1997, and Bowlus et al., 2001), Denmark (Bunzel et al., 2001) and France (Bontemps et al., 2000). Bunzel et al. (2001) compare different versions of the model using Danish data. This study provides a similar exercise in the context of the Finnish labour market.

We estimate different variants of the Burdett-Mortensen model using maximum likelihood from a sample of workers who entered unemployment in Finland during 1992. In the analysis we distinguish between separate segments of the labour market by stratifying the data according to education, sex and age. The estimation results are discussed and compared across the model specifications as well as across the worker groups.² We begin with the homogeneous version of the model. This simplest version of the model gives a poor fit to the data because the shape of the theoretical wage distribution is at odds with the observed distribution. We

proceed by estimating an augmented specification of the homogenous model by assuming that the wage data are subject to measurement error (e.g. Christensen and Kiefer, 1994a). This is followed by analysis of the theoretical extensions of the model in which firms differ in their unobserved labour productivity. We consider specifications with a finite number of firm types (Bowlus et al., 1995, 2001, and Bowlus, 1997) as well as continuous productivity heterogeneity (Van den Berg and Van Vuuren, 2000, and Bontemps et al., 1999, 2000). Although the fundamental structure of the model remains unchanged, the source and interpretation of observed wage differentials depend upon the underlying assumptions on measurement error and productivity heterogeneity.

The results of the paper are useful in a variety of ways. First, a comparison of the fit of different extensions is informative on what aspects of the theory and empirical specification are likely to be important to obtain an acceptable fit to the Finnish data. Second, by investigating variation in the parameter estimates across the model specifications, we can test how robust the estimates of the fundamental parameters are with respect to different ways of deviating from the homogenous version of the model. Third, our results are informative about the relationship between unemployment durations, job durations and wages in the Finnish labour market.

The rest of the paper is organized as follows. In the next section the concepts of the equilibrium search theory are discussed. Section 3 introduces the data and provides some summary statistics. Estimation results from various empirical specifications are reported in Section 4. The final section concludes.

2. Equilibrium search theory

In this section we briefly describe equilibrium search theory along the lines of Burdett and Mortensen (1998) and Bontemps et al. (2000). We begin with a pure search model in which all jobs are equally productive and all workers are identical. Then we introduce employer heterogeneity in the model but retain the assumption

¹ See Mortensen and Pissarides (1999), Van den Berg (1999), Rogerson et al. (2005), and Eckstein and Van den Berg (2007) for the surveys of search models.

² Unlike Bunzel et al. (2001), we include also a model specification with continuous productivity dispersion in our set of the equilibrium search models to be compared.

of homogeneity of the worker population throughout the paper.

2.1 Equilibrium search with identical agents

2.1.1 Worker behaviour

The supply side of the labour market is populated by a continuum of *ex ante* identical workers. Behaviour of workers is characterised with the standard job search model with search on the job. In particular workers are assumed to be risk-neutral agents who are maximising the expected present value of future income stream with infinite horizon. Workers in the labour market are either employed or unemployed. Each worker is facing a known distribution of wage offers F with associated jobs, from which he randomly samples wage offers both on and off the job. Wage offers arrive at the Poisson rate λ_0 when unemployed and at the Poisson rate λ_1 when employed. Unemployed workers search for an acceptable job and employed workers for a better job. Jobs are destroyed at the Poisson rate δ , in which case the worker who holds the job is laid off and becomes unemployed.

Jobs are identical apart from the wage associated with them. As a result, employed workers are willing to move into higher-paying jobs whenever the opportunity arises. The current wage thus serves as the reservation wage for employed workers. In the limiting case of zero discounting, the reservation wage of the unemployed worker can be expressed as

$$(1) \quad r = b + (\kappa_0 - \kappa_1) \int_r^h \frac{1 - F(z)}{1 + \kappa_1 [1 - F(z)]} dz,$$

where b is the value of non-market time, including unemployment benefits net of search costs, $\kappa_0 = \lambda_0/\delta$, $\kappa_1 = \lambda_1/\delta$ and h is the upper bound of the support of F (see Burdett and Mortensen, 1998). This equation defines the reservation wage r as a function of the structural parameters of the model.

From (1) one can see how the possibility of search on the job affects the optimal search strategy of an unemployed worker. If wage offers arrive more frequently for the unemployed

than for the employed ($\lambda_0 > \lambda_1$), the reservation wage r exceeds the value of non-market time b . In that case it is more rewarding to search while unemployed and the worker rejects wage offers in the interval (b, r) , even though this causes a utility loss in the short run. If the offers arrive at a higher rate when employed ($\lambda_0 < \lambda_1$), entering employment improves the subsequent chances of receiving higher offers. It follows that the worker is willing to accept also some jobs that pay less than b . When the arrival rate is independent of employment status ($\lambda_0 = \lambda_1$), the worker is indifferent between searching while employed and while unemployed. Any job that compensates for the foregone value of non-market time is acceptable in this case and thus $r = b$.

2.1.2 Firm behaviour

The demand side of the labour market consists of a continuum of *ex ante* identical firms. The firms are assumed to use only labour inputs in production. Each worker generates a flow of revenue p to his employer. We assume that p is independent of the size of the workforce and refer to p as the (labour) productivity of the firm. The firm sets its wage so as to maximise the expected steady-state profit flow taking the optimal search behaviour of workers and wages set by other firms as given. To attract workers the firm posts wage offers, among which workers randomly search using a uniform sampling scheme. Contrary to the competitive setting, the presence of search frictions in the labour market generates dynamic monopsony power for wage-setting firms. As workers cannot find a higher-paying job instantaneously, firms can offer wages strictly smaller than marginal labour productivity.

The expected profit flow of a firm paying wage w in a steady state is given by

$$(2) \quad \pi(p, w) = (p - w) l(w),$$

where $l(w)$ is the expected size of the workforce (associated with a given F). The firm would employ as many workers as possible to maximise its profit flow as long as $p > w$. Since the current wage serves as the reservation wage for employed workers, the number of workers availa-

ble to the firm in equilibrium increases with the wage offered. In other words, the labour supply curve the firm is facing with is upward-sloping. The firm takes the function l as given and offers a wage that maximises its expected steady-state profit flow. Obviously, a firm never offers a wage above p as the profits would be negative, nor it offers a wage less than r as such a wage would not attract any workers.

Equally productive firms must receive the same expected profit flow in equilibrium. This does not mean that wage offers need to be equal, however. A firm paying a higher wage makes a lower profit per worker but makes it up in volume as the higher wage attracts more workers from other firms and enables the firm to retain them for a longer time. It follows that some firms choose to offer low wages with a cost of high labour turnover, while others pay higher wages and experience lower labour turnover. Due to this trade-off between the wage offered and labour turnover, the same expected profit level can be attained by paying different wages.

2.1.3 Steady-state outcomes

Denote the fixed size of the labour force with m and the steady-state number of unemployed workers with u . In a steady state the flows into and out of unemployment are equal, so that $\delta(m-u) dt = \lambda_0 u dt$ for a short time interval dt . Thus the steady-state unemployment rate is

$$(3) \quad \frac{u}{m} = \frac{\delta}{\delta + \lambda_0} = \frac{1}{1 + \kappa_0},$$

Using an analogous argument we can derive the steady-state *earnings* distribution G , the cross-section wage distribution of currently employed workers, associated with a given wage offer distribution F . By equating worker flows into and out of jobs paying w or less, we find the following steady-state relationship between the wage offer and earnings distributions:

$$(4) \quad G(w) = \frac{F(w)}{\delta + \lambda_1 [1 - F(w)]} \cdot \frac{\lambda_0 u}{m - u} \\ = \frac{F(w)}{1 + \kappa_1 [1 - F(w)]}$$

for all w on the common support of F and G . Since workers tend to move up the wage range over time, the earnings distribution lies to the right of the wage offer distribution, or more formally, G first-order stochastically dominates F as $F(w) - G(w) \geq 0$ for all w and $\kappa_1 \geq 0$. The discrepancy between the earnings and wage offer distributions depends on κ_1 which is equal to the expected number of wage offers during a spell of *employment* (which may consist of several consecutive job spells) and can be thought of as a relative measure of competition among firms for workers.

By equating the inflow and outflow of workers to the firm offering wage w , we find that the expected size of the workforce in such a firm can be expressed as

$$(5) \quad l(w) = \frac{\lambda_0 u + \lambda_1 G(w) (m - u)}{\delta + \lambda_1 [1 - F(w)]} \\ = \frac{\kappa_0 (1 + \kappa_1) m}{(1 + \kappa_0) (1 + \kappa_1 [1 - F(w)])^2},$$

provided that the size of the firm population is normalized to one. Obviously l is increasing in w and continuous where F is continuous. Note that no assumptions on the shape of F have been made so far. Since workers quit and are hired and laid off at random intervals, the workforce of the firm is a random variable, varying around its expected value l over time. As a consequence, also the profit flow is a random variable.

Burdett and Mortensen (1998) prove that there exists a unique non-cooperative equilibrium which consists of a triple $(r, F, \bar{\pi})$, such that (i) r satisfies (1) given F and (ii) each w on the support of F maximises $\pi(p, w)$, yielding the expected steady-state profit flow equal to $\bar{\pi}$.³ Furthermore, they show that the equilibrium solutions for F and G are absolutely continuous with the common support $[r, h]$.

Since the expected profit flow is $\bar{\pi}$ across all firms in equilibrium, it holds in particular that $\pi(r) = \bar{\pi} = \pi(w)$ for all w on the support of F .

³ Trivial solutions are ruled out by making the natural assumptions that $\infty > p > b$ and $\infty > \kappa_i > 0$ for $i = 0, 1$.

Taking this together with (5) gives the equilibrium wage offer distribution⁴

$$(6) \quad F(w) = \frac{1 + \kappa_1}{\kappa_1} \left(1 - \sqrt{\frac{p-w}{p-r}} \right),$$

with the associated density

$$(7) \quad f(w) = \frac{1 + \kappa_1}{2\kappa_1 \sqrt{(p-r)(p-w)}}, \quad w \in [r, h].$$

Moreover, by substituting (6) into (1) and recognising that $F(h) = 1$, we can write the bounds of support of F as

$$(8) \quad r = \alpha b + (1 - \alpha) p,$$

$$(9) \quad h = \beta b + (1 - \beta) p,$$

where the weights are given by

$$(10) \quad \alpha = \frac{(1 + \kappa_1)^2}{(1 + \kappa_1)^2 + (\kappa_0 - \kappa_1) \kappa_1},$$

$$(11) \quad \beta = \frac{1}{(1 + \kappa_1)^2 + (\kappa_0 - \kappa_1) \kappa_1}.$$

In other words, both the support and functional form of the wage offer distribution depends only on the structural parameters of the model. The fact that h is a weighed average of b and p further implies that the highest wage offered in the market is strictly smaller than p .

The main outcome of equilibrium search theory with on-the-job search is that wages are dispersed in equilibrium even when all workers and firms are homogenous. When the arrival rates of job offers, λ_0 and λ_1 , tend to infinity, the equilibrium earnings distribution G converges to a mass point at p , and both the steady-state unemployment rate and the equilibrium profit rate tends to zero. Thus the competitive

solution emerges as a limiting case when search frictions disappear. As a second extreme, if only unemployed workers receive offers ($0 < \lambda_0 < \infty$ and $\lambda_1 = 0$), the firms cannot increase their workforce by offering higher wages. Thus all firms offer the same wage equal to r which in turn converges to b . In this case the equilibrium earnings distribution G limits to a mass point at b , and the Diamond's (1971) paradoxical monopsony solution emerges. Moreover, all employment would be uniformly distributed across the firms as $l(w) = m - u$ by virtue of (5) and (3).

Other strong predictions follow from the simple model outlined above. First, workers with longer employment history are predicted to be more likely to be located at the upper end of the wage distribution. This is because wage growth in the model results from job-to-job transitions. Second, the model implies a positive relationship between the size of workforce and the wage paid by the firm. Firms offering higher wages grow at a larger size because a higher wage attracts more workers to a firm from other firms and reduces the quit rate, $\lambda_1 [1 - F(w)]$. These results are driven by on-the-job search.

An interesting prediction of the model is that a change in the unemployment benefit b does not affect equilibrium unemployment as long as $b < p$. For example, an increase in b increases the reservation wage r by virtue of (1). However, to retain a positive workforce, firms offering wages below the new value of r must react by increasing their wage offers which in turn affects the wage offers of other firms. The net result is that the exit rate out of unemployment and thus the unemployment level remain unchanged. Using a similar reasoning one can see that a decrease in b does not affect unemployment either.

2.2 Employer heterogeneity

The model of the previous section makes several predictions which can be expected to be consistent with empirical data. However, a closer look at (7) reveals that the density of the wage offer distribution (and, consequently, that of the earnings distribution) is strictly increasing and convex on its whole support. This con-

⁴ The associated equilibrium solutions for the earnings distribution follows directly from the steady-state relationship outlined in (4).

tradicts with the shape of wage distributions usually observed in the data as empirical wage densities are typically unimodal and skewed with a long right tail. This calls to doubt whether the simple equilibrium search model with identical agents can provide an acceptable fit to the wage data. To make the model more realistic, we extend the basic model by allowing for employer heterogeneity. In the first case we introduce heterogeneity assuming a discrete distribution for productivity across firms along the lines of Mortensen (1990) and Burdett and Mortensen (1998). This is followed by an extension of Bontemps et al. (2000) which allows for continuous productivity dispersion.

2.2.1 A finite number of firm types

Assume that there are q types of firms which differ in their labour productivity such that $p_1 < p_2 < \dots < p_q$. Let γ_i be the fraction of firms with productivity p_i or less. Keeping all other aspects of the model unchanged, Mortensen (1990) and Burdett and Mortensen (1998) show that the equilibrium solution in this case results in a complete segmentation of the wage offer range among firm types. Optimal wage setting implies that the wage offered increases with productivity and that all firms of type i offer wages on the interval $[w_i, \bar{w}_i)$, where the bounds of intervals are such that $\bar{w}_i = w_{i+1}$ for $i = 1, 2, \dots, q-1$. In addition, the lowest wage offered is equal to the reservation wage, so that $w_1 = r$. We also define that $\bar{w}_q = h$ in order to be consistent with our previous notation.

In equilibrium all firms of given type must have the same expected profit flow, so that $(p_i - w)l(w) = (p_i - w_i)l(w_i)$, where $l(w)$ is as defined in (5), holds for all firms with productivity p_i offering a wage on the interval $[w_i, \bar{w}_i)$. This implies the following equilibrium distribution of wage offers:

$$(12) \quad F(w) = \frac{1 + \kappa_1}{\kappa_1} \left(1 - \frac{1 + \kappa_1 (1 - \gamma_{i-1})}{1 + \kappa_1} \cdot \sqrt{\frac{p_i - w}{p_i - w_i}} \right), \quad w \in [w_i, \bar{w}_i),$$

with the associated density

$$(13) \quad f(w) = \frac{1 + \kappa_1 (1 - \gamma_{i-1})}{2\kappa_1 \sqrt{(p_i - w)(p_i - w_i)}},$$

$$w \in [w_i, \bar{w}_i),$$

where $\gamma_i = F(\bar{w}_i)$, with the convention that $\gamma_0 = 0$. As shown in Mortensen (1990) and in Burdett and Mortensen (1998), an equilibrium is characterized by $(r, F, \bar{\pi}_1, \dots, \bar{\pi}_q)$, where (i) r is the common reservation wage satisfying (1), (ii) F is the wage offer distribution given in (12) and (iii) $\bar{\pi}_i = (p_i - w)l(w)$ is the expected steady-state profit flow of firms with productivity p_i offering wages on the interval $[w_i, \bar{w}_i)$, $i = 1, 2, \dots, q$. In general, when there are q types of firms, the resulting distribution of wage offers F is absolutely continuous with the support $[r, h]$ and has $q-1$ “kinks” corresponding to the wage cuts $(\bar{w}_1, \bar{w}_2, \dots, \bar{w}_{q-1})$.⁵

Recall from the previous section that the equilibrium wage densities are convex over their whole support when all jobs are equally productive. This property is at odds with the long flat right tail commonly observed in the wage data. In contrast, the model of this section with productivity heterogeneity implies that the wage density f is discontinuous at the wage cuts, between which it exhibits locally increasing patterns. This theoretical distribution can take the functional form able to mimic the shape observed in the data. Moreover, since more productive firms offer higher wages, they attract more workers, face lower quit rates and, consequently, are larger on average. High productivity firms make also more profit on average in equilibrium than less productive firms.

2.2.2 A continuum of firm types

Bontemps et al. (2000) (see also Bontemps et al., 1999, and Burdett and Mortensen, 1998)

⁵ Van den Berg (2003) points out that there may exist other equilibria characterized by a different reservation wage and a different number of active firm types. This possibility arises from the two-way relationship between the reservation wage of the unemployed and minimum productivity level in use. In other words, the reservation wage affects the minimum level of profitable productivity and vice versa.

propose an alternative extension of the Burdett-Mortensen model which allows for continuous productivity dispersion. For this specification we suppose that productivity p is continuously distributed across active firms according to the distribution function Γ , with support $[p, \bar{p}]$, where $p \geq r$. Bontemps et al. (2000) show that only one wage offer can be profit-maximizing for each firm and the optimal wage offer increases with productivity. It follows that there exists a direct map between the productivity distribution Γ and wage offer distribution F :

$$(14) \quad F(K(p)) = \Gamma(p),$$

where $K(p)$ denotes the wage offered by firms with productivity p and K is an increasing and continuous function. Stated differently, the fraction of offers no higher than $K(p)$ equals the fraction of firms with productivity p or less.

Given that K is strictly increasing in p , firms with the lowest productivity p must offer a wage equal to the reservation wage r and the highest wage in the market is offered by firms with the highest productivity \bar{p} . Since the wage offer is unique for any p in the interior of Γ , the optimal wage offer $w = K(p)$ solves the first-order condition $\partial\pi(p, w)/\partial w = 0$. For given F and p this condition writes as

$$(15) \quad 2\kappa_1 f(w)(p - w) - (1 + \kappa_1 [1 - F(w)]) = 0,$$

provided that $w = K(p) \geq r$. The first-order condition gives an implicit function of the wage offer of a firm with productivity p given κ_1 and the distribution of wages offered by other firms. As shown in Bontemps et al. (2002), this implies the following optimal wage policy for a firm with productivity p :

$$(16) \quad K(p) = p - (1 + \kappa_1 [1 - \Gamma(p)])^2 \int_r^p \frac{dz}{(1 + \kappa_1 [1 - \Gamma(z)])^2},$$

By substituting r from (1) into (16), we can express $K(p)$ in terms of the structural parameters $(b, \lambda_0, \lambda_1, \delta, p, \Gamma)$.

Given Γ an equilibrium is characterized by (r, F) , where (i) the common reservation wage r satisfies (1) and (ii) the wage offer distribution the workers face while searching F is given by (14) and (16).⁶ As before the equilibrium distribution of wage offers F is absolutely continuous over its support $[r, h]$. Neither K nor F has a closed-form expression in general. Despite this the model outlined above makes some restrictions on the shape of the wage distributions, excluding certain shapes for F and G . One of these restrictions says that for given κ_1 , $f(w)(1 + \kappa_1 [1 - F(w)])$ decreases over the whole support of F (i.e. the second-order condition of profit maximization holds everywhere). This prediction can be used as a specification test in the empirical analysis of the model. It imposes the restriction how steeply f can increase for given κ_1 . Where the density f is decreasing the condition is obviously met regardless of the value of κ_1 . For large κ_1 the condition allows f to increase quite steeply, but for small κ_1 the condition is violated even if f increases only slightly.

When productivity is dispersed across firms, the shape of F obviously depends on the shape of Γ . However, whereas $\kappa_1 > 0$ is necessary and sufficient for wage dispersion, it should be noted that productivity dispersion is neither necessary nor sufficient for wage dispersion. In the presence of productivity differentials the degree of wage dispersion may be decomposed into two parts: one which is due to search frictions analogously to the homogeneous case and another which results from variation in productivity across firms. In particular it can be shown that the equilibrium of the homogeneous version of the Burdett-Mortensen models emerges as a limiting case when the degree of productivity dispersion goes to zero (see Bontemps et al., 2000, for details).

⁶ To be specific, there can be a single equilibrium, multiple equilibria or equilibrium may not exist at all, depending on the values of structural parameters of the model. Bontemps et al. (2000) point out that it is in general hard to differentiate between alternative cases. In the text we arbitrarily assume that the unique equilibrium exists.

3. Data

3.1 Sample details

Christensen and Kiefer (1997) discuss data requirements for identification of the structural parameters of the Burdett-Mortensen model. They show that the model can be estimated from data on individual labour market histories where at least some of the workers are observed with both unemployment duration and job duration with the associated wage. Empirical analysis of this study is based on a sample of individuals drawn from the worker data of the Integrated Panel of Finnish Companies and Workers (the IP data).⁷ Underlying source of information on workers in the IP data is the Employment Statistics (ES) database of Statistics Finland. The ES database is a longitudinal database which combines information from over 20 administrative registers. Since 1987 the ES database has been updated regularly, and it covers effectively all people with a permanent residence in Finland.

Each individual in the ES database who holds a job at the last week of the year is associated to his or her employer with a company and establishment identifier. The worker panel of the IP data covers all people from the ES database with an identifier of the private-sector employer at least in one of the years between 1988 and 1996. As a result, the underlying worker panel covers practically *all* persons who have been employed in the private sector during the period 1988–1996 (at least at the end of one year). The total number of persons in the IP data is slightly below two million. For these people a set of variables, collected by combining the annual records of the ES database, is available over the period 1988–1996.

In this paper we focus on a subsample of the worker panel of the IP data. As a first step we select all individuals between the ages of 16 and 65 who entered unemployment during 1992. In choosing a sample of unemployed workers we follow the practice of Bowlus et al. (2001) and Bunzel et al. (2001). We exclude workers who

⁷ A detailed description of the IP data is given in Korkeamäki and Kyyrä (2000). More recently, the Finnish Longitudinal Employer-Employee data (FLEED) have replaced the IP data.

have been self-employed as well as those have been employed by the public sector or non-profit organization during the period 1990–1996. These groups are excluded as the underlying model does not describe their labour market experiences.

For all individuals selected in the sample, we record the duration of the unemployment spell (d) and information on whether unemployment ended because of finding a job ($c_d = 0$) or for some other reason ($c_d = 1$).⁸ Unemployment spells not followed by a job are treated as right-censored in the empirical analysis. This may occur due to a drop out of the labour force, participation in the active labour market programme or the spell continuing beyond the observation period. It should be stressed that we treat unemployment spells ended in a job replacement programme as right-censored as well. Thus we make a difference between finding a job from the open labour market and becoming employed by labour administrative measures.

For those workers who found a job, we further record the accepted wage (w) and the duration of the subsequent job spell (j) along with the reason for termination. The wage rate is computed using information on annual earnings and the days worked. A job spell may end in a layoff ($a = 0$, $c_j = 0$), a quit ($a = 1$, $c_j = 0$) or be right-censored ($c_j = 1$). Job spells followed by unemployment are classified to be ended in a layoff, whereas job spells consecutively followed by another job spell with a new employer are interpreted to be ended in a quit for a better job. We identify changes in employer by comparing establishment identifiers attached to workers on the basis of the employer.⁹ Job spells terminated due to a drop out of the labour force and those continuing beyond the observation period are treated as right-censored. All durations are measured in months, and wages are

⁸ If the worker has several unemployment spells started in 1992, we choose the first one.

⁹ The establishment identifier is available only at the last week of each year, and hence we do not observe the employer for jobs that started and ended within the calendar year. Subsequent job spells result in a quit only if we actually observe the change in the establishment identifier. Hence, we may be underestimating the number of quits to some extent.

converted into monthly rates to match the duration measures.

Recall that our theoretical model is concerned with the population of homogeneous workers. While all workers are different in practice, we cannot allow the parameters to be different for each individual as the model will be of no use at all in such a case. Instead we assume that the labour market consists of a large number of segments, each of which forms a single market of its own. These segments are assumed to differ from each others according to observed characteristics of workers. To deal with this kind of heterogeneity, the model can be applied separately to each group of workers, allowing for all parameters to vary freely across the groups. This approach corresponds to controlling observed heterogeneity with a complete set of discrete regressors (Christensen and Kiefer, 1994a). To pursue this approach, we stratify the data by education, sex and age. We categorize education as follows: lower vocational education or less (11 years or less), upper vocational education (12–13 years), lower university (a Bachelor degree or the lowest level of university education, 13–15 years), and upper university (a Master degree or higher, 16 years or more). The age groups considered are: 16–21 years, 22–30 years, 31–50 years, and 51–65 years. As only few workers with lower or upper university education are aged below 22 or above 50, we combine the two lowest age groups as well as the two highest age groups for these education groups.

It is worth noting that the segmentation assumption does not preclude firms from employing different types of workers provided that the production function is additive in different types of labour inputs (e.g. Van den Berg, 1999). In this case, there may be several types of jobs with possibly different productivity levels within a firm but only a particular type of workers are suitable for each type of job. Since workers of different type do not compete for the same jobs, wage offers for workers occupying a particular labour market segment are independent of labour market conditions in other segments. This is an unrealistic assumption since heterogeneous workers (say, women and men with the same education) compete partly for the same

jobs in the real-world labour market. One may try to relax this restriction by introducing a more realistic production technology but it will probably make the model completely intractable. This unappealing implication of the segmentation assumption should be kept in mind when interpreting our empirical results.

As some of the estimation procedures used are sensitive to outliers in the wage data, some concern needs to be taken with our wage measures. Since the monthly wages are computed from annual earnings without information on hours worked, the wage data can be expected to contain some measurement error. To deal with outliers in the wage data, we first require that all wages must be at least 80% of the lowest salary grade of the central government, after which we trim the lowest and highest 3% of wage observations in each subgroup.

Finally, the maximum size of the estimation sample is restricted to 3,000 observations. This is because the computational burden of the estimation method for the model with a discrete distribution of productivity increases rapidly with the sample size and the number of firm types. Thus, we have drawn random subsamples of 3,000 workers (after trimming the wage data) from the underlying groups, defined by education, sex and age, which exceed this size threshold.

3.2 Descriptive statistics

It is worth emphasizing that the period under investigation is exceptional one. An overheating period of the Finnish economy in the last years of the 1980s was followed by a deep recession in the early 1990s. The annual change in the GDP was negative during the period 1991–1993, and in the worst year, 1991, the GDP decreased by over 7%. According to the Labour Force Survey, the unemployment rate rose from 3.2% in 1990 to over 16% in 1993, remaining at the level beyond 14.5% until 1996. The labour market experiences of the sampled workers thus took place in a period of record high and stable unemployment. We should keep this in mind when interpreting the results.

Table 1 gives some descriptive statistics for the worker groups to be analysed. The aggrega-

Table 1: Summary statistics

	N	\bar{w}	w_{\min}	w_{\max}	\bar{d}	c_d	\bar{j}	c_j	a
<hr/>									
Lower vocational and less									
Men, 16-21	17,051	7,819	4,625	16,713	7.56	.78	12.54	.32	.20
Men, 22-30	40,145	9,526	4,845	23,464	10.49	.58	12.28	.20	.22
Men, 31-50	73,142	10,483	5,067	26,669	11.58	.57	12.52	.22	.20
Men, 51-65	20,129	10,320	4,940	27,002	16.43	.73	9.94	.20	.19
Women, 16-21	7,604	7,071	4,546	18,325	7.07	.80	12.24	.33	.17
Women, 22-30	13,721	7,638	4,607	18,858	10.06	.71	14.63	.31	.22
Women, 31-50	31,212	7,932	4,685	18,172	11.94	.66	16.71	.29	.14
Women, 51-65	13,724	7,871	4,647	17,741	19.17	.79	13.95	.22	.10
<hr/>									
Upper vocational									
Men, 16-21	9,544	7,843	4,659	15,324	5.83	.78	11.91	.41	.27
Men, 22-30	11,233	9,298	4,834	22,003	8.81	.57	14.92	.28	.27
Men, 31-50	8,993	11,742	5,278	28,770	11.22	.60	17.82	.31	.22
Men, 51-65	1,811	13,645	5,724	27,963	18.08	.79	14.02	.30	.12
Women, 16-21	6,769	6,919	4,535	14,799	5.41	.73	11.45	.37	.25
Women, 22-30	11,002	7,734	4,680	16,232	7.51	.60	15.72	.31	.26
Women, 31-50	6,273	9,298	4,837	22,809	10.35	.65	17.60	.32	.23
Women, 51-65	695	10,284	5,028	23,591	18.98	.82	13.98	.26	.07
<hr/>									
Lower university									
Men, 16-30	5,042	10,531	5,185	23,859	7.36	.59	17.77	.35	.28
Men, 31-65	4,495	13,755	5,338	37,047	9.28	.54	18.88	.34	.24
Women, 16-30	1,831	8,541	4,725	16,842	6.55	.56	17.31	.36	.32
Women, 31-65	1,512	11,241	4,938	28,597	10.83	.61	17.97	.36	.25
<hr/>									
Upper university									
Men, 16-30	502	12,883	5,636	27,255	7.21	.47	20.87	.35	.55
Men, 31-65	1,321	17,874	5,667	53,488	12.12	.56	22.18	.38	.32
Women, 16-30	526	11,756	5,683	34,397	6.33	.51	18.09	.40	.38
Women, 31-65	543	12,887	4,831	34,092	10.32	.59	16.64	.34	.32

Notes: N is the number of observations in the underlying population from which the estimation sample was drawn. \bar{w} is the average accepted wage (FIM) in the estimation sample. w_{\min} and w_{\max} are the minimum and maximum of observed wages (FIM) in the estimation sample respectively. \bar{d} and \bar{j} are the average durations of unemployment and job spells respectively. c_d is the share of censored unemployment spells, and c_j is the share of censored job spells in the estimation sample. a is the share of uncensored job spells ending in a quit for a better job.

gate number of observations in the underlying group (N) is given in the first column of the table. Where N exceeds 3,000, all sample statistics are computed from a random subsample of 3,000 workers, describing the sample to be used in the estimations. It appears that the size of the underlying worker group is much lower for highly educated groups. This does not reflect only the education structure of the labour force but also a lower incidence of unemployment among more educated workers. The fact that the period under investigation is characterized by high unemployment levels is reflected to the figures in the table. Unemployment durations are relatively long with a high rate of censoring,

and most of subsequent job spells ended in a layoff.

Unemployment duration increases with age and is exceptionally high among low educated workers aged over 50. There are no clear differences in the average duration of unemployment by sex. The rate of censoring in the unemployment data is found to be very high. It is also worth emphasizing that the average duration of censored spells is over two times higher than that of uncensored spells (not shown in the table). This is because long-duration spells of unemployment are often terminated by labour administrative measures. This explains partly the higher censoring rates for the groups with

the longest unemployment durations. Young job seekers are often regarded as a special target group of the labour administrative measures, resulting in a relatively high censoring rate for workers aged under 22. There are no large differences in job duration across age groups. Highly educated workers experience slightly longer job spells and are more likely to quit for a better job. Compared to unemployment spells, job spells are longer on average and the rate of censoring in the job duration data is much lower.

Wages increase with age at least up until the interval 31–50 years of age. There are no clear wage differentials between workers at the two lowest levels of education, whereas higher education yields slightly higher return. Given education and age women receive uniformly lower wages than men do. Despite the trimming procedure there are still wage observations which are relatively low compared to minimum requirements, reflecting some measurement problems in the wage data. Empirical wage densities for each worker group are shown in Figures 1 to 3 in the Appendix, where the thick solid lines represent the kernel density estimates obtained using Gaussian kernels with the bandwidth chosen by a rule of thumb. Other lines depict the predicted densities obtained from the different specifications of the equilibrium search model and they will be discussed later on. The empirical wage densities are generally unimodal and skewed with a long right tail.

4. Econometric analysis

We have derived the explicit solutions for the equilibrium wage offer distribution with and without employer heterogeneity. From the assumptions underlying the theoretical model it is straightforward to derive distributions for unemployment and job durations as well. Knowledge about these distributions allows us to write down the likelihood function for the various specifications of the model. In the next section we derive the general form of the likelihood function without specifying the functional form for the wage offer distribution. In the subsequent sections we report the empirical results.

Some technical details on the estimation procedures are given in the Appendix.

4.1 The log-likelihood function

The structural parameters of interest are $(\lambda_0, \lambda_1, \delta, r)$ with a scalar p for the homogeneous model, the set (p_1, p_2, \dots, p_q) of productivity terms for the model with a discrete distribution of productivity, and Γ for the model with a continuous distribution of productivity. With the corresponding parameter estimates in hand, we can obtain an estimate for b using (1).¹⁰ Since our data were drawn from the inflow of unemployment, we observe a spell of unemployment along with the post-unemployment destination for each individual in the data. For those whose unemployment ended in a new job we further observe the wage rate accepted as well as the duration of the subsequent job along with the reason for termination. Since we do not have complete information on all observations, the possibility of censored observations on unemployment and job durations is explicitly accounted for using censoring indicators c_d and c_j .

The likelihood contribution from an individual who is unemployed for d periods, accepts then a job with an associated wage w , keeps that job for j periods until he gets laid off ($a = 0$) or finds another job ($a = 1$) has a general form

$$(17) \quad \ell = \varphi(d) [f(w)\phi(j, a|w)]^{1-c_d},$$

where φ is the density function of unemployment duration, f is the density of the wage offer distribution and ϕ is the density function of job duration and destination conditional on the accepted wage w . The censoring indicator for unemployment duration c_d takes a value of zero if the unemployment spell is followed by a new job, and a value of one otherwise.

Since job offers arrive at the Poisson rate λ_0 and all offers are acceptable to the unemployed

¹⁰ It can be argued that b is the 'deep' structural parameter of the model rather than r , which follows implicitly from the optimal search strategy of unemployed workers. However, this distinction does not make any difference in practice due to the one-to-one relationship between b and r outlined in (1).

in equilibrium, unemployment duration d is exponentially distributed with intensity parameter λ_0 , so that

$$(18) \quad \varphi(d) = \lambda_0^{1-c_d} e^{-\lambda_0 d}.$$

As workers search randomly among employers using a uniform sampling scheme, the wage offers are random draws from the equilibrium wage offer distribution F . To derive the conditional distribution of job duration j and destination a , we can use the standard competing risks framework for exponential duration models. Recall that layoffs occur at the Poisson rate δ and alternative offers arrive at the Poisson rate λ_1 . Since only wage offers exceeding the current wage will be accepted, the actual quit rate is $\lambda_1(1-F(w))$, the probability of receiving an offer times the probability that the received offer is acceptable given the current wage w . Conditional on the current wage w , the job duration j has an exponential distribution with intensity parameter $\delta + \lambda_1(1-F(w))$. Exit from this job into unemployment occurs with probability $\delta/[\delta + \lambda_1(1-F(w))]$ and exit into a higher-paying job with probability $\lambda_1(1-F(w))/[\delta + \lambda_1(1-F(w))]$. Putting these together yields

$$(19) \quad \phi(j, a | w) = [(1-a)\delta + a\lambda_1(1-F(w))]^{1-c_j} e^{-(\delta+\lambda_1[1-F(w)])j}.$$

Substituting (18) and (19) into (17) and taking logarithm gives the individual contribution to the log-likelihood function:

$$(20) \quad \log \ell = (1-c_d)(1-c_j) \ln [(1-a)\delta + a\lambda_1(1-F(w))] - \lambda_0 d + (1-c_j) (\ln \lambda_0 + \ln f(w) - (\delta + \lambda_1[1-F(w)])j).$$

Estimations of different specifications of the Burdett-Mortensen model will all be based on (20), the only difference between the specifications being the functional form assumed for F and f .¹¹ Recall that the shape of the equilibrium

wage offer distribution generally does not depend on λ_0 . This observation taken together with (20) suggests that λ_0 is identified from the unemployment duration data only. It follows that the estimator of λ_0 is stochastically independent of all other parameters of the model, being robust with respect to different model specifications. Despite the apparent simplicity of the log-likelihood function (20), its maximization is a difficult task, and some nonstandard procedures are called for. We discuss these difficulties in the Appendix.

4.2 Identical workers and firms

We begin our empirical analysis with the simplest specification of the model with homogeneous workers and firms. Equilibrium solutions for F and f are given by (6) and (7) respectively. For the reasons given in the Appendix, we estimate r and h using the sample minimum and maximum respectively. These order statistics estimates of (r, h) for each group can be found from Table 1, where they correspond to the minimum and maximum accepted wage (i.e. $\hat{r} = w_{\min}$ and $\hat{h} = w_{\max}$). Estimates of other parameters of the model are presented in Table 2.

It is found that λ_0 is uniformly higher than λ_1 , suggesting that wage offers arrive more frequently when unemployed than when employed. This corresponds to the case where the reservation wage r exceeds the value of non-market time b . Moreover, as δ is uniformly higher than λ_1 , jobs are more likely to end in a layoff than in a quit for a better job. Ignoring workers aged below 22, λ_0 decreases with age, being exceptionally low for less educated workers aged over 50. Moreover, λ_0 increases with education, though not uniformly. In contrast, λ_1 does not exhibit any clear patterns with respect to education nor with respect to age. Less educated women aged over 50 have the lowest chances of

quit or layoff, i.e. observations on a are not crucial for identification. The separate identification of λ_1 and δ even without knowledge of f follows from the fact that the conditional job hazard decreases with w . A higher wage does not affect the layoff rate but implies a lower quit rate and the extent of this effect depends on λ_1 .

¹¹ Christensen and Kiefer (1997) show that identification of all structural parameters of the model does not necessarily require information on whether the job spell ends in a

Table 2: Estimation results with identical agents

	λ_0	λ_1	δ	p	b
<u>Lower vocational and less</u>					
Men, 16-21	.0285 (.0011)	.0099 (.0011)	.0470 (.0024)	42,697 (3,271)	2,461 (229)
Men, 22-30	.0399 (.0011)	.0130 (.0009)	.0547 (.0018)	58,533 (2,844)	754 (279)
Men, 31-50	.0375 (.0010)	.0117 (.0008)	.0531 (.0018)	70,946 (3,629)	332 (312)
Men, 51-65	.0166 (.0006)	.0138 (.0013)	.0701 (.0029)	78,136 (5,373)	4,532 (205)
Women, 16-21	.0282 (.0012)	.0070 (.0009)	.0486 (.0025)	62,871 (6,164)	1,753 (257)
Women, 22-30	.0292 (.0010)	.0082 (.0007)	.0401 (.0017)	50,235 (3,240)	1,228 (264)
Women, 31-50	.0286 (.0009)	.0051 (.0005)	.0384 (.0015)	65,841 (5,314)	811 (238)
Women, 51-65	.0109 (.0004)	.0054 (.0008)	.0519 (.0024)	78,050 (9,152)	3,988 (114)
<u>Upper vocational</u>					
Men, 16-21	.0385 (.0015)	.0128 (.0013)	.0405 (.0022)	29,884 (1,751)	1,740 (298)
Men, 22-30	.0494 (.0014)	.0118 (.0008)	.0390 (.0014)	46,673 (1,950)	-2,358 (416)
Men, 31-50	.0353 (.0010)	.0080 (.0006)	.0327 (.0012)	71,490 (3,825)	-3,464 (535)
Men, 51-65	.0116 (.0006)	.0071 (.0013)	.0454 (.0029)	94,126 (13,130)	4,687 (342)
Women, 16-21	.0505 (.0018)	.0113 (.0010)	.0458 (.0022)	33,238 (1,982)	635 (291)
Women, 22-30	.0534 (.0015)	.0106 (.0007)	.0357 (.0013)	33,173 (1,426)	-1,366 (346)
Women, 31-50	.0335 (.0010)	.0079 (.0006)	.0324 (.0013)	55,604 (3,089)	-1,489 (419)
Women, 51-65	.0095 (.0009)	.0035 (.0014)	.0506 (.0055)	152,083 (55,257)	3,970 (337)
<u>Lower university</u>					
Men, 16-30	.0562 (.0016)	.0096 (.0007)	.0296 (.0011)	48,443 (2,054)	-7,453 (697)
Men, 31-65	.0495 (.0013)	.0075 (.0005)	.0292 (.0010)	91,523 (4,474)	-14,893 (1,031)
Women, 16-30	.0675 (.0025)	.0119 (.0010)	.0284 (.0014)	28,814 (1,272)	-5,081 (705)
Women, 31-65	.0357 (.0015)	.0080 (.0009)	.0293 (.0017)	66,729 (4,988)	-4,881 (895)
<u>Upper university</u>					
Men, 16-30	.0735 (.0047)	.0180 (.0020)	.0184 (.0018)	34,727 (1,350)	-16,266 (2,945)
Men, 31-65	.0363 (.0016)	.0078 (.0008)	.0218 (.0013)	110,639 (7,135)	-20,876 (2,368)
Women, 16-30	.0770 (.0050)	.0102 (.0015)	.0249 (.0023)	63,666 (5,355)	-26,298 (3,901)
Women, 31-65	.0395 (.0027)	.0123 (.0019)	.0310 (.0029)	64,873 (6,069)	-5,850 (1,781)

Notes: Standard errors are in parentheses.

finding a job when unemployed, but women with university education tend to receive more offers than their male counterparts when unemployed. Layoff rate δ decreases with education but there is little difference by sex. Workers aged below 22 and those aged over 50 are more likely to be laid off than other workers.

Productivity p increases with age and is often higher for men than for women with some exceptions. Young workers with upper vocational education are found to be less productive than their less educated co-workers. Otherwise productivity differentials across education groups do not exhibit very clear insight. The value of non-market time b appears to be positive among workers with the lowest level of education. For more educated groups b is typically negative, being more negative for higher education levels. Negativity of b reflects the need to interpret this

parameter not only a function of unemployment benefits but also of search costs and perhaps even of the disutility of unemployment (Bunzel et al., 2001). There is also a tendency for b to be lower for women than men among less educated workers aged below 22 and over 50, while the reverse is true among workers between 22 and 50 years old.

Equilibrium unemployment rates, as implied by the estimates of λ_0 and δ , are record high, being over 50% for some groups (not reported in the table). It is obvious that our sample drawn from the inflow of unemployment is not a representative sample of the labour force but workers with poor employment prospects are likely to be over represented. In particular, the layoff rates are expected to be much higher in our sample than among the employed population on average, which leads to the overwhelmingly

large unemployment rates. Although a flow sample is a natural choice to explain differences in post-unemployment wages, it is obvious that our results do not describe the labour market as a whole very well.

Overall all structural parameters of the model are estimated accurately and their estimates allow for a meaningful economic interpretation. The fit to the wage data is less satisfactory, however. This is illustrated in Figures 1 to 3 in the Appendix where empirical wage distributions and predicted wage offer distributions obtained from the different specifications of the model are shown. The predicted theoretical density for the pure homogeneity model is computed by inserting the parameter estimates into (7). While the empirical (kernel) densities are unimodal and skewed with a long right tail, the equilibrium search model with identical agents restricts the predicted densities to be increasing and convex over the whole support. Such a shape is obviously not supported by the data. The predicted densities are flat over their whole support, leading to a poor fit to the wage data in all worker groups.

Recall that our data do not contain information on working time. This with some other inaccuracies in the available data suggests that our wage variables are subjected to measurement error. The estimates of the structural parameters of the model can be expected to be affected by measurement errors. The order statistics estimators of (r, h) are clearly sensitive to measurement error. The dependence of (r, h) on other parameters of the model implies that the maximum likelihood estimates of frictional parameters (λ_1, δ) are also affected by measurement error in wage data (Van den Berg and Ridder, 1993). Taking the possibility of measurement error in wages explicitly into account may hence improve the performance of maximum likelihood estimation.

Following Christensen and Kiefer (1994a) and Bunzel et al. (2001), we assume that the wage data are subject to a multiplicative error term with a Pearson Type V distribution with unit mean. Estimation results for the homogeneous model with measurement error in wages are reported in Table 3. While the underlying theoretical model remains unchanged, the estima-

tion procedure deals with a more complex measurement process now (see the Appendix for details). In contrast to the previous case, we now estimate r and h along with other parameters using maximum likelihood.

The results are missing for four groups of low-educated women as the estimates of r and h converged to the same value in their cases. Compared to the previous results, λ_1 is now uniformly much higher and δ uniformly slightly lower, while λ_0 is of course not affected by the introduction of measurement errors. Among workers with an upper university degree λ_1 now exceeds δ , while the reverse still holds for other groups. Moreover p is uniformly lower and b uniformly higher than previously. The presence of measurement errors in the wage data suggests that a range of wage offers is narrower than previously, so less variation in p and b is required to explain the observations in the data. The estimates of r and h are generally very close to each other, resulting in a small difference between p and b . As another implication, dispersion in the sequence of wages that the worker can receive over time is predicted to be quite narrow.

The predicted densities for *observed* wages obtained from the measurement error model are shown in Figures 1 to 3 in the Appendix. It is evident that the measurement error specification results in a rather good (statistical) fit to the wage data. In particular, both tails are captured quite nicely and the mode point is very close. As a consequence of the improved fit to the wage data, the estimates of frictional parameters are likely to be more appropriate here. On the other hand, the finding that measurement errors account for such a large part of the observed wage variation can be viewed as a failure of the theoretical model as it implies that the theory is unable to explain wage dispersion within the labour market segments. In other words, the results suggest that search frictions do not play an important role in explaining wage dispersion. But this is a too hasty conclusion since workers and jobs are likely to be different within the narrowly defined labour market segments as well.

Table 3: Estimation results with identical agents and measurement error in wages

	λ_0	λ_1	δ	p	b	σ
Low voc. & less						
Men, 16-21	.0285 (.0011)	.0236 (.0027)	.0437 (.0023)	9,278 (920)	7,052 (454)	.3012 (.0114)
Men, 22-30	.0399 (.0011)	.0317 (.0024)	.0508 (.0018)	11,915 (848)	8,121 (475)	.3417 (.0108)
Men, 31-50	.0375 (.0010)	.0291 (.0022)	.0495 (.0017)	11,329 (996)	9,995 (545)	.3739 (.0100)
Men, 51-65	.0166 (.0006)	.0339 (.0033)	.0658 (.0028)	14,336 (1,499)	8,855 (517)	.3743 (.0173)
Women, 16-21	–	–	–	–	–	–
Women, 22-30	–	–	–	–	–	–
Women, 31-50	.0286 (.0009)	.0122 (.0013)	.0366 (.0015)	8,013 (1,022)	7,849 (420)	.2700 (.0069)
Women, 51-65	.0109 (.0004)	.0124 (.0018)	.0502 (.0024)	11,457 (1,793)	7,026 (391)	.2883 (.0123)
Upper vocational						
Men, 16-21	.0385 (.0015)	.0300 (.0031)	.0363 (.0021)	10,900 (660)	5,611 (409)	.2294 (.0183)
Men, 22-30	.0494 (.0014)	.0286 (.0020)	.0353 (.0013)	10,682 (635)	8,056 (540)	.3320 (.0091)
Men, 31-50	.0353 (.0010)	.0189 (.0015)	.0301 (.0012)	16,688 (1,268)	8,244 (889)	.3549 (.0170)
Men, 51-65	.0116 (.0006)	.0133 (.0025)	.0439 (.0029)	34,816 (4,476)	8,013 (652)	.2657 (.0327)
Women, 16-21	.0505 (.0018)	.0301 (.0029)	.0412 (.0021)	7,436 (479)	6,464 (364)	.2594 (.0076)
Women, 22-30	–	–	–	–	–	–
Women, 31-50	–	–	–	–	–	–
Women, 51-65	.0095 (.0009)	.0075 (.0031)	.0496 (.0054)	40,043 (13,510)	5,802 (674)	.2217 (.0516)
Lower university						
Men, 16-30	.0562 (.0016)	.0228 (.0016)	.0264 (.0011)	15,660 (612)	4,556 (607)	.2506 (.0151)
Men, 31-65	.0495 (.0013)	.0183 (.0013)	.0265 (.0010)	22,672 (1,127)	4,951 (1,007)	.3550 (.0201)
Women, 16-30	.0675 (.0025)	.0267 (.0024)	.0251 (.0014)	9,190 (570)	7,545 (827)	.3018 (.0103)
Women, 31-65	.0357 (.0015)	.0195 (.0022)	.0265 (.0016)	13,554 (1,640)	9,305 (1,332)	.3832 (.0188)
Upper university						
Men, 16-30	.0735 (.0047)	.0431 (.0047)	.0140 (.0016)	17,131 (528)	2,864 (1,355)	.2071 (.0210)
Men, 31-65	.0363 (.0016)	.0197 (.0020)	.0189 (.0012)	27,896 (1,760)	6,656 (1,901)	.4462 (.0388)
Women, 16-30	.0770 (.0050)	.0294 (.0044)	.0205 (.0022)	12,490 (1,350)	9,911 (2,646)	.3409 (.0206)
Women, 31-65	.0395 (.0027)	.0305 (.0050)	.0271 (.0028)	19,978 (1,861)	6,219 (1,642)	.4045 (.0604)

Notes: Standard errors are in parentheses

4.3 Discrete productivity dispersion

Next we consider the extended model with a discrete distribution of productivity. Here we do not allow for measurement error in wages, so comparisons with the homogeneous version of the model can be done in a straightforward manner. The individual contribution to the log-likelihood function is still given by (20), the only difference compared to the homogeneous case without measurement errors being that F and f are now given by (12) and (13) respectively. Bowlus et al. (1995) develop an estimation method which is able to deal with the ill-behaved likelihood function of this extension (see the Appendix for details). Once the other parameters of the model are estimated, an estimate of b can be obtained using

$$(21) \quad b = r - \frac{\lambda_0 - \lambda_1}{\lambda_1} \sum_{i=1}^q (p_i - \underline{w}_i) \cdot \left(1 - \mu_i^2 - \frac{2\delta(1 - \mu_i)}{\delta + \lambda_1(1 - \gamma_{i-1})} \right),$$

which follows from the substitution of (12) into (1).

Order statistics estimates for $(r, h) = (w_1, \bar{w}_q)$ can be found from Table 1 as previously. Other parameter estimates are shown in Tables 4 to 6. Estimates of the frictional parameters (λ_1, δ) are generally very close to the estimates obtained from the homogeneous model with measurement error in wages. Compared to the corresponding estimates from the measurement error specification, there is a tendency for λ_1 to be slightly smaller while δ does not exhibit any

Table 4: Estimation results with discrete productivity dispersion

	λ_0	λ_1	δ	q	\bar{p}	b
<hr/>						
Lower vocational and less						
Men, 16-21	.0285 (.0011)	.0219 (.0023)	.0444 (.0023)	5	21,539	4,266
Men, 22-30	.0399 (.0011)	.0288 (.0019)	.0507 (.0018)	6	28,438	4,049
Men, 31-50	.0375 (.0010)	.0234 (.0016)	.0499 (.0017)	5	36,920	3,761
Men, 51-65	.0166 (.0006)	.0298 (.0026)	.0662 (.0028)	6	37,507	5,793
Women, 16-21	.0282 (.0012)	.0166 (.0020)	.0464 (.0025)	4	27,227	4,069
Women, 22-30	.0292 (.0010)	.0193 (.0017)	.0370 (.0017)	6	22,808	3,971
Women, 31-50	.0286 (.0009)	.0123 (.0012)	.0365 (.0015)	6	29,556	3,465
Women, 51-65	.0109 (.0004)	.0122 (.0017)	.0503 (.0024)	4	36,217	4,715
<hr/>						
Upper vocational						
Men, 16-21	.0358 (.0015)	.0254 (.0024)	.0373 (.0021)	4	17,497	3,827
Men, 22-30	.0494 (.0014)	.0254 (.0016)	.0353 (.0013)	4	23,945	2,491
Men, 31-50	.0353 (.0010)	.0181 (.0013)	.0301 (.0012)	5	35,160	2,320
Men, 51-65	.0116 (.0006)	.0133 (.0023)	.0439 (.0029)	3	53,937	5,972
Women, 16-21	.0505 (.0018)	.0276 (.0024)	.0420 (.0021)	5	15,697	3,636
Women, 22-30	.0534 (.0015)	.0230 (.0015)	.0323 (.0013)	6	17,499	2,547
Women, 31-50	.0335 (.0010)	.0174 (.0013)	.0298 (.0013)	4	27,277	2,913
Women, 51-65	.0095 (.0009)	.0072 (.0029)	.0498 (.0054)	2	77,089	4,798
<hr/>						
Lower university						
Men, 16-30	.0562 (.0016)	.0213 (.0014)	.0267 (.0011)	5	25,167	275
Men, 31-65	.0495 (.0013)	.0180 (.0012)	.0266 (.0010)	6	41,202	-1,758
Women, 16-30	.0675 (.0025)	.0245 (.0019)	.0250 (.0014)	6	16,807	373
Women, 31-65	.0357 (.0015)	.0205 (.0021)	.0263 (.0016)	6	29,217	2,460
<hr/>						
Upper university						
Men, 16-30	.0735 (.0047)	.0410 (.0043)	.0138 (.0016)	5	19,919	-863
Men, 31-65	.0363 (.0015)	.0208 (.0020)	.0188 (.0012)	6	45,053	-584
Women, 16-30	.0770 (.0050)	.0257 (.0033)	.0206 (.0022)	5	26,853	-3,965
Women, 31-65	.0395 (.0027)	.0265 (.0039)	.0275 (.0028)	5	33,265	2,109

Notes: Standard errors are in parentheses. q is the number of firm types and $\bar{p} = \sum_{i=1}^q (\gamma_i - \gamma_{i-1}) p_i$ is the average productivity across firms.

systematic differences. Overall the differences in these estimates are so moderate that one can draw basically the same conclusions concerning the frictional parameters from this model and from the homogeneous model with the measurement error.

To get an idea of productivity differences, the average productivity across the firms is computed and shown in Table 4. Conditional on the education level, the average productivity \bar{p} tends to increase with age. Except for workers aged over 50, there is a tendency for \bar{p} to be lower for workers with upper vocational education than for those with lower vocational education. Overall these differences across the worker groups are in line with the findings from the homogeneous model, even though the absolute values of productivity estimates are quite different.

Namely, the average productivity estimates \bar{p} are approximately only half of the corresponding productivity estimates obtained from the homogeneity model without measurement error but are clearly higher than the estimates from the measurement error specification. Furthermore, it turns out that the value of non-market time b is typically positive, though there are few groups for which b takes a negative value. Differences in b with respect to education and age are similar to those observed in the case of the homogeneous model.

Estimates of productivity parameters, wage cuts and their weights are shown in Tables 5 and 6. Individual productivity values and wage cuts, say p_i and \bar{w}_i , exhibit increasing patterns with respect to age among groups with university education while the picture is less clear for less

Table 5: Estimation results with discrete productivity dispersion, continued

	p_1	p_2	p_3	p_4	p_5	p_6
<u>Lower vocational and less</u>						
Men, 16-21	13,643	30,392	40,487	44,072	108,892	
Men, 22-30	17,295	17,190	19,007	24,177	41,443	136,864
Men, 31-50	22,671	37,564	47,012	56,809	148,293	
Men, 51-65	20,879	26,750	35,892	40,360	62,652	211,517
Women, 16-21	11,969	33,784	114,231	679,097		
Women, 22-30	12,354	17,426	25,163	42,016	47,827	125,113
Women, 31-50	16,398	17,708	24,134	28,632	45,129	162,682
Women, 51-65	19,360	43,418	123,891	540,357		
<u>Upper vocational</u>						
Men, 16-21	13,375	19,258	26,360	62,415		
Men, 22-30	15,146	21,668	27,620	75,768		
Men, 31-50	22,338	27,545	30,704	41,592	100,363	
Men, 51-65	38,261	72,673	141,549			
Women, 16-21	9,757	15,158	29,653	39,350	87,954	
Women, 22-30	11,946	13,236	16,207	26,410	56,156	56,535
Women, 31-50	16,759	26,769	48,668	195,431		
Women, 51-65	49,328	152,647	721,354			
<u>Lower university</u>						
Men, 16-30	18,213	20,020	24,891	35,263	113,867	
Men, 31-65	25,616	45,985	48,888	83,992	156,640	446,879
Women, 16-30	12,812	13,507	14,208	17,024	22,032	41,729
Women, 31-65	17,605	25,184	30,545	44,957	59,843	127,858
<u>Upper university</u>						
Men, 16-30	16,589	16,960	19,525	21,083	40,111	
Men, 31-65	26,719	28,891	31,640	44,102	58,421	168,181
Women, 16-30	17,355	18,697	25,678	50,459	151,310	
Women, 31-65	22,495	26,773	34,878	40,743	95,614	

Notes: All p_i 's are statistically significant at the 5 per cent level.

educated workers. Of course, these kinds of comparisons are complicated by the fact that the number of firm types q varies across the groups. Given the shape of the empirical wage distribution, it is not very surprising that the bulk of firms is found to be low productivity ones. Firms with the lowest level of productivity represent over half of all firms in each submarket as $\gamma_1 > .5$ holds for all worker groups. Some productivity terms take very high values in some groups of workers, though their relative weights are very low (see the associated γ_i 's). It is worth noting that all wage differentials that cannot be explained by differences in the frictional parameters are attributed to productivity differences. Thus the productivity parameters may capture also other sources of wage dispersion than pure productivity differences (Bowlus, 1997).

In Figures 1 to 3 in the Appendix the estimated density functions of the model with discrete productivity dispersion are characterized by discontinuous jumps at the estimated wage cuts $(\bar{w}_1, \dots, \bar{w}_{q-1})$, between which the densities exhibit locally increasing patterns. It turns out that the model with discrete productivity dispersion is able to capture the shape of the wage distribution quite well but has some difficulties with the both tails of the distribution. In particular the estimated density has generally the left tail which is too fat compared to the observed wage distribution. As a result, the model has problems also to match the mode point accurately. These failures are not unique to the Finnish data but appear in the previous empirical applications of the same model as well (see Bowlus et al., 1995, Bowlus, 1997, Bunzel et

Table 6: Estimation results with discrete productivity dispersion, continued

	\bar{w}_1	\bar{w}_2	\bar{w}_3	\bar{w}_4	\bar{w}_5	γ_1	γ_2	γ_3	γ_4	γ_5
<u>Lower voc. & less</u>										
Men, 16-21	8,700	10,704	11,755	12,511		.7857	.8917	.9297	.9543	
Men, 22-30	9,768	10,008	10,273	11,889	15,874	.6143	.6493	.6807	.8053	.9420
Men, 31-50	12,458	14,803	15,684	16,329		.7477	.8620	.8933	.9110	
Men, 51-65	11,330	11,478	12,988	13,132	18,042	.7283	.7403	.8183	.8247	.9467
Women, 16-21	7,352	9,303	14,857			.8027	.9153	.9927		
Women, 22-30	7,913	8,150	10,445	11,352	12,211	.7080	.7357	.8880	.9173	.9410
Women, 31-50	8,270	8,647	9,378	9,629	12,102	.6623	.7290	.8063	.8270	.9383
Women, 51-65	8,849	11,177	13,265			.7947	.9433	.9823		
<u>Upper vocational</u>										
Men, 16-21	9,379	10,197	11,728			.7967	.8673	.9450		
Men, 22-30	9,563	10,242	14,532			.6310	.6810	.9067		
Men, 31-50	12,429	13,539	14,280	17,680		.6340	.7100	.7527	.8757	
Men, 51-65	16,607	19,303				.7914	.8767			
Women, 16-21	7,118	8,296	10,438	10,646		.7297	.8663	.9517	.9573	
Women, 22-30	8,074	8,349	9,603	11,238	11,580	.6483	.6957	.8377	.9157	.9213
Women, 31-50	10,213	11,963	18,128			.7013	.8103	.9770		
Women, 51-65	13,736	17,585				.8653	.9686			
<u>Lower university</u>										
Men, 16-30	12,405	12,811	13,030	17,481		.7490	.7897	.8030	.9563	
Men, 31-65	15,407	21,470	21,528	25,153	31,672	.7200	.9040	.9057	.9520	.9903
Women, 16-30	8,763	8,951	9,742	10,257	12,828	.5913	.6199	.7297	.7763	.9206
Women, 31-65	10,846	13,638	14,769	15,852	21,243	.6153	.7864	.8373	.8635	.9533
<u>Upper university</u>										
Men, 16-30	12,846	14,301	16,066	17,204		.5553	.7085	.8255	.8872	
Men, 31-65	16,284	17,287	19,185	23,020	29,622	.5637	.6181	.7062	.8022	.9103
Women, 16-30	13,065	13,086	16,651	22,223		.7102	.7122	.8796	.9592	
Women, 31-65	13,888	15,327	17,958	21,768		.6160	.6979	.7914	.9006	

al., 2001, and Bowlus et al., 2001). Additionally, there are difficulties in explaining wage observations at the upper end of the distribution which is often thin, covering wide ranges. Adding more firm types serves as a way of obtaining a more accurate fit to the right tail of the distribution. The estimation procedure aims to attach different firm types to each of these observations, leading to implausibly high productivity values sometimes. However, these high productivity values have only a minor overall effect as their weights are very low (in terms of the associated values of γ_i 's).

It should be stressed that the trimming procedure applied to the wage data is related to the number of heterogeneity terms needed to match the right hand tail of the wage distribution. Indeed by trimming a higher fraction of wage values from the upper end of the distribution leads to a smaller choice of q , with the highest pro-

ductivity parameters p_i being in a more reasonable range and the associated values of γ_i being well below one. Changing the upper value of trimming has of course a direct effect on the estimate of h . A brief sensitivity analysis done with the different trimming thresholds suggests that the other parameters and conclusions of interest are reasonably robust, however.

4.4 Continuous productivity dispersion

Next we turn our attention to the version of the model with a continuous productivity distribution. The individual log-likelihood contribution has the same general form as previously defined in (20) but the equilibrium distribution of wage offers is now given by $F(w) = \Gamma(K^{-1}(p))$. This is a highly nonlinear function of unknown parameters and does not have a closed-form expression in general, suggesting that the standard

maximum likelihood estimation could be very cumbersome. The first point to note is that the integrals within the expressions for F and f must be evaluated numerically in each iteration. An additional difficulty follows from the fact that the productivity distribution Γ is not generated by the model but it must be taken as exogenously given. This is not a problem as such but, as argued by Bontemps et al. (2000), the most well-known parametric specifications for Γ are unlikely to generate the wage offer distribution consistent with the shape usually observed in the wage data.

To deal with these difficulties Bontemps et al. (2000) propose a flexible estimation procedure which does not restrict Γ to belong in any parametric family (see the Appendix for details). In the first step the wage distribution is estimated using some nonparametric method, which does not impose any restrictions on the shape of equilibrium wage distributions. Conditional on the nonparametric estimates of F and f , maximum likelihood estimation in the second step relies only on assumptions about the behaviour of individual workers who are taking the wage offer distribution as given. In other words, the assumptions about the wage-setting strategies of firms do not affect the estimation of frictional parameters $(\lambda_0, \lambda_1, \delta)$. Only in the final step of the estimation procedure, that part of the model which describes firm behaviour (i.e. the first-order condition of firm's problem) is exploited to estimate productivity dispersion. Bontemps et al. (2000) emphasise that the estimates of the frictional parameters can be expected to be consistent under a wide range of assumptions on the demand side of the story. This class of models includes, among others, the specifications of equilibrium search models outlined and estimated in the previous sections.

The estimates of the frictional parameters are given in Table 7, whereas the kernel estimates of f are shown in Figures 1 to 3 in the Appendix. It appears that λ_1 is generally very close to the estimates obtained from the homogeneous model with measurement error and from the model with discrete productivity dispersion while δ is almost identical. Since the estimates of $(\lambda_0, \lambda_1, \delta)$ in this setting can be expected to be robust

with respect to different mechanisms determining the wage distributions, we can conclude that all specifications of the Burdett-Mortensen model generating an acceptable fit to the wage data produce appropriate estimates for the frictional parameters. This suggests that to the extent we are concerned with the estimation of frictional parameters it is not so important whether the deviations from the theoretical distribution predicted by the homogeneous model are attributed to measurement error or to employer heterogeneity as long as the shape of the wage distribution is captured by the specification. This is essentially the same result as found by Bunzel et al. (2001) with the Danish data.

The average productivity values in the last column are computed by taking the averages of point estimates p over workers who found a job (see (28) in the Appendix). However, the estimated relationship between the wage offer and productivity is not consistent with the theory. The condition that $f(w)(1 + \kappa_1 [1 - F(w)])$ decreases everywhere is violated for small wages in all worker groups. In other words, the model fails to capture a steeply increasing wage density observed at the lower end of the wage distribution. Recall that the model with a discrete distribution of productivity also fails to explain the shape of the left tail of the wage distribution.

The increasing pattern of $f(w)(1 + \kappa_1 [1 - F(w)])$ on small wages implies that the relationship $K(p)$ is downward-sloping for small values of p . This suggests the wages offered by some less productive firms are not optimal as they could increase their profits by reducing their wages. A binding minimum wage is an obvious candidate to explain this failure of the model, as it may prevent low paying firms from lowering their wages further. As pointed out by Van den Berg and Van Vuuren (2000), omitted worker heterogeneity may also provide an explanation. To see this, suppose that workers within a given labour market segment are heterogeneous with respect to their value of non-market time (this heterogeneity may result, for example, from differences in unemployment benefits or in the value of leisure). In this case workers will apply different reservation wages when searching from unemployment. Thus a firm which lowers

Table 7: Estimation results with continuous productivity dispersion

	λ_0	λ_1	δ	\bar{p}
<u>Lower vocational and less</u>				
Men, 16-21	.0285 (.0009)	.0234 (.0029)	.0437 (.0034)	20,921
Men, 22-30	.0399 (.0014)	.0291 (.0021)	.0508 (.0026)	28,443
Men, 31-50	.0375 (.0016)	.0254 (.0022)	.0495 (.0025)	34,973
Men, 51-65	.0166 (.0007)	.0298 (.0030)	.0658 (.0040)	36,894
Women, 16-21	.0282 (.0010)	.0179 (.0025)	.0454 (.0032)	25,973
Women, 22-30	.0292 (.0010)	.0194 (.0019)	.0368 (.0022)	22,520
Women, 31-50	.0286 (.0011)	.0118 (.0011)	.0366 (.0019)	30,019
Women, 51-65	.0109 (.0006)	.0117 (.0016)	.0502 (.0034)	38,504
<u>Upper vocational</u>				
Men, 16-21	.0385 (.0014)	.0293 (.0026)	.0362 (.0029)	16,134
Men, 22-30	.0494 (.0017)	.0262 (.0020)	.0352 (.0019)	23,465
Men, 31-50	.0353 (.0012)	.0183 (.0014)	.0301 (.0018)	33,879
Men, 51-65	.0116 (.0006)	.0129 (.0026)	.0440 (.0043)	56,461
Women, 16-21	.0505 (.0023)	.0280 (.0026)	.0412 (.0029)	15,990
Women, 22-30	.0534 (.0019)	.0232 (.0017)	.0322 (.0017)	17,573
Women, 31-50	.0335 (.0013)	.0184 (.0015)	.0296 (.0016)	26,052
Women, 51-65	.0095 (.0012)	.0077 (.0034)	.0495 (.0075)	71,444
<u>Lower university</u>				
Men, 16-30	.0562 (.0018)	.0224 (.0014)	.0264 (.0014)	24,535
Men, 31-65	.0495 (.0018)	.0174 (.0011)	.0266 (.0013)	42,881
Women, 16-30	.0675 (.0026)	.0245 (.0025)	.0251 (.0018)	16,971
Women, 31-65	.0357 (.0015)	.0185 (.0020)	.0265 (.0020)	31,605
<u>Upper university</u>				
Men, 16-30	.0735 (.0067)	.0384 (.0054)	.0139 (.0021)	20,448
Men, 31-65	.0363 (.0018)	.0189 (.0018)	.0190 (.0016)	48,668
Women, 16-30	.0770 (.0051)	.0271 (.0042)	.0204 (.0026)	24,157
Women, 31-65	.0395 (.0031)	.0294 (.0051)	.0272 (.0035)	30,916

Notes: Standard errors are in parentheses. \bar{p} is the average productivity over workers who found a job.

its wage offer may become unattractive for some groups of workers. The firms should take this effect into account when setting wages, which may explain the failure of the theoretical model at the lower end of the wage distribution.

5. Conclusion

This paper has provided quite an extensive structural empirical analysis of various specifications of the Burdett-Mortensen model. We found that, in the absence of measurement error in wages, the equilibrium search model with identical agents does not fit to the wage data. This failure is due to the prediction of the theory that the shape of the equilibrium wage density is convex which is at odds with the wage

data. Introduction of the measurement error in wages or employer heterogeneity in terms of labour productivity across firms provides a way of making the model more flexible. These extended versions of the basic model proved to give a much better fit to the wage data. The frictional parameters of the model, i.e. the lay-off rate and arrival rates of job offers, were found to be fairly robust across the model specifications which fit to the wage data. Stated differently, it does not make much difference for the estimates of the frictional parameters whether the wage distribution is matched by allowing for the measurement error in wages or unobserved productivity differences. However, the estimates of the other parameters – the value of non-market time and productivity terms – vary across the different specifications to a large extent.

In the case of the homogeneous model with the measurement error in wages almost all wage dispersion was attributed to the measurement error. This indicates that the model without (unobserved) worker or employer heterogeneity cannot explain the observed wage variation. Although the equilibrium models with employer heterogeneity match the overall shape of the wage distributions relatively well, they do have problems in explaining the shape of the lower end of the wage distribution, and in particular the testable theoretical conditions implied by the model with continuous productivity dispersion were rejected in the empirical analysis. Of course, one can expect that incorporating employer heterogeneity and measurement error into the same model provides a way of capturing the shape of the lower end of the wage distribution as well. On the other hand, unobserved heterogeneity and measurement errors allowed for in a sufficiently flexible form can be used to 'explain' any discrepancy between the theory and data.

One may call into question whether introducing more unobservables in the empirical analysis of equilibrium search models stands for any progress. As long as the shape of empirical wage distributions is captured mainly by unobservables, we cannot be very satisfied with our empirical analysis. Since the equilibrium search theory describes the joint behaviour of workers and firms, a natural way to proceed is to exploit data on both sides of the labour market. Matched firm-worker data are currently available for many countries, including Finland. Such data contain information on firm sizes, labour turnover and productivity, opening up new possibilities for testing existing models and, more importantly, for estimating more general models (see Postel-Vinay and Robin, 2002, and Christensen et al., 2005, for steps in this direction). But it may be difficult to make significant progress without reconsidering the theoretical structure of the model. In particular, the demand side of the equilibrium search story raises doubts. The assumption that the only choice the firm has to make is to set its wage optimally is obviously not realistic. A consequence is that the labour demand side of the Burdett-Mortensen model is too naive for describing

observed labour turnover in the data. In reality firms face external shocks which lead to the creation and destructions of jobs as wages cannot be fully adjusted in response to these shocks. Equilibrium search models with endogenous job creation and destruction are necessarily much more complicated than the Burdett-Mortensen model. Therefore, a serious challenge for empirical equilibrium analysis of the labour market is to introduce endogenous labour turnover without making the resulting model empirically completely intractable.

References

- Bontemps, C., J.-M. Robin, and G. J. van den Berg (1999).** "An equilibrium job search model with search on the job and heterogeneous workers and firms." *International Economic Review* 40, 1039–1074.
- Bontemps, C., J.-M. Robin, and G. J. van den Berg (2000).** "Equilibrium search with continuous productivity dispersion: Theory and nonparametric estimation." *International Economic Review* 41, 305–358.
- Bowlus, A. (1997).** "A search interpretation of male-female wage differentials." *Journal of Labour Economics* 15, 625–657.
- Bowlus, A., N. M. Kiefer, and G. R. Neumann (1995).** "Estimation of equilibrium wage distributions with heterogeneity." *Journal of Applied Econometrics* 10, S119–S131.
- Bowlus, A., N. M. Kiefer, and G. R. Neumann (2001).** "Equilibrium search models and the transition from school to work." *International Economic Review* 42, 317–343.
- Bunzel, H., B. J. Christensen, P. Jensen, N. M. Kiefer, L. Korsholm, L. Muus, G. R. Neumann, and M. Rosholm (2001).** "Specification and estimation of equilibrium search models." *Review of Economic Dynamics* 4, 90–126.
- Burdett, K., and D. T. Mortensen (1998).** "Wage differentials, employer size, and unemployment." *International Economic Review* 39, 257–273.
- Christensen, B. J., and N. M. Kiefer (1994a).** "Measurement error in the prototypical job-search model." *Journal of Labor Economics* 12, 618–639.
- Christensen, B. J., and N. M. Kiefer (1994b).** "Local cuts and separate inference." *Scandinavian Journal of Statistics* 21, 389–407.
- Christensen, B. J., and N. M. Kiefer (1997).** "Inference in non-linear panel with partially missing observations: The case of the equilibrium search model." *Journal of Econometrics* 79, 201–219.
- Christensen, B. J., R. Lentz, D. T. Mortensen, G. R. Neumann, and A. Werwatz (2005).** "On-the-job search and the wage distribution." *Journal of Labor Economics* 23, 31–59.
- Diamond, P. A. (1971).** "A model of price adjustment." *Journal of Economic Theory* 3, 156–168.
- Eckstein, Z., and G. J. van den Berg (2007).** "Empirical labour search: A survey." *Journal of Econometrics* 136, 531–564.

- Goffe, W. L., G. D. Ferrier, and J. Rogers (1994).** “Global optimization of statistical functions with simulated annealing.” *Journal of Econometrics* 64, 65–99.
- Kiefer, N. M., and G. R. Neumann (1993).** “Wage dispersion with homogeneity: The empirical equilibrium search model.” In *Panel Data and Labour Market Dynamics*. 57–74. Eds. H. Bunzel, P. Jensen, and N. C. Westergård-Nielsen. Amsterdam: North-Holland.
- Korkeamäki, O., and T. Kyyrä (2000).** Integrated panel of Finnish companies and workers. VATT-Discussion Paper 226, Helsinki.
- Mortensen, D. T. (1990).** “Equilibrium wage distributions: A synthesis.” In *Panel Data and Labor Market Studies*. 279–296. Eds. J. Hartog, G. Ridder, and J. Theeuwes, Elsevier Science Publishers B.V., North-Holland.
- Mortensen, D. T., and C. A. Pissarides (1999).** “New developments in models of search in the labor market.” In *Handbook of Labor Economic* Vol. 3, 2567–2624. Eds. O. Ashenfelter, and D. Card. Amsterdam: North-Holland.
- Postel-Vinay, F., and J.-M. Robin (2002).** “Equilibrium wage dispersion with worker and employer heterogeneity.” *Econometrica* 70, 2295–2350.
- Rogerson, R., R. Shimer, and R. Wright (2005).** “Search-theoretic models of the labour market: A survey.” *Journal of Economic Literature* XLIII, 959–988.
- Van den Berg, G. J. (1999).** “Empirical inference with equilibrium search models of the labor market.” *The Economic Journal* 109, F283–F306.
- Van den Berg, G. J. (2003).** “Multiple equilibria and minimum wages in labour markets with informational frictions and heterogeneous production technologies.” *International Economic Review* 44, 1337–1357.
- Van den Berg, G. J., and G. Ridder (1993).** “Estimating an equilibrium search model from wage data.” In *Panel Data and Labour Market Dynamics*. 43–56. Eds. H. Bunzel, P. Jensen, and N. C. Westergård-Nielsen. Amsterdam: North-Holland.
- Van den Berg, G. J., and G. Ridder (1998).** “An equilibrium search model of the labor market.” *Econometrica* 66, 1183–1221.
- Van den Berg, G. J., and A. van Vuuren (2000).** “Using firm data to assess the performance of equilibrium search models of the labour market.” *Annales d’Economie et de Statistique* 67–68, 227–256.

Appendix

A.1 This appendix gives details of the estimation procedures. Identical workers and firms without measurement error

Substituting (6) and (7) into (20) and summing over the observations gives the log-likelihood function in terms of $(\lambda_0, \lambda_1, \delta, p, r)$. The properties of the maximum likelihood estimators are not standard in this case, however. This is due to the dependence of the support of F on the unknown parameters of the model. Kiefer and Neumann (1993) show that one can simplify the estimation problem considerably by reparametrizing the model from $(\lambda_0, \lambda_1, \delta, p, r)$ to

$(\lambda_0, \lambda_1, \delta, h, r)$ and using order statistics to estimate the bounds of the support of F . To pursue this route, we solve the system (8) and (9) for p and b to obtain

$$(22) \quad p = \frac{\beta}{\beta - \alpha} r + \frac{\alpha}{\alpha - \beta} h,$$

$$(23) \quad b = \frac{\beta - 1}{\beta - \alpha} r + \frac{\alpha - 1}{\alpha - \beta} h,$$

where α and β are given by (10) and (11) respectively. Using (22) to substitute p out of the expressions for F and f , we can rewrite the log-likelihood function in terms of $(\lambda_0, \lambda_1, \delta, r, h)$.

Following Kiefer and Neumann (1993), we estimate r and h using the sample minimum and maximum respectively. In the second step the frictional parameters are estimated by maximizing the likelihood function with respect to $(\lambda_0, \lambda_1, \delta)$ conditional on the estimates of (r, h) . Due to the properties of order statistics, the estimates of (r, h) are superconsistent, converging to their true values at a rate faster than \sqrt{n} , where n is the sample size. This suggests that ignoring variation in the estimates of (r, h) does not affect asymptotic inference about $(\lambda_0, \lambda_1, \delta)$, and therefore we can treat (r, h) as fixed in the maximum likelihood estimation of the frictional parameters, even though they are not exogenous (see Christensen and Kiefer, 1994b). Given the consistent estimates of $(\lambda_0, \lambda_1, \delta, r, h)$, we can estimate p and b using (22) and (23) and compute their standard errors using the delta method (see Bunzel et al., 2001, for details).

A.2 Identical workers and firms with measurement error

To deal with measurement errors, we assume that the wage observation in the data, say x , is the product of the true unobserved wage w and an error term ε , so that $x = w \cdot \varepsilon$. The multiplicative measurement error is assumed to be independently and identically distributed across individuals and to be independent of all other variables in the model. Following Christensen and Kiefer (1994a) and Bunzel et al. (2001), we assume that ε has a Pearson Type V distribution with unit mean, variance σ^2 and density

$$(24) \quad g_\varepsilon(\varepsilon) = \frac{(1 + \sigma^{-2})^{2+\sigma^{-2}}}{\tilde{\Gamma}(2 + \sigma^{-2}) \cdot \varepsilon^{3+\sigma^{-2}}} \cdot \exp\left(-\frac{1 + \sigma^{-2}}{\varepsilon}\right),$$

where $\tilde{\Gamma}$ denotes the gamma function. A consequence of allowing for the measurement error is that we need to add an integral for each wage observation in the individual likelihood contribution. Formally, we replace (17) by

$$(25) \quad \ell = \varphi(d) \left(\int_{x/h}^{x/r} \phi\left(j, a \left| \frac{x}{\varepsilon} \right.\right) \cdot f\left(\frac{x}{\varepsilon}\right) \frac{1}{\varepsilon} g_\varepsilon(\varepsilon) d\varepsilon \right)^{1-c_d},$$

where φ , ϕ and f are as given in the previous section, and $1/\varepsilon$ is the Jacobian of the transformation between the true and observed wages given the error term. In this case we do not use order statistics for (r, h) but estimate them simultaneously with $(\lambda_0, \lambda_1, \delta, \sigma)$ by maximizing the likelihood function based on (25), in which the integral is evaluated numerically in each iteration. The presence of measurement errors makes the support of the distribution of observed wages independent of the unknown parameters, so the maximum likelihood estimation is standard in this case.

With the estimates of $(\lambda_0, \lambda_1, \delta, r, h)$ in hand, the estimates of (p, b) can be computed using (22) and (23) as before. In Figures 1–3 the predicted densities for *observed* wages are obtained by inserting the parameter estimates into

$$\int_{x/h}^{x/r} f\left(\frac{x}{\varepsilon}\right) \frac{1}{\varepsilon} g_\varepsilon(\varepsilon) d\varepsilon.$$

A.3 Discrete productivity dispersion

In addition to the problem that the bounds of the support of F depend on unknown parameters, the estimation of the model specification with a discrete productivity distribution is further complicated by the fact that the likelihood function is not differentiable at the wage cut points

$(\bar{w}_1, \dots, \bar{w}_{q-1})$. An estimation method which can deal with these complications has been developed by Bowlus et al. (1995). Once again it is convenient to reparametrize the model in a similar fashion as was done in the homogeneous case. Evaluating (12) at $w = \bar{w}_i$ and solving for p_i gives the following relationship between the productivity terms, wage cuts and the fractions of firm types:

$$(26) \quad p_i = \frac{1}{1 - \mu_i^2} \bar{w}_i - \frac{\mu_i^2}{1 - \mu_i^2} w_i,$$

where

$$(27) \quad \mu_i = \frac{1 + \kappa_1(1 - \gamma_i)}{1 + \kappa_1(1 - \gamma_{i-1})} \in (0, 1),$$

$$i = 1, 2, \dots, q.$$

Substituting p_i 's out of (12) and (13) using (26) allows us to write the likelihood function in terms of $(\lambda_0, \lambda_1, \delta, r, h, \bar{w}_1, \dots, \bar{w}_{q-1}, \gamma_1, \dots, \gamma_{q-1})$. Since there are kinks at the wage cuts in F , the density f and hence the likelihood function is discontinuous at these points.

Bowlus et al. (1995) show that the maximum likelihood estimates of wage cuts $(\bar{w}_1, \dots, \bar{w}_{q-1})$ come from the set of observed wages. As in the homogenous case, we use order statistics to estimate (r, h) . Conditional on these estimates, the likelihood function can be maximized using an iterative procedure with two steps in each iteration. In the first step the likelihood function is maximized with respect to $(\bar{w}_1, \dots, \bar{w}_{q-1})$ holding $(r, h, \lambda_0, \lambda_1, \delta)$ fixed, using simulated annealing which randomly searches over the possible wage cut combinations according an optimal stopping rule.¹² Given the estimates of $(\bar{w}_1, \dots, \bar{w}_{q-1})$, the corresponding discontinuity points in the wage offer distribution $(\gamma_1, \dots, \gamma_{q-1})$ are estimated by observed frequencies in the wage data. In the second step the likelihood function is maximized with respect to $(\lambda_0, \lambda_1, \delta)$ conditional on $(r, h, \bar{w}_1, \dots, \bar{w}_{q-1}, \gamma_1, \dots, \gamma_{q-1})$. Since this part of the maximization problem is smooth,

¹² For simulated annealing, see Goffe et al. (1994) or Bowlus et al. (1995).

Figure 1. Wage offer densities for workers with lower vocational education or less.

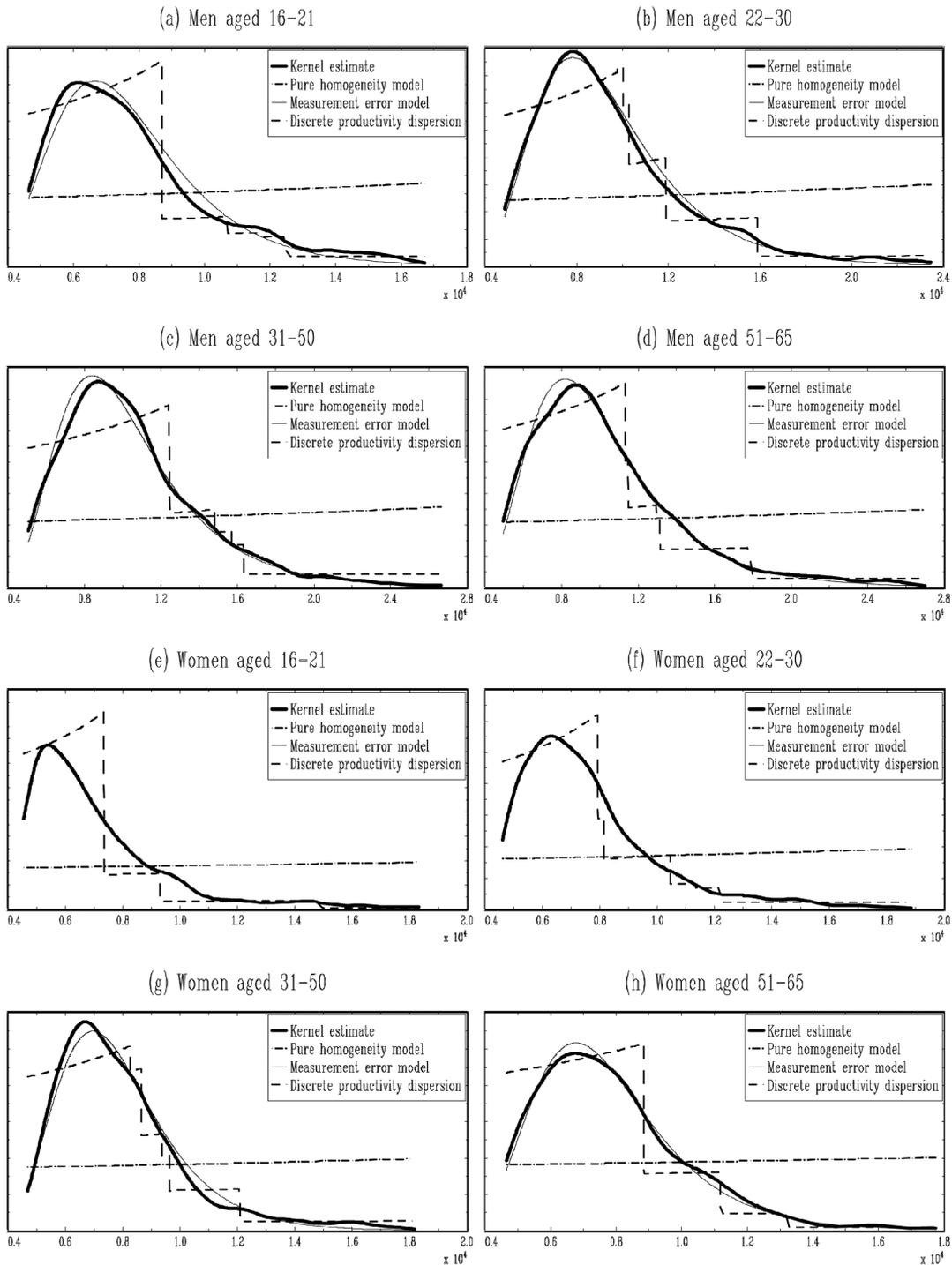


Figure 2. Wage offer densities for workers with upper vocational education.

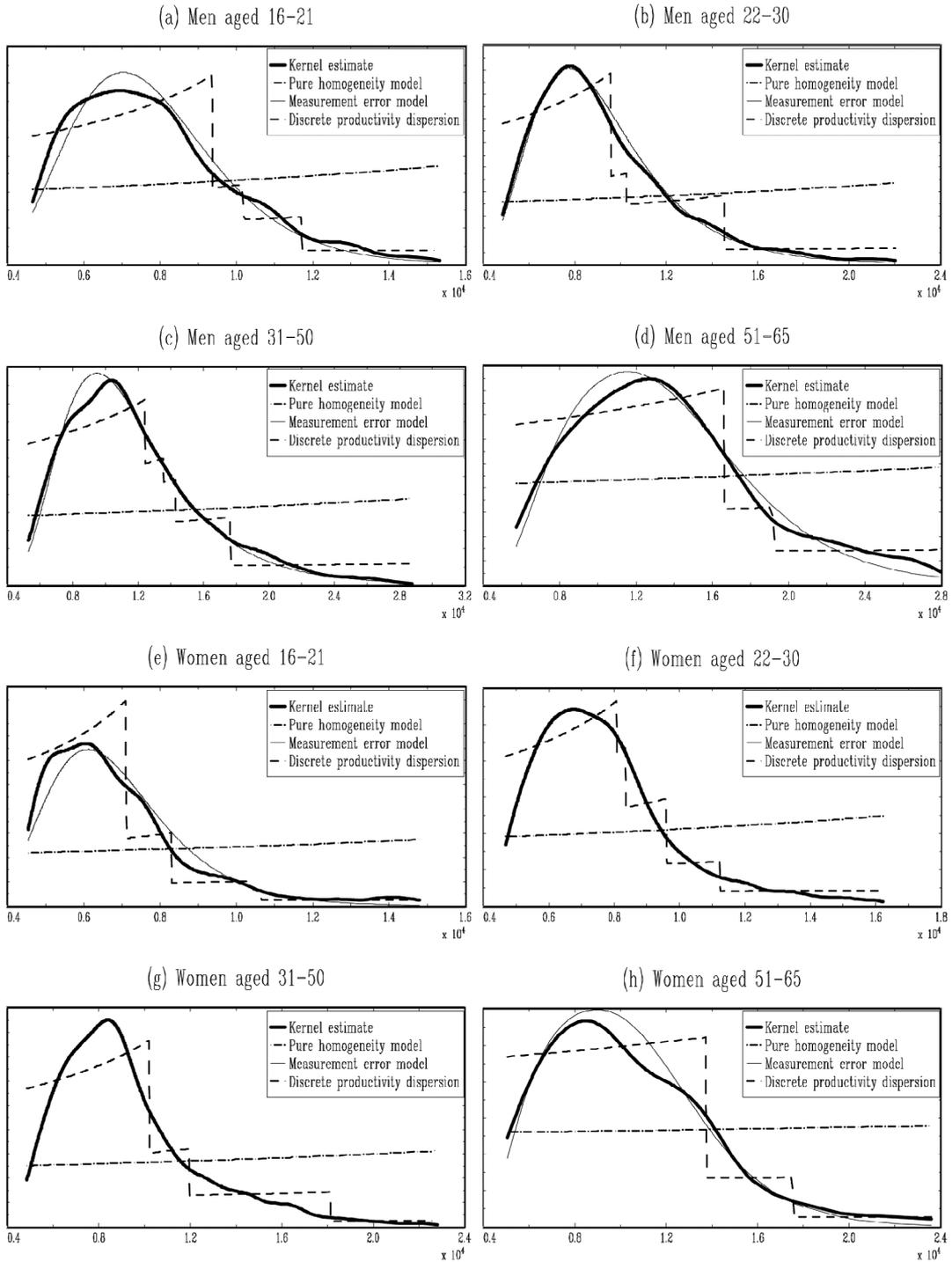
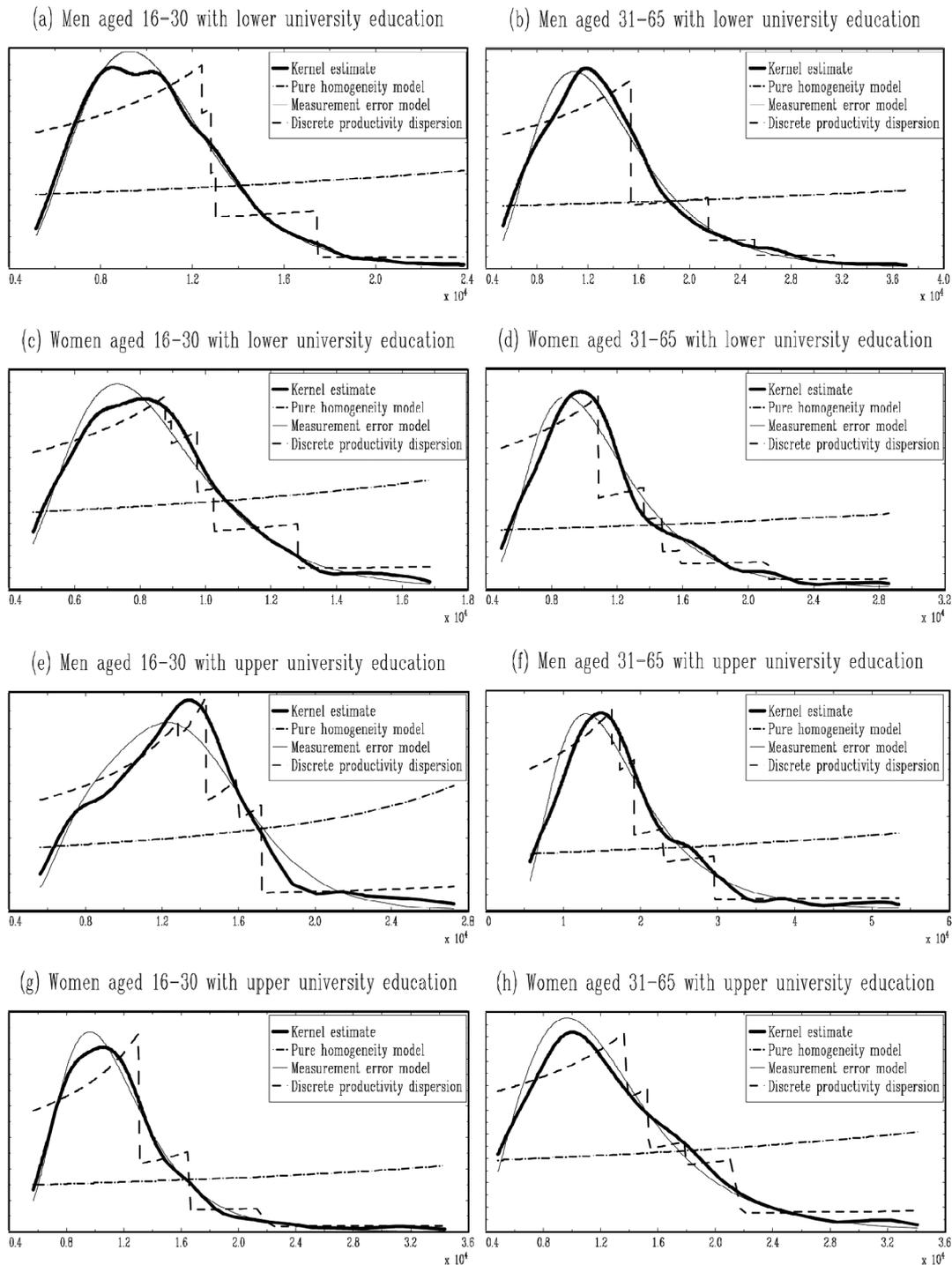


Figure 3. Wage offer densities for workers with university education.



standard maximum likelihood algorithms can be applied. These two steps are then iterated until convergence occurs.

In addition to the order statistics estimators for (r, h) , the maximum likelihood estimators of the wage cuts in the wage offer distribution $(\bar{w}_1, \dots, \bar{w}_{q-1})$ also converge to their true value at a rate faster than \sqrt{n} . It follows that they are asymptotically independent of the maximum likelihood estimator of $(\lambda_0, \lambda_1, \delta)$ and the theory of local cuts by Christensen and Kiefer (1994b) justifies conditioning on them in the second step of the procedure. The iterative separate maximization can be shown to lead to a joint maximum of the likelihood function on convergence.

There is no formal test for choosing a value of q , the number of firm types. However, the authors of this estimation technique argue that the likelihood ratio test of one value of q against another based on the standard χ^2 -criterion can be expected to work reasonably well in practice. This is so even though the exact distribution of the test statistics is not known due to the non-regular estimation procedure. Thus, we choose the number of firm types by comparing two times the improvement in the log-likelihood function with each additional firm type to the $\chi^2_{0.05}$ critical value.¹³ Once the other parameters of the model are estimated, unobserved heterogeneity terms (p_1, \dots, p_q) can be estimated using (26) and (27).

A.4 Continuous productivity dispersion

The estimation method of Bontemps et al. (2000) for the model with a continuous productivity distribution involves three steps. The first step of the procedure is to estimate F and f from the wage data using some nonparametric procedure. In the second step the likelihood function is maximized with respect to $(\lambda_0, \lambda_1, \delta)$ conditional on the nonparametric estimates of F and f . As the final step $p = K^{-1}(w)$ and $\gamma(p)$ are esti-

mated using¹⁴

$$(28) \quad K^{-1}(w) = w + \frac{1 + \kappa_1 [1 - F(w)]}{2\kappa_1 f(w)},$$

$$(29) \quad \gamma(p) = \frac{2\kappa_1 f(w)^3}{\kappa_1 f(w)^2 - f'(w) (1 + \kappa_1 [1 - F(w)])},$$

where F, f and f' are replaced by their nonparametric estimates from the first step and $\kappa_1 = \lambda_1/\delta$ by its maximum likelihood estimate from the second step.¹⁵ To estimate the wage offer density f , we apply the standard Gaussian kernel density estimator and choose the bandwidth by a rule of thumb that minimizes the mean integrated square error. Corresponding estimates of F and f' are then obtained by integration and differentiation of the kernel density estimate. Standard errors are finally obtained by bootstrapping the whole estimation procedure outlined above.¹⁶

It is worth emphasising that the estimation method of Bontemps et al. (2000) is very simple in computational respect. Substitution of the kernel estimates in the likelihood function circumvents the need for numerical integrations, while the third step does not require even iterations. This is a clear advantage compared to the estimation procedure of Bowlus et al. (1995) for the case of discrete productivity dispersion.

¹⁴ The first equation is simply the first-order condition (15) solved for $p = K^{-1}(w)$. The second equation can be found by differentiating the first equation with respect to w and noting that $(K^{-1})'(w) = f(w)/\gamma(p)$ by virtue of the relationship $F(w) = \Gamma(K^{-1}(w))$.

¹⁵ Obviously, the final step requires the denominator of (29) to be positive. This holds if the second-order condition of the firm's problem is satisfied or, in other words, if $f(w)(1 + \kappa_1 [1 - F(w)])$ decreases over the support of F .

¹⁶ Our estimation procedure differs slightly from that used by Bontemps et al. (2000) because of the different sampling scheme. Contrary to the inflow sample of unemployment, Bontemps et al. (2000) use a sample from the French Labour Force Survey drawn from the stock of employed and unemployed workers. Consequently, wage observations in their data come from G , not from F as in our data. As such they estimate G (instead of F) using a nonparametric procedure and then recover the associated F using the equilibrium flow relationship. They also replace (28) and (29) by the corresponding equations expressing p and γ as functions of G, g, g' and κ_1 , where g is the density of the earnings distribution and g' its derivative.

¹³ A Monte Carlo evidence of Bowlus et al. (2001) indicates a tendency towards overfitting the number of heterogeneity types using this criterion. However, they further find that choosing a value of q greater than the true value has only a minor effect on the estimates of (λ_1, δ) , while the order statistics estimators of (r, h) and the ML estimator of λ_0 are obviously unaffected by the value of q chosen.