Semi-nonparametric estimation of an equilibrium search model

Koning, P.; van den Berg, G.J.; Ridder, G.

1998

aculteit der Economische Wetenschappen en Econometrie

Serie research memoranda

Semi-Nonparametric Estimation of an Equilibrium Search Model

Pierre Koning
Gerard J. van den Berg
Geert Ridder

Research Memorandum 1998-3 January 1998

/a/l/e/r/t/

applied
labour
economics
research
team

vrije Universiteit *amsterdam*

Semi-Nonparametric Estimation of an Equilibrium Search Model

Pierre Koning *
Gerard J. van den Berg †
Geert Ridder ‡

January 28, 1998

Abstract

We specify and estimate an equilibrium job search model with productivity differences across labor market segments. The model allows for two types of unemployment: frictional unemployment due to search frictions and structural unemployment due to wage floors. Wage floors exist because of high unemployment benefits or binding minimum wages. The productivity distribution is estimated semi-nonparametrically along the lines of Gallant-Nychka, using Hermite series approximation. We decompose the total unemployment rate and we examine the effect of changes in the minimum wage.

*Ministry of Social Affairs and Employment, The Netherlands.
†Department of Economics, Free University Amsterdam, De Boelelaan 1105, NL-1081 HV Amsterdam, The Netherlands; CEPR, and Tinbergen Institute.
‡Department of Economics, Free University Amsterdam, and Tinbergen Institute.

The Organization for Strategic Labor Market Research (OSA) allowed us to use the OSA panel data. Thanks to Ruud Koning for many valuable suggestions.

Keywords: wages, unemployment, job mobility, minimum wage, Hermite series, identification.

1 Introduction

Most studies on the impact of minimum wages on employment and wages exploit the fact that minimum wages vary over time or across labor market segments.' As an alternative, Meyer and Wise (1983a,b) propose a method in which the employment effect of minimum wages can be inferred using data on the cross-sectional distribution of wages only, and without any variation in minimum wages at all. This method rests upon two assumptions. First, in the absence of a mandatory minimum wage, the wage earnings distribution of individual workers follows has a functional form that is recoverable (see Flinn and Heckman (1982)). This means that the distribution belongs to a family of distributions that is not closed under truncation. Thus the untruncated distribution can be fully identified its truncated version (for example, the lognormal distribution belongs to this family of distributions). Second, above a certain level (close to the minimum wage) the wage earnings distribution is unaffected by the imposition of a minimum wage. With these two assumptions it is possible to deduce the number of workers who should be earning a wage at or below the minimum from the estimated truncated wage earnings distribution.

The Mever and Wise technique has been used by e.g. Albæk and Madsen (1987) and 'Van Soest (1989). Obviously, the fundamental drawback is that the estimated effect of the minimum wage is sensitive to functional form assumptions on the distribution of wage earnings. In particular, the probability mass below the minimum wage is obtained by extrapolation of a distribution to a region where there are no observations, and alternative assumptions on the distribution of wages can result in different estimates of the effect of (changes in) minimum wages.[2] In contrast to this, Koning, Ridder and Van den Berg (1995) present a model in which the effect of minimum wages can be identified without making untestable distributional assumptions. They specify and estimate an equilibrium search model which allows for two types of unemployment: unemployment due to search frictions (frictional unemployment) and unemployment due to minimum wages (structural unemployment). This is an interesting distinction, as it allows one to infer the group of individuals that also would be unemployed in the absence of minimum wages. Due to the structural setup of the model, it is possible to identify the rate of structural unemployment with data on (censored) unemplovment durations and/or the fraction of unemployment. In addition? the rate of structural unemployment is also equal to the probability mass of the individual

'See Card and Krueger (1995) and Dolado et al. (1996) for recent surveys of the literature.

[2]See Dickens, Machin and Manning (1994) for an evaluation of the Meyer and Wise approach.

productivity distribution below the minimum wage. This can be used as a check on the specification of the productivity distribution.

In this paper we relax the distributional assumptions made by Koning, Ridder and Van den Berg (1995) with respect to the productivity distribution of individual workers. For this we follow the so called 'semi-nonparametric' (SNP) estimation method developed by Gallant and Nvchka **(1987).** The basic idea of this estimation method is that any proper density can be approximated by a Hermite series. Within the context of the model, the advantages of the SNP approach are twofold. First, it allows us to estimate the truncated productivity density more accurately, resulting in better estimates of e.g. the elasticity of the rate of structural unemployment with respect to the minimum wage. Second, it reduces the impact of misspecification of the productivity distribution on the implied rate of structural unemployment. This rate is overidentified, and a more flexible functional form may help us reducing the biasing impact of this on the parameter estimates of the model.

The organization of this paper will be as follows. Section 2 discusses the Burdett-Mortensen model and the extension to heterogeneous agents. Section 3 describes the data we use in the empirical analysis, and the likelihood function is derived in Section 4. Section 5 discusses the identification of the model, as well as the SNP estimation method. In Section 6 we conduct two simulation experiments in order to test for the accuracy of the SNP method, as well as to examine the impact of misspecification on the parameter estimates of the model. In Section 7 we present the estimation results of the equilibrium search model. Section 8 concludes. Most of the exposition in Sections 2, 3 and 4 will be kept brief, as the same material is discussed more extensively in Koning, Ridder and Van den Berg (1995).

2 The equilibrium search model

2.1 Equilibrium search with identical agents

Consider a labor market with identical agents, i.e. a market in which all workers are equally productive at all firms. Even in this case, the Burdett-Mortensen model has a dispersed equilibrium wage (offer) distribution, i.e. the law of one price does not hold. Allowing for heterogeneity in the productivity of workers adds to the equilibrium dispersion of wages, and it allows us to distinguish between frictional and structural unemplovment.

We assume that there are large fixed numbers of workers and firms participating in the labor market (formally a continuum of each). The measure of workers is denoted by m, whereas the measure of firms is normalized to one. Workers receive job offers at given Poisson rates, λ_0 when unemployed and λ_1 when employed, with $0 \leq \lambda_0, \lambda_1 < \infty$. Job offers are independent random drawings from the distribution $F(w)$ of wage offers. When an offer arrives, the worker must decide whether to accept the offer or to reject it and continue searching for a. better one. Workers become unemployed at the exogenous separation rate δ $(0 < \delta < \infty)$. During unemployment, the worker receives unemplovment benefits b $(0 \leq b < \infty)$. A firm posts a wage that is the same for all workers, and it does not bargain over this wage. In the basic model, the marginal value product of any worker at any firm is the same. It is denoted by p. The firms have a linear production function, so that the average and marginal product are equal. Individuals and firms are assumed to maximize their expected wealth.

Assuming that the wage offer distribution is known and stationary over time and that wages are constant in jobs, the supply-side of this model is equivalent to the standard job search model with search on the job (see e.g. Mortensen (1986). Thus, the optimal strategy of an unemployed worker has the reservation wage property. The reservation wage r is

$$r = b + (\lambda_0 - \lambda_1) \int_r^\infty \frac{\overline{F}(w)}{\delta + \lambda_1 \overline{F}(w)} dw \tag{1}$$

. with $\overline{F} = 1 - F$. Employed workers accept any wage offer that exceeds their current wage.

It is important to distinguish between the distribution of wages offered to individuals, which is the *wage offer distribution F,* and the distribution of wages received bv workers who are currently employed. The latter is referred to as the *earnings distribution,* and we denote this distribution by G. In equilibrium, the flow of workers out of jobs with a given wage is equal to the inflow in such jobs. Similarly, the flows into and out of unemployment are equal. Firms that offer a wage lower than the reservation wage of the unemployed do not attract any worker and therefore cannot survive. The market is only viable if there is a positive gain from trade, i.e. if $p > b$. Under these assumptions we have

$$G(w) = \frac{\delta F(w)}{\delta + \lambda_1 \overline{F}(w)} \qquad w \geq r \tag{2}$$

$$F(r) = 0 \tag{3}$$

$$\frac{u}{m} = \frac{\delta}{\delta + \lambda_0} \tag{4}$$

In (4), u is the measure of unemployed workers. Thus, ***u/m*** is the rate of frictional unemployment in this market. It is determined by the rates δ and λ_0. Note that with full information on the location of jobs, i.e. in the absence of search frictions, $\lambda_0 = \infty$ and $u = 0$. *Frictional unemployment* should be distinguished from *structural unemployment,* which occurs if the unemployment benefits, or more generally the value of leisure, exceeds the value product p. This type of structural unemployment is voluntary, because workers are better off if they are unemployed.

If there is a mandatory minimum wage, denoted by w_L, then wage offers must exceed this wage. If $p < w_L$, then firms do not employ any worker, and there is structural unemployment. This type of unemployment is involuntary if $b < p < w_L$, because workers would supply labor if the minimum wage would be lower than p. Hence, if $p > \max(w_L, b)$ then there is frictional unemployment equal to u, while if the reverse holds there is voluntary or involuntary structural unemplovment equal to ***m***.

For the moment, assume that $p > \max(w_L, b)$. The steady-state level of production is determined by the size of the steady-state work force l of the firm. That work force depends on the wage w set by the firm, the reservation wage r of the unemployed, and the distribution F of wages set by other firms competing for the same workers. Each firm chooses w to maximize its steady-state profit flow π, which, given r and F, equals $(p - w)l(w; r, F)$.

A noncooperative steadv-state equilibrium solution consists of a reservation wage r and a wage offer distribution F such that (i) r satisfies **(1)** given F, and (ii) every w in the support of F maximizes the steady-state profit flow π. Burdett and Mortensen (1998) prove that there is a unique equilibrium and they give closed-form solutions. The distributions F and G have probability density functions f and g with a support equal to $[\underline{w}, \overline{w}]$, with

$$\begin{aligned} \underline{w} &= \max(w_L, r) \\ \overline{w} &= \left[\frac{\lambda_1}{\delta + \lambda_1}\right]^2 \underline{w} + \left[1 - \left[\frac{\delta}{\delta + \lambda_1}\right]^2\right] p \end{aligned} \tag{5}$$

The equilibrium wage offer c.d.f and p.d.f. are

$$F(w) = \frac{\delta + \lambda_1}{\lambda_1}\left[1 - \sqrt{\frac{p - w}{p - \underline{w}}}\right] \qquad \text{on} \quad [\underline{w}, \overline{w}] \tag{6}$$

$$f(w) = \frac{\delta + \lambda_1}{2\lambda_1 \sqrt{p - \underline{w}}} \frac{1}{\sqrt{p - w}} \qquad \text{on} \qquad [\underline{w}, \overline{w}] \tag{7}$$

Substitution of *(6)* in (1) gives

$$r = \frac{(\delta + \lambda_1)^2 b + (\lambda_0 - \lambda_1)\lambda_1 p}{(\delta + \lambda_1)^2 + (\lambda_0 - \lambda_1)\lambda_1} \qquad \text{if} \qquad r \geq w_L. \tag{8}$$

If $r < w_L$, the equilibrium reservation wage is not given by (8). However in that case the reservation wage is not effective, because the lowest wage offer is w_L. If $r \geq w_L$ then r and $\overline{w}$ are weighted averages of b and p. Otherwise $\overline{w}$ is a weighted average of p and w_L. Note that r is smaller than b iff λ_0 is smaller than λ_1. In that case the unemployed accept wages below b, because it is easier to find a higher paying job if employed. Allowing for this possibility is important given the empirical evidence on the relative size of b and r *(see* e.g. Narendranathan and Nickel1 (1985) and Van den Berg (1990)).

Using (2), the equilibrium wage (or earnings) density is

$$g(w) = \frac{\delta\sqrt{p - \underline{w}}}{2x} \frac{1}{(p - w)^{3/2}} \qquad \text{on} \qquad [\underline{w}, \overline{w}]. \tag{9}$$

Note that both f and g are increasing densities. The wage distribution is related to the income distribution, and there is abundant empirical evidence that the income distribution does not have an increasing density. We return to this issue below. For an employed individual earning a wage w, the exit rate out of that job equals $\delta + \lambda_1 \overline{F}(w)$. This rate decreases in w, which is consistent with a number of empirical studies on job durations (Lindeboom and Theeuwes (1991) and Van den Berg (1992)).

2.2 Heterogeneity in value products

In reality, workers and firms are obviously not identical. All parameters of the model. *i.e.* λ_0, λ_1, 6, b and p vary over workers and/or firms. As argued in Van den Berg and Ridder (1998), there are basically two ways to introduce heterogeneity: within the market and between markets. Heterogeneity within the market means that there is one labor market within which heterogeneous workers and firms interact. Heterogeneitv between markets means that the labor market is segmented and consists of a large number of separate submarkets within which workers and firms are homogeneous. We follow the latter approach, and we assume that we observe a mixture of homogeneous markets. Conceivably we can stratify on all the parameters. In the present context, dispersion in p is particularly relevant, since it allows for the possibility of structural unemployment (namely when $p < \max(w_L, b)$). Because we assume that p follows a continuous distribution, we have effectively a continuum of submarkets which differ in the value product of workers.

As noted above, there is abundant empirical evidence that the income distribution does not have an increasing density as is predicted by the model with identical agents. Allowing for heterogeneity in p may improve the fit to the observed wage offer and earnings distribution. To see this we consider the following transformation of w (note that we acknowledge the dependence of the support of w on p)

$$y = \frac{p - w}{p - \underline{w}(p)} \tag{10}$$

so that the excess wage $w - \underline{w}(p)$ satisfies

$$w - \underline{w}(p) = (1 - y)(p - \underline{w}(p)) \qquad \eta^2 \leq y \leq 1. \tag{11}$$

with $\eta = \delta/(\delta + \lambda_1)$

Equation (11) describes the wage determination in the Burdett-Mortensen model. The excess wage $w - \underline{w}(p)$ is a fraction of the excess productivity $p - \underline{w}(p)$. This fraction is a random variable with a distribution that depends on λ_1/δ, which is the expected number of wage offers during a spell of *employment* (i.e. a spell that starts with the acceptance of a job from unemployment and ends with a layoff). This ratio is a measure of the speed at which the worker climbs the job (and wage) ladder, with $y = 1$ corresponding to the bottom, $w = \underline{w}(p)$, and $y = \eta^2$ to the top, $w = \overline{w}(p)$, of this ladder. From (11) it follows that the moments of $w - \underline{w}(p)$ in either the wage offer or the earnings distribution are the product of $(p - \underline{w}(p))^n$ and an expression that only depends on η. By choosing an appropriate distribution of the productivity p, the moments of the observed wage offer or earnings distribution can be matched. Hence, we expect that an acceptable fit to the data depends on the allowance for sufficient heterogeneity in p. This is confirmed in Van den Berg and Ridder (1993).

3 The data

The model is estimated with the OSA (Netherlands Organization for Strategic Labor Market Research) Labor Supply Panel Survey. This panel started in

1985. Presently four waves are available (April-May 1985, August-October 1986, August-October 1988, and August-November 1990).

In the OSA panel a random sample of households in the Netherlands is followed over time. Because the study concentrates on individuals who are between 15 and 61 years of age and who are not full-time students, only households with at least one person in this category are included. All individuals (and in all cases the head of the household) in this category are interviewed. The first wave consists of 4020 individuals (in 2132 households).

The data allow for a reconstruction of the sequence of labor market states occupied by the respondents and the sojourn times and income levels in these states. Part of the information is retrospective. For example, the first wave (in 1985) contains information on the labor market histories from January 1, 1980 until the date of the interview. The following labor market positions are distinguished: employment (job-to-job changes are recorded), self-employment, unemployment, and not-in-labor-force (subdivided into military service, full-time education, and other activities not related to the labor market).

In this paper we restrict attention to respondents who were participating in the panel as of the first wave. Individuals who were self-employed at certain dates of the time span covered by the survey were deleted, since it is likely that the behavior of such individuals, at least in a certain period, deviates substantially from the behavior as described by the model. For similar reasons, we do not use information on respondents who are observed to be nonparticipant in the labor market at certain dates. Finally, we delete observations for which the reported wage is smaller than the legal minimum wage.[3]

As a result, we have a sample of 1767 individuals. Of these, 12% were unemployed at the date of the first interview. In our sample, 34% participate in all four waves of the panel, while 33% only participate in the first wave. The income changes at transitions occurring before the date of the first interview (April 1985) are only recorded to lie in one of a few broad intervals. This makes the information on spells ending before this date relatively inaccurate in comparison to spells ending after it. For the latter spells we observe exact income levels. Therefore, the first spell used is the spell which is in progress at the date of the first interview. For computational reasons, information on subsequent spells is not used either.

The benefits level b is taken to be the average in the subsample. The mandatory minimum wage w_L equals 1000 Dutch guilders (monthly) for respondents younger than 23 years of age, and 1450 Dutch guilders for older respondents. In

[3]An alternative way to deal with this problem is to assume that wages are measured with error; see Van den Berg and Ridder (1998).

the overall sample, w_L is taken to be the average in the sample.

4 The likelihood function

If we only allow for heterogeneity in p, then the parameters of the model are λ_0, λ_1, 6 and the parameters of the distribution of p. We estimate these parameters from observed labor market histories of a sample of individuals. These labor market histories are recorded with the observation plan of Section 3. In this observation plan a random sample from all individuals in a particular age bracket, who were either employed or unemployed in May 1985, was drawn. The individuals in the sample were followed until either their first transition from the state occupied in May 1985 or the end of the observation period, i.e. November 1990. Their labor market history from January 1, 1980 onward was also reconstructed. Hence the individuals in the sample can be followed backward in time starting in May 1985. In the observation plan they are followed until either their first transition from the state occupied in May 1985 or the end of the observation period, i.e. January 1, 1980. For the younger age groups we observe many transitions from (un)employment to full-time education or, for the men, military service. The corresponding unemployment or job spells are treated as censored spells.

In its simplest form. the model has no observed explanatory variables, and the dependent variables' are aspects of the individual labor market histories. To be specific, the dependent variables include the labor market position (employed or unemployed) at the date of the first interview and the elapsed and residual duration in that position. If an unemployed individual finds a job before the end of the observation period. we observe the re-employment wage w_0. Further? we observe the wages of employed individuals at the time of the first interview, w_1, and, if the job spell of this individual ends before the observation period, the type of transition at the end of this spell. This is either a transition into unemployment, or into another job.

The derivation of the joint distribution of the observed variables is in two steps. First, we derive the joint distribution conditional on a particular value of p. Next, we integrate with respect to p to obtain the joint distribution of the observable variables. Below, we merely list the distributions of some of the main ingredients of the likelihood function. See Koning, Ridder and Van den Berg (1995) for more details.

For a given value of $p > \max(b, w_L)$, the employment and job durations are exponentially distributed. Following Ridder (1984), this implies that the corresponding elapsed and residual durations are independent and also exponentially distributed. If $p < b$, there is no gain from trade, and if $p < w_L$, the minimum wage is too high to employ the workers. In both cases the workers are permanently unemployed. If $p \geq \max(b, w_L)$, then an individual is frictionally unemployed with probability $\delta/(\delta + \lambda_0)$. In our data we have $w_L > b$. Hence the structural unemployment rate u_S equals $H(w_L)$. The unconditional probability that an individual is unemployed at the time of the first interview equals

$$H(w_L) + (1 - H(w_L))\frac{\delta}{\delta + \lambda_0} \tag{12}$$

In the sequel, we assume that the censoring times for the elapsed and residual unemployment durations are stochastically independent of the corresponding durations, *i.* e. we assume that censoring is independent. For the unemployed the joint conditional distribution of the elapsed and residual unemployment durations is degenerate if $p < w_L$. In that case the elapsed and residual unemployment durations are both infinite, hence the observed durations are censored. Because the observation period is at most 129 months, we never observe the elapsed and residual unemployment durations to be infinite, but if they are both censored we allow for this possibilitv. Given the length of the observation period, we consider this to be a reasonable assumption. If $p \geq w_L$, the elapsed and residual unemployment durations at the time of the first interview are independently and exponentially distributed with parameter λ_0. Note that these durations can be censored as well.

The conditional distribution of the accepted wage from unemployment w_0 is only defined if $p \geq w_L$. We denote this conditional p.d.f. by $f(w_0|p)$, with f as in (7). Note that $\underline{w}(p)$ and $\overline{w}(p)$ are as in equations (5) and (8).

If the individual is employed at the date of the first interview, the conditional distributions are onlv defined if $p \geq w_L$. The densitv function of wages of employed individuals at the time of the first interview conditional on p is denoted by $g(w_1|p)$, with g as in (9), again with $\underline{w}(p)$ and $\overline{w}(p)$ as in (5) and (8).

Further, the elapsed and residual job duration at the time of the first interview are independently exponentially distributed with parameter $\delta + \lambda_1\overline{F}(w_1|p)$. Note that this hazard rate out of the job is conditional on p. A subsequent transition to a higher paying job has probability

$$\frac{\lambda_1\overline{F}(w_1|p)}{\delta + \lambda_1\overline{F}(w_1|p)} \tag{13}$$

and the complement of this is the probability of a transition into unemployment.

The individual likelihood contribution is the joint distribution of the observables, which is obtained by integration with respect to the distribution of p. The integration interval for p depends on observed values of paid or accepted wages. If one of these variables is observed, it follows that $\underline{w}(p) \leq w_i \leq \overline{w}(p)$ with $i = 0, 1$. As a consequence, $p > w_L$, so that the individual can not be permanently unemployed. To derive the support of p for a given value of w_0 or w_1, it is convenient to distinguish between the following four regimes,

I.	$b \leq w_L$,	$\lambda_1 > \lambda_0$
II.	$b \leq w_L$,	$\lambda_1 \leq \lambda_0$
III.	$b > w_L$,	$\lambda_1 > \lambda_0$
IV.	$b > w_L$,	$\lambda_1 \leq \lambda_0$

Because b and w_L are known and constant within a segment of the labor market, we know whether the segment is in one of the first two regimes or in one of the last two. In our sample we always have $b \leq w_L$. This implies that permanent unemployment is due to $p < w_L$, *i.e.* a high minimum wage, rather than $p < b$, *i.e.* a high level of unemployment benefits. If, in addition, λ_0 is larger than λ_1, then regime II applies. In this regime the lower bound of $w_i, i = 0, 1$ is a function of p. To demonstrate this, we first define $\underline{p}$ as

$$\underline{p} = w_L + \frac{(\delta + \lambda_1)^2}{(\lambda_0 - \lambda_1)\lambda_1}(w_L - b) \tag{14}$$

Low-productivity workers with $0 < p < b$ are unemployed because they have a high value of non-employment. Low-productivity workers with $b < p < w_L$ are willing to work, but the high minimum wage makes firms unwilling to hire them. In markets with $w_L < p < \underline{p}$ the lowest wage is equal to the minimum wage, and in markets with $p > \underline{p}$ the lowest wage is equal to the reservation wage which is increasing in p. In regime I, $r(p) < b$ and $\underline{w}(p) = w_L$ for all $p \geq w_L$. The equilibrium search model induces a positive correlation between p and the wage. Note that workers are always paid less than their marginal product.

We do not have complete information on all observations due to censoring or missing data. On the assumption that censoring is uninformative and that data are missing at random, one can easily modify the likelihood to deal with these complications. The likelihood of the sample is obtained by multiplication of the individual contributions.

Koning, Ridder and Van den Berg (1995) assume p to follow a lognormal distribution. Formal Chi-square tests that compare the fitted distribution to the observed wage earnings distribution lead to rejection of this distribution. It is not clear to what extent this affects the estimates of the parameter values of the model. Thus, it is interesting to assume that p follows a more flexible distribution.

5 Identification and semi-nonparametric estimation

5.1 Distinguishing between structural and frictional unemployment

The likelihood function combines two types of data: wage data and duration data. Both types of data play a different role in the identification of the parameters. First, we show that the fraction of (un)employed individuals and the unemployment spells just identify the structural unemployment rate. The rate of unemployment which can be identified from the fraction of unemployed individuals equals $u_S + (1 - u_S)\frac{\delta}{\delta+\lambda_0}$ (fractions of structural and frictional unemployment). Further, if there is permanent unemployment, then the distribution of unemployment durations is defective. However, if there is also frictional unemployment then one can estimate the re-employment hazard λ_0. Thus, in principle the unemployment spells identify u_S and λ_0. This, together with the information on labor market positions at the time of the first interview, just identifies δ.

The job durations are distributed as a non-scalar mixture of exponential distributions. The mixing distribution is that of p truncated at $\max(b, w_L)$. Given this distribution one can identify λ_1 and δ from the job durations and the subsequent destination. As a result, δ becomes overidentified.

Next we turn to the wage data. Recall from (11) that excess wages can be rewritten in terms of a fraction y of the excess productivity $p - \underline{w}(p)$. The distribution of y is independent of that of $p - \underline{w}(p)$. In fact, it can be shown that the earnings and accepted wage distribution also allow us to (over)identify λ_1/δ, but as shown in Van den Berg and Ridder (1993) there is not much information on this ratio in the wage data. As a result, the moments of the distribution of p truncated at $\max(b, w_L)$ from the earnings distribution and the distribution of accepted and paid wages. If we restrict the productivity distribution to belong to the class of recoverable distributions, the wage data (over)identify the (untruncated) distribution of p and hence u_S.

In sum, in the model the rate of structural is identified from three sources of information:
(*i*) the fraction of unemployment spells that are censored
(*ii*) the rate of unemployment
(*iii*) the distribution of observed wages.
The overidentification of the rate of structural unemployment can be used as a test on the model. In particular, we can perform an overidentifying restriction

test on $u_S = H(w_L)$ to detect a possible source of misspecification in the model. If we estimate the model with u_S as a free parameter, u_S will be identified from (*i*) and (*ii*), whereas $H(w_L)$ is identified from (*iii*). This is used in Section 7 as a misspecification test on the model.

Although the structural unemployment rate is identified, the effect of a change in the minimum wage on that rate is not identified without further assumptions. However, because we can identify the truncated (at the minimum wage) distribution of p from the wage data, we can recover the untruncated density in w_L by multiplication by $1 - H(w_L)$. Multiplication of the untruncated density $h(w_L)$ by w_L gives the semi-elasticity of the structural rate of unemployment with respect to the minimum wage. We can recover this semi-elasticity for the current minimum wage and higher levels of the minimum wage, but not for levels below the current minimum. Hence, we must invoke a continuity assumption to estimate the semi-elasticity for these lower levels.

5.2 Semi-Nonparametric estimation

Generally, the aim of semi- or nonparametric estimation methods is to weaken the impact of distributional assumptions on parameter estimates of a model. In the present context, if the productivity distribution is misspecified, this may affect the estimates of the structural parameters (λ_0, λ_1 and δ), as well as the rate of structural unemployment. As an alternative to the lognormal distribution, we propose the SNP approach developed by Gallant and Nychka (1987) to estimate the productivity distribution. The advantage of the SNP approach is that densities are specified as a very flexible functional form, which can approximate every density satisfying 'mild regularity conditions'. Basically, the unknown density function is approximated by a Hermite series. The fattest tails the SNP approach allows for are t-like, whereas the thinnest tails are thinner than normal ones. Further, the SNP density family allows for any sort of skewness or kurtosis. Due to the parameterization of the functional form, recoverability of the distribution is maintained.

Applications of the SNP approach include Gabler, Laisney and Lechner (1993), Melenberg and Van Soest (1993) and Van der Klaauw and Koning (1996).[4] Generally, it is found that the computational burden of the SNP approach increases with the number of terms in the polynomial (K). This is mainly due to numerical

[4] Gabler, Laisney and Lechner (1993) apply the semi-nonparametric approach to estimate a binary-choice model, whereas Melenberg and Van Soest (1993) and Van der Klaauw and Koning (1996) apply it to estimate a sample selection model.

problems involved in the computation of the density approximation, as well as in the optimization algorithm which is used to find the optimal parameter values.

In the SNP approach any smooth density can be approximated by a Hermite series. Our analysis starts from the lognormal distribution. For the distribution of p, we have

$$h(p) = \left[\Sigma_{i=0}^{K} \alpha_i \left(\frac{\ln(p) - \tau}{\eta} \right)^i \right]^2 \exp[-\frac{(\ln(p) - \tau)^2}{\eta^2}] \frac{1}{p} \tag{15}$$

This family of distributions is referred to as the family of SNP distributions. It is clear that the density is guaranteed to be nonnegative. Note that for $K = 0$, $\mu_p = \tau$, $\sigma_p = \frac{\eta}{\sqrt{2}}$ and $\alpha_0 = \pi^{-\frac{1}{4}}$, we have the lognormal form.[5]

To obtain a proper density, it is apparent that a restriction has to be imposed to ensure integration to 1. One way to do this is to derive an explicit restriction on the parameters of the density. For computational reasons, it is more convenient however to scale the improper density $h(p)$. Define

$$S = \int_0^\infty h(p) dp \tag{16}$$

and restrict the Hermite series to integrate to 1:

$$h^*(p) = \frac{h(p)}{S} \tag{17}$$

S can be computed easily using the moments of the (log)normal distribution. This also applies to the computation of moments of the truncated (log)normal distribution, as in our model. In Appendix A we present the recursion formulae which are used to compute the moments of the SNP distribution. Finally, it is clear that the proper density $h^*(p)$ remains unchanged if each α_i is multiplied by a scalar. Therefore we normalize by setting:

$$\alpha_0 = \pi^{-\frac{1}{4}} \tag{18}$$

This restriction ensures that S is equal to 1 for $K = 0$.

Gallant and Nychka show that the consistent estimation of an unknown density requires the number of terms in the Hermite polynomial K increases with the sample size. We assume that the true density is a member of the SNP class of densities, and treat the SNP specification as fully parametric. As a result, the standard framework of inference can be used to determine the significance

[5]In most applications the Hermite series is equivalent to the normal density for $K = 0$.

of higher order polynomial terms. This will also be the approach in the estimation of the equilibrium search model, where for $K = 0$ we have the lognormal distribution as a benchmark.[6]

In what follows we estimate the equilibrium search model of Koning, Ridder and Van den Berg (1995) semi-nonparametrically. To get an idea of the accuracy of the SNP method, and possible biases in the estimates of the structural parameters of the model, we first conduct two simulation experiments. Next, we estimate the model, using the SNP method. Special attention will be focused on the overidentifying restriction which follows from the imposed recoverability condition.

6 Simulation experiments

We conduct two Monte Carlo experiments to examine whether SNP estimation of the productivity distribution yields consistent estimates of the structural parameters of the model. Moreover, we analyze the effect of misspecification on the transition parameters (λ_0, λ_1 and δ) and the rate of structural unemployment, u_S, which equals $H(\max(b, w_L))$. In both experiments, we assume the following true parameter values: λ_0=0.05, λ_1=0.06, δ=0.003. Further, we set $w_L = 1400$ and assume $b < w_L$. Note that in general the parameter values we choose are close to the estimated parameter values in Koning, Ridder and Van den Berg (1995). Since $\lambda_0 < \lambda_1$, and from $b < w_L$, it follows that $r < b < w_L$ (Regime I). Consequently, the minimum wage is binding and structural unemployment equals $H(w_L)$.

In the first experiment, the productivity distribution H is assumed to be a mixture of two lognormal distributions. For instance, this may be the case if the labor market consists of two sectors with distinct productivity distributions.

$$H(p) = \gamma H_1(p) + (1 - \gamma) H_2(p) \quad \gamma \in [0, 1] \tag{19}$$

The parameters of H_1 and H_2 are denoted as μ_1, σ_1 and μ_2, σ_2. We set γ=0.5, μ_1=7.5, σ_1=0.15, μ_2=7.75, and σ_2=0.25. The implied fractions of structural and frictional unemployment equal 0.033 and 0.055, respectively. A single lognormal

[6] Gabler, Laisney and Lechner (1993) characterize a 'fully SNP approach' as an estimation procedure which includes the estimation of K. Because such a procedure does not exist, they propose to start the SNP estimation procedure with $K = 3$, and to test against higher order terms in the polynomial. They argue that for this number of terms in the polynomial the SNP density is very flexible, while the computational burden to find the optimal parameters is relatively small.

distribution fails to capture the peak at the mode of H, as well as the extreme right-hand tail, which is too thin compared to the mixture distribution. As we already stated in Koning, Ridder and Van den Berg (1995), the overall sample distribution of paid wages has the same properties.

In the second experiment we assume that the individual productivities are Pareto distributed with the minimum wage as the lower bound:

$$H(p) = 1 - \left(\frac{w_L}{p}\right)^{\theta} \quad p \geq w_L \tag{20}$$

$$= 0 \quad p < w_L \tag{21}$$

We set θ equal to 2.5. From the assumed c.d.f. it is clear that the probability mass below the minimum wage equals zero. As a result, all unemployment is frictional and equals 0.057. Flinn and Heckman (1982) show that the Pareto distribution does not belong to the class of recoverable distributions. Since the density is discontinuous in w_L, it is clear that this distribution belongs to a qualitatively different class of distributions than the family of SNP distributions. In this sense, the experiment can be considered as a worst case scenario. The family of Pareto distributions is one of the most popular distributions that is used to model wage offer distributions in a structural empirical search framework. Bontemps, Robin and Van den Berg (1997), who estimate an equilibrium search model in which the productivities of firms within a single labor market are allowed to differ, find that the right-hand tail of the estimated productivity distribution is similar to that of a Pareto distribution. It is interesting to examine how well the SNP density is able to approximate the Pareto distribution, as well as the implied rate of structural unemployment and examine the impact of misspecification on the parameters of the model.

Appendix B describes the data generating process of the experiments. Within each experiment we draw 100 samples with 1000 individual observations. We estimate the model semi-nonparametrically for $K = 0, 1, 2, 3, 4$. It should be stressed at this point that the unemployment durations we sample are right-censored if they exceed a period of 60 months. As a result, the fraction of unemployment durations that are censored (over)identifies the rate of structural unemployment.

The means of the estimated parameters obtained for the two designs are presented in Tables 1 and 2.[7] What is clear from the first experiment is that

[7] The optimization algorithm did not converge for all replications. In such cases, the tolerance criterion for the gradients was not satisfied after 100 iterations, and the final iteration values were not included in the MC results. However, those values did not differ substantially from those that were obtained when convergence was achieved within 100 iterations.

the SNP approach improves the fit to the data significantly up to 3 terms in the polynomial (at a 5% significance level). Figures 1 and 2 show that for $K > 1$ the estimated density of the productivity function mimics the true density very well. All structural parameters converge to their true values with an increasing number of terms in the polynomial. λ_1 is underestimated for $K = 0$ (lognormal distribution), and this negative bias vanishes for $K > 1$. Further, the implied probability mass below the minimum wage is overestimated for $K = 0$, but for $K = 3$ it equals its true value as well. The estimated semi-elasticity of the rate of structural unemployment with respect to the minimum wage (which equals $h(w_L)w_L$) is estimated quite well for all K. This semi-elasticity can be interpreted as a check on the fit of the density at the truncation point, which is the minimum wage.

In the second experiment, we find the results to be less satisfactory (see Table 2, as well as Figures 3 and 4). Although the fit of the model increases with the number of terms in the polynomial in the SNP specification, the estimated parameters do not converge to their true values. The estimated rate of structural unemployment decreases from 15.1% for $K = 0$ to 8.3% for $K = 4$. For all K the model is not able to explain the size of the peak at the minimum wage, whereas in general the fitted density is too fat in the area to the right of the peak. From this we conclude that the family of SNP densities is incapable of fitting the true Pareto density. This seems to affect the estimates of the structural parameters λ_1 and δ in two ways. First, as we have argued in the previous sections, the moments of the wage earnings distribution are determined by the moments of the productivity distribution and λ_1/δ. Presumably the misspecification of the distribution of productivities is partly compensated by a decrease in λ_1/δ to obtain a better fit to the wage data. Second, the individual productivity levels that exceed the minimum wage are overestimated. Thus, individual workers seem to earn low wages, compared to their maximum wage earnings, which lie close to their productivity level. This means that the acceptance probability of job offers for employed workers is overestimated. as they have relatively low positions on the wage ladder. As a result, the exit rate into another job becomes overestimated, and a decrease in λ_1, the job offer arrival rate, is needed to compensate for this. If we include more terms in the polynomial the negative bias in λ_1 remains unaffected, or even increases. Also the estimates of λ_0 and δ are biased, and the inclusion of additional terms in the polynomial does not help in obtaining more accurate estimates of these parameters.

The results from the experiments suggest that the SNP estimation method is a useful tool to detect and minimize any sources of misspecification. provided that the 'true' productivity distribution is continuous over the whole support. The misspecification of the productivity distributions seems to affect the estimates of λ_1 and the rate of structural unemployment in particular. It is interesting to compare these results with the results obtained from the SNP estimation with the observed data.

7 Estimation results

Table 3 presents the parameter estimates of the model based on the whole sample for $K = 0, 1, 2, 3, 4$.[8] Figures 5 and 6 plot the fitted and observed wage earnings distribution, respectively, for the whole sample. It is clear that the inclusion of extra terms in the polynomial improves the fit of the model substantially. In particular, it turns out that for $K > 1$ the SNP specification is capable of capturing the peak at the mode of the sample distribution, as well as the relatively fat right-hand tail (see Figures 5 and 6). Considering the sensitivity of the structural parameters to the assumed lognormal distribution, the estimate of λ_1 is most affected by misspecification of the productivity distribution. Generally, λ_1 increases if we allow the distribution to become more flexible. This finding is in line with the results obtained from the first simulation experiment. For a sufficient number of terms in the polynomial the parameter estimate of λ_1 gets close to or even exceeds the estimate of λ_0.

The estimated models can also be used to examine the robustness of the estimated rate of structural unemployment. This rate decreases from 5.2% for $K = 0$ (lognormal) to 1.4% for $K = 4$. This indicates that the estimate of u_S is sensitive to functional form assumptions. Similar evidence is found for the implied semi-elasticity of the rate of structural unemployment with respect to the minimum wage, which decreases from 0.41 for $K = 0$ to 0.21 for $K = 4$ (we return to the latter result below).

We also estimate the equilibrium search model with u_S as a free parameter. Thus we do not impose the restriction $H(w_L) = u_S$ that follows from the recoverability of H. This enables us to test the overidentifying restriction on the rate of structural unemployment for different K. Table 4 presents the estimation results of this version of the model for K =0, 2, and 3. For $K = 1$ and $K = 4$ the parameter values do not converge. As stated above, in the model with $u_S = H(w_L)$ convergence is achieved up to $K = 4$. This finding suggests that the recoverability of the productivity distribution implies a restriction, and an extra

[8]For $K = 5$ the optimization algorithm did not converge after 100 iterations.

parameter is needed to take account of this. Indeed, combining Tables 3 and 4, Likelihood Ratio tests on the (overidentifying) restriction on the rate of structural unemployment are rejected for all SNP specifications. If the productivity distribution is assumed to be lognormal ($K = 0$), the estimated rate which is based on the recovered productivity distribution equals 14% ($H(w_L)$), whereas the information on the fraction of unemployed and the unemployment durations predicts it to equal 1%.

From Table 4 we see that for an increasing K the rate of structural unemployment seems to converge to the rate based on the unemployment (duration) data. This is to be expected, as the more flexible SNP form of the productivity density allows for a smaller probability mass below the minimum wage. Further, the results in Table 4 indicate that the rate of structural unemployment which is based on the recovered productivity distribution is extremely sensitive to the number of terms in the polynomial. For example, for $K = 2$ the implied rate equals 3.4%, whereas for $K = 3$ it equals 55%.

Clearly, the SNP approach helps in estimating the truncated productivity distribution more accurately, but additional information is needed for a more accurate estimate of the probability mass of H below the minimum wage. This picture is confirmed if we look at the estimation results for the age group 39-61 (Tables 5 and 6, and Figures 7 and 8). In Koning, Ridder and Van den Berg (1995) it was found that this group has a relatively large rate of structural unemployment (5.2%). This is mainly due to the relatively high variance of productivities for this age group. We examine the robustness of this result by following the SNP approach. Again, the inclusion of additional terms in the polynomial improves the likelihood of the model up to $K = 4$, and the estimated rate of structural unemployment converges to the rate which is predicted by the fraction of unemployment and the fraction of censored unemployment spells (1.4%).

From the results it seems that frictional unemployment is more important than structural unemployment. However, this should be taken with caution, as structurally unemployed individuals may well be underrepresented in our sample. These individuals will never find a job, so they may classify themselves as being a nonparticipant when being questioned on their labor market state. In that case they are not included in the sample, so that the sample unemployment rate is an underestimate of the total unemployment rate. Including nonparticipants in the sample is feasible, but in that case one would have to make a distinction between individuals who do not participate by choice, and individuals who would participate if the wage floor were not binding. The former are what one could call genuine nonparticipants, whereas the latter are structurally unemployed. The current data do not allow for such a distinction.

Finally, let us turn to the implied effect of increases of the minimum wage on the rate of structural unemployment. With a lognormal (or similar) specification for the productivity distribution, it is conceivable that the estimated shape of the wage density just above the minimum wage is affected by the requirement that the structural unemployment rate should fit the wage probability mass below the minimum wage. With the SNP approach, such a problem does not exist. It turns out that, indeed, the effect is overestimated under the lognormality assumption. The SNP approach leads to a semi-elasticity of around 0.2 for the whole sample. This means that if the minimum wage increases by 10% then the structural unemployment rate increases by 2 percent points.[9]

8 Conclusion

In this paper we have used the SNP (semi-nonparametric) approach developed by Gallant and Nychka (1987) to examine the robustness of the estimation results of the equilibrium search model of Koning, Ridder and Van den Berg (1995). This model is based on the equilibrium search model as developed by Burdett and Mortensen, and it is extended in such a way that we can distinguish between frictional and structural unemployment. Structural unemployment results from the allowance of productivity heterogeneity of workers. In the homogeneous model, unemployment is frictional, originating from the fact that people have to wait some time to find a job. In the heterogeneous model, individuals are permanently (structurally) unemployed if their productivity is smaller than the mandatory minimum wage.

The advantage of the SNP approach is that densities can be specified by a very flexible form, which can approximate any "proper" density. This helps us to reduce the impact of misspecification in the model, *i.e.* misspecification of the productivity distribution of individual workers. The rate of structural unemployment is particularly sensitive to such misspecification, as it is overidentified from the wage data if the productivity distribution has a particular functional form.

[9] As was to be expected, our estimate is more in line with the estimates in Van den Berg and Ridder (1998) than with those in Koning, Ridder and Van den Berg (1995). Recall that the latter study assumes a lognormal productivity distribution whereas the former assumes a flexible discrete productivity distribution. The setup in the former study is different in a number of additional respects (see above). The estimated minimum wage effects in the present study are also in line with Van Soest (1989), who uses a different methodology as well as a different Dutch data set.

The estimation results as well as the simulation exercises indicate that the SNP approach serves as a very useful tool to reduce the impact of distributional assumptions on the structural estimates of the model. If we allow for a sufficiently flexible SNP specification, then the implied rate of structural unemployment is determined by unemployment-related data information, notably the fraction of censored unemployment spells and the fraction of unemployed workers. The results of the simulation experiments show that the rate of structural unemployment is well identified with such data. The structural unemployment estimates that are obtained from extrapolation of the truncated productivity distribution are found to very sensitive with respect to functional form assumptions on this distribution. The estimate of the structural unemployment rate corresponding to the latter approach is often estimated to be implausibly high.

The results indicate that most unemployment in the sample is due to labor market frictions. However, it should be kept in mind that structurally unemployed individuals presumably are underrepresented in samples like ours. Specifically, since these individuals will never find a job, they may classify themselves as non-participants. The current data do not allow us to identify which self-assigned non-participants are structurally unemployed.

Table 1: Monte Carlo estimation results of mixture of lognormal distributions, 100 replications, 1000 observations, K=0, K=1, K=2, K=3, and K=4. Standard errors of average estimated parameter values between parentheses

	true value	K=0	K=1	K=2	K=3	K=4
λ_0	0.050	0.050	0.050	0.050	0.050	0.050
		(0.00014)	(0.00014)	(0.00016)	(0.00016)	(0.00016)
λ_1	0.060	0.047	0.052	0.061	0.061	0.060
		(0.00096)	(0.0011)	(0.0015)	(0.0014)	(0.0014)
δ	0.0030	0.0031	0.0030	0.0030	0.0030	0.0030
		(1.21e-5)	(1.16e-5)	(1.33e-5)	(1.24e-5)	(1.23e-5)
τ		7.63	7.37	7.75	7.62	7.52
		(0.00076)	(0.0010)	(0.0017)	(0.0070)	(0.0076)
η		0.34	0.40	0.26	0.29	0.30
		(0.00078)	(0.00092)	(0.0012)	(0.0023)	(0.0026)
α_1		-	1.11	-0.73	-0.37	0.15
			(0.0035)	(0.010)	(0.023)	(0.051)
α_2		-	-	0.81	0.23	-0.18
				(0.013)	(0.031)	(0.026)
α_3		-	-	-	0.23	0.22
					(0.013)	(0.013)
α_4		-	-	-	-	0.091
						(0.011)
$H(w_L)$	0.033	0.055	0.046	0.045	0.033	0.026
$h(w_L)w_L$	0.41	0.46	0.44	0.51	0.48	0.39
u_{fric}	0.055	0.054	0.055	0.053	0.055	0.055
u_{tot}	0.088	0.109	0.101	0.098	0.088	0.081
Log L		-17050.6	-17040.0	-17030.4	-17028.3	-17027.3
convergence within 100 iterations[a]		98	98	95	88	89

[a]: We excluded the parameter values from the sample in case of non-convergence

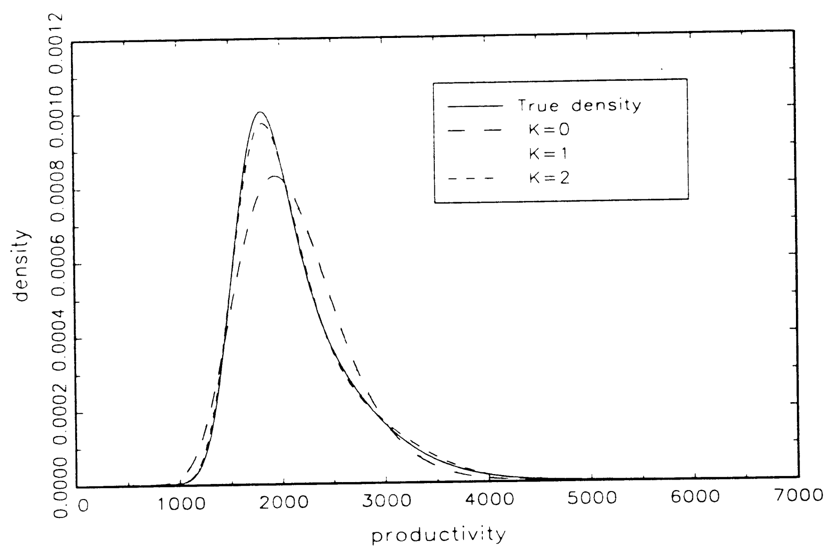

Figure 1: Simulation experiment with mixture of lognormal distributions: Fitted and true productivity distribution for K=0, K=1, and K=2.

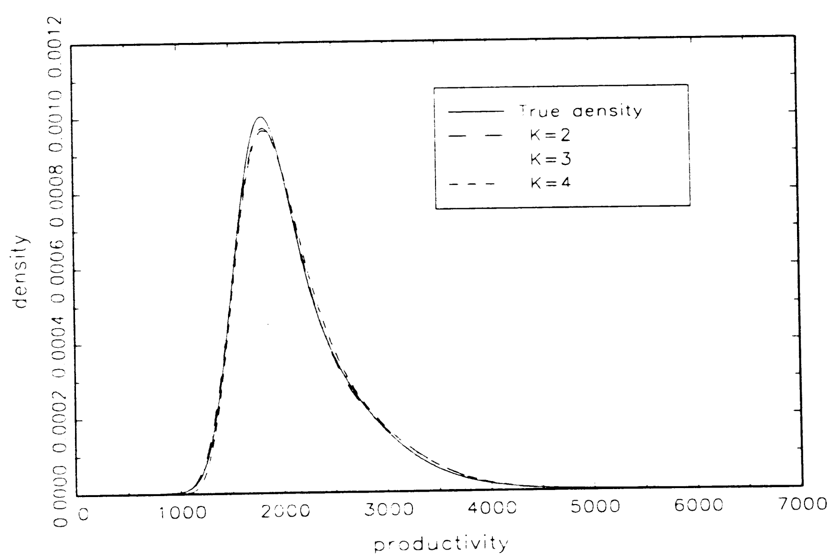

Figure 2: Simulation experiment with mixture of lognormal distributions: Fitted and true productivity distribution for K=2, K=3, and K=4.

Table 2: Monte Carlo estimation results of Pareto distribution, 100 replications, 1000 observations, K=0, K=1, $K = 2$, K=3, and K=4; standard errors of average estimated parameter values between parentheses

	true value	K=0	K=1	K=2	K=3	K=4
λ_0	0.050	0.051	0.051	0.051	0.052	0.052
		(0.00017)	(0.00017)	(0.00018)	(0.00017)	(0.00017)
λ_1	0.060	0.036	0.035	0.033	0.033	0.033
		(0.00035)	(0.00036)	(0.00057)	(0.00045)	(0.00048)
δ	0.0030	0.0029	0.0029	0.0029	0.0029	0.0029
		(1.34e-5)	(1.38e-5)	(1.70e-5)	(1.42e-5)	(1.36e-5)
τ		7.60	7.23	7.83	7.62	7.42
		(0.0012)	(0.0013)	(0.0017)	(0.0018)	(0.0013)
η		0.49	0.56	0.34	0.38	0.40
		(0.0016)	(0.0018)	(0.0010)	(0.0016)	(0.0017)
α_1		-	1.00	-1.11	-0.80	0.21
			(0.0016)	(0.0087)	(0.0074)	(0.0051)
α_2		-	-	1.43	0.091	-0.77
				(0.012)	(0.0034)	(0.0036)
α_3		-	-	-	0.48	0.20
					(0.0033)	(0.0023)
α_4		-	-	-	-	0.22
						(0.0013)
$H(w_L)$	0.000	0.151	0.135	0.105	0.089	0.083
$h(w_L)w_L$		0.68	0.71	0.82	0.85	0.87
u_{fric}	0.057	0.046	0.047	0.048	0.049	0.049
u_{tot}	0.057	0.197	0.182	0.153	0.138	0.132
Log$L$		-17614.7	-17586.1	-17491.2	-17475.6	-17466.5
convergence within 100 iterations		98	98	98	97	95

[a]: We excluded the parameter values from the sample in case of non-convergence

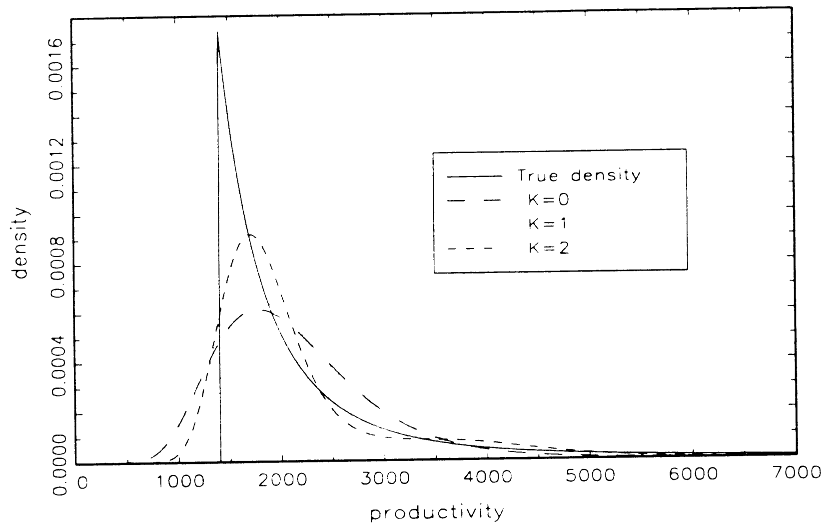

Figure 3: Simulation experiment with Pareto distribution: Fitted and true productivity distribution for K=0, K=1, and K=2.

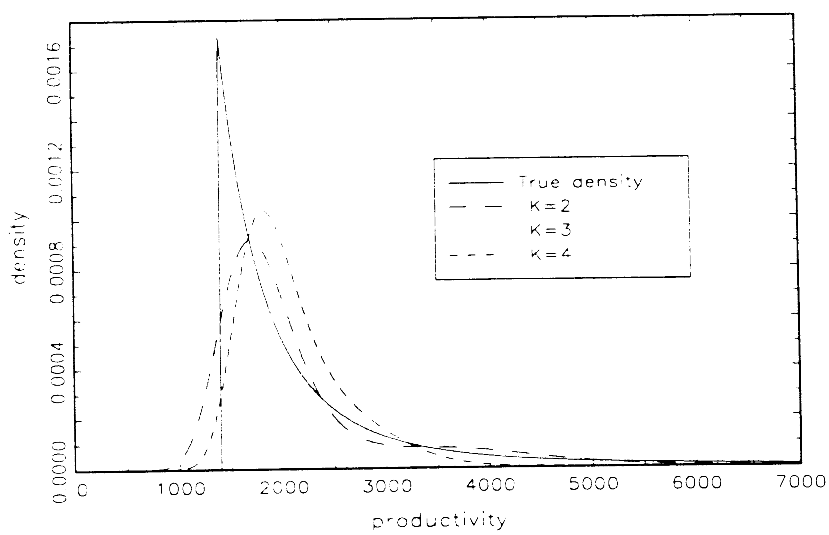

Figure 4: Simulation experiment with Pareto distribution: Fitted and true productivity distribution for K=2, K=3, and K=4.

Table 3: Parameter estimates for the whole sample; rates are per month, standard errors between parentheses, N=1767

Specification	K=0	K=1	K=2	K=3	K=4
λ_0	0.036	0.038	0.039	0.038	0.037
	(0.0019)	(0.0019)	(0.0019)	(0.0019)	(0.0019)
λ_1	0.023	0.029	0.031	0.036	0.038
	(0.0024)	(0.0038)	(0.0042)	(0.0048)	(0.0056)
δ	0.0033	0.0034	0.0034	0.0033	0.0033
	(0.00015)	(0.00015)	(0.00015)	(0.00015)	(0.00015)
τ	7.69	7.40	7.88	7.69	7.52
	(0.010)	(0.018)	(0.010)	(0.032)	(0.021)
η	0.40	0.44	0.29	0.33	0.36
	(0.0071)	(0.0090)	(0.0056)	(0.011)	(0.011)
α_1	-	1.06	-1.10	-0.65	0.39
		(0.076)	(0.075)	(0.14)	(0.15)
α_2	-	-	1.18	0.022	-0.68
			(0.11)	(0.16)	(0.065)
α_3	-	-	-	0.39	0.14
				(0.048)	(0.054)
α_4	-	-	-	-	0.19
					(0.025)
$H(w_L)$	0.052	0.037	0.021	0.015	0.014
$h(w_L)w_L$	0.41	0.35	0.26	0.22	0.21
u_{fric}	0.079	0.079	0.080	0.079	0.079
u_{tot}	0.131	0.115	0.101	0.095	0.093
u_{obs}	0.118	0.118	0.118	0.118	0.118
Log L	-21669.4	-21633.6	-21558.8	-21544.8	-21541.1

Table 4: Parameter estimates for whole sample, u_S as a free parameter; rates are per month, standard errors between parentheses, N=1767

Specification	K=0	K=2	K=3
λ_0	0.037	0.037	0.037
	(0.0019)	(0.0020)	(0.0020)
λ_1	0.041	0.040	0.041
	(0.0066)	(0.0062)	(0.0064)
δ	0.0032	0.0033	0.0032
	(0.00015)	(0.00012)	(0.00012)
τ	7.60	7.87	7.48
	(0.017)	(0.014)	(0.029)
η	0.47	0.31	0.41
	(0.015)	(0.010)	(0.038)
α_1	-	-1.05	0.54
		(0.074)	(0.24)
α_2	-	1.03	-1.12
		(0.11)	(0.21)
α_3	-	-	0.63
			(0.068)
$H(w_l)$	0.138	0.034	0.554
$h(w_L)w_L$	0.68	0.17	0.15
u_S	0.010	0.010	0.010
	(0.0029)	(0.0029)	(0.0029)
u_{fric}	0.080	0.080	0.080
u_{tot}	0.090	0.090	0.090
u_{obs}	0.118	0.118	0.118
Log L	-21604.6	-21548.9	-21537.2

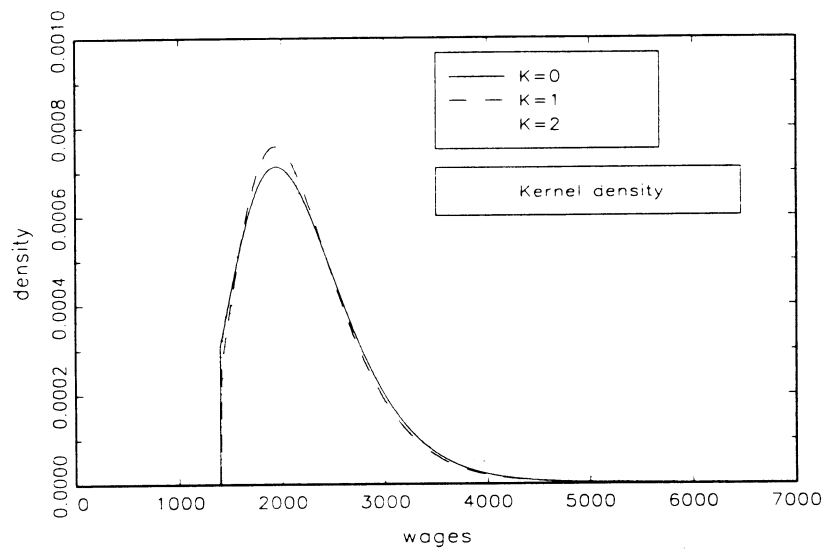

Figure 5: Fitted and nonparametric wage earnings density for whole sample; K=0, K=1 and K=2 (standard normal kernel, h=92.1)

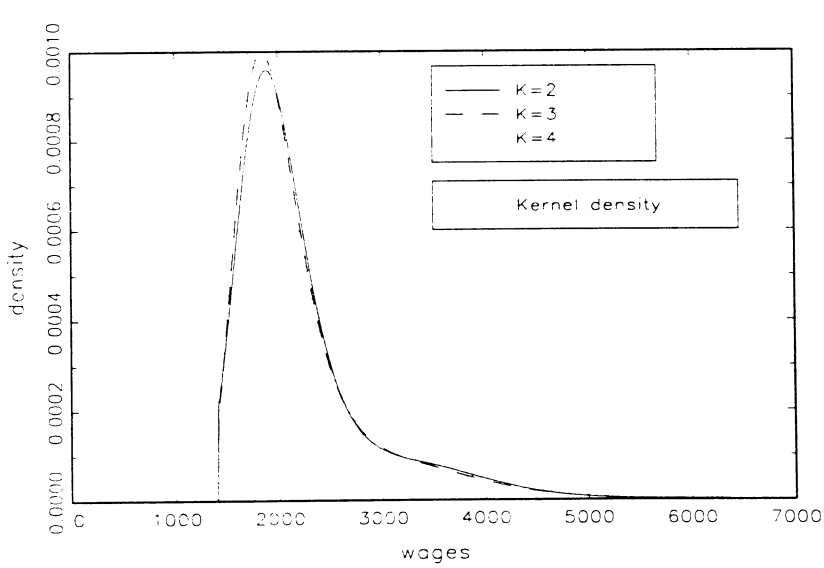

Figure 6: Fitted and nonparametric wage earnings density for whole sample; K=2, K=3 and K=4 (standard normal kernel, h=92.1)

Table 5: Parameter estimates for age category 39-61, rates are per month, standard errors between parentheses, N=635

Specification	K=0	K=1	K=2	K=3	K=4
λ_0	0.030	0.030	0.030	0.030	0.030
	(0.0029)	(0.0028)	(0.0027)	(0.0027)	(0.0027)
λ_1	0.0093	0.0090	0.0089	0.0089	0.0088
	(0.0022)	(0.0019)	(0.0017)	(0.0017)	(0.0016)
δ	0.0022	0.0022	0.0022	0.0022	0.0022
	(0.00021)	(0.00016)	(0.00016)	(0.00016)	(0.00015)
τ	7.79	7.47	7.93	7.73	8.05
	(0.017)	(0.029)	(0.010)	(0.016)	(0.029)
η	0.44	0.50	0.29	0.33	0.26
	(0.016)	(0.017)	(0.011)	(0.012)	(0.011)
α_1	-	1.00	-0.69	-0.53	0.099
		(0.13)	(0.092)	(0.10)	(0.18)
α_2	-	-	1.19	0.080	0.30
			(0.18)	(0.056)	(0.27)
α_3	-	-	-	0.52	-0.72
				(0.060)	(0.10)
α_4	-	-	-	-	0.45
					(0.067)
$H(w_L)$	0.050	0.034	0.017	0.016	0.010
$h(w_L)w_L$	0.30	0.28	0.22	0.10	0.16
u_{fric}	0.065	0.066	0.068	0.067	0.067
u_{tot}	0.115	0.101	0.085	0.083	0.077
u_{obs}	0.113	0.113	0.113	0.113	0.113
Log L	-7494.1	-7463.5	-7440.8	-7433.6	-7429.2

Table 6: Parameter estimates for age category 39-61, u_S as a free parameter; rates per month, standard errors between parentheses, N=635

Specification	K=0	K=2	K=3
λ_0	0.030	0.030	0.030
	(0.0029)	(0.0027)	(0.0027)
λ_1	0.0090	0.0089	0.0089
	(0.0021)	(0.0017)	(0.0017)
δ	0.0022	0.0022	0.0022
	(0.00020)	(0.00016)	(0.00016)
τ	7.74	7.93	7.41
	(0.028)	(0.018)	(0.050)
η	0.50	0.30	0.43
	(0.026)	(0.014)	(0.024)
α_1	-	-0.68	2.73
		(0.095)	(1.70)
α_2	-	1.13	-3.48
		(0.19)	(1.72)
α_3	-	-	1.72
			(0.66)
$H(w_L)$	0.098	0.020	0.90
$h(w_L)w_L$	0.53	0.099	0.00
u_S	0.014	0.014	0.014
	(0.0051)	(0.0052)	(0.0052)
u_{fric}	0.066	0.068	0.068
u_{tot}	0.080	0.082	0.082
u_{obs}	0.113	0.113	0.113
Log L	-7477.8	-7440.4	-7426.9

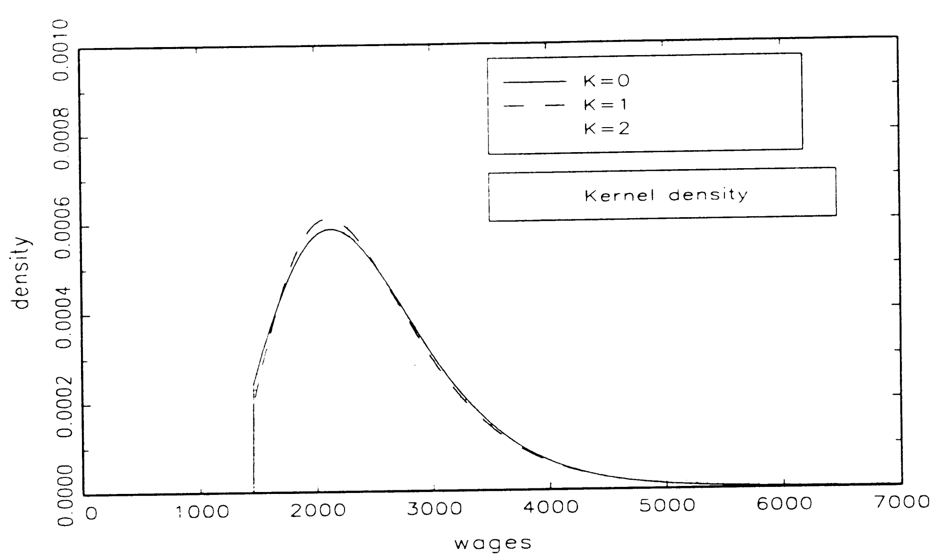

Figure 7: Fitted and nonparametric wage earnings density for age 39-61; K=0, K=1 and K=2 (standard normal kernel, h=218.7)

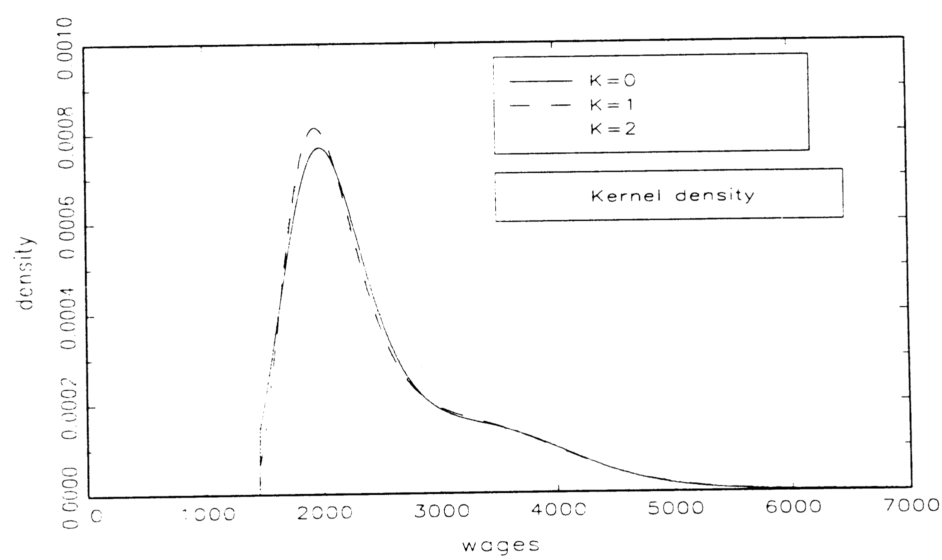

Figure 8: Fitted and nonparametric wage earnings distribution for age 39-61; K=2, K=3 and K=4 (standard normal kernel, h=218.7)

Appendix A. Recursion formulae of the truncated normal distribution

Define

$$u = \frac{\ln(p) - \tau}{\eta}$$

and

$$I_k(a,b) = \int_a^b u^k \exp(-u^2)\, du \qquad b \geq a \geq 0$$

We denote $\Phi(.)$ as the c.d.f. of the standard normal distribution. Using partial integration, we obtain the following recursion formulae:

$$\begin{aligned} I_0(a,b) &= \sqrt{\pi}\ \left(\Phi(\sqrt{2}b) - \Phi(\sqrt{2}a)\right) \\ I_1(a,b) &= -\frac{1}{2}\ \left(\exp(-a^2) - \exp(-b^2)\right) \\ I_k(a,b) &= \frac{1}{2}\left[(k-1)I_{k-2}(a,b) + a^{k-1}\exp(-a^2) - b^{k-1}\exp(-b^2)\right] \quad k \geq 2 \end{aligned}$$

Note that S can be easily computed with the above formulae, with $(a,b) = (-\infty, \infty)$. To compute the moments of the productivity distribution, we use $(a,b) = (\frac{\ln(w_L)-\tau}{\eta}, \infty)$ as the support.

Appendix B. The data generating process of the simulation experiments

In this appendix we briefly discuss the procedure that is used to draw the observations on individual labor market histories. The procedure is as follows:

(i): We draw individual worker productivity p out of productivity distribution H.

(ii): If $p < w_L$, a worker is structurally unemployed (x=1), and the corresponding unemployment spell is right-censored at 60 months. If $p \geq w_L$, a worker is either frictionally unemployed with probability $\delta/(\delta+\lambda_0)$, or employed with probability $\lambda_0/(\delta+\lambda_0)$.

$(iii-a)$: If an individual worker (with $p \geq w_L$) is frictionally unemployed ($x = 1$), we draw an unemployment spell t_{0f}. These spells are exponentially distributed with parameter λ_0. We apply right-censoring if the unemployment spell lasts longer than 60 months. Next, for individuals with uncensored spells we draw the accepted wage out of unemployment, w_0, from $F(w|p)$.

$(iii-b)$: If an individual worker (with $p \geq w_L$) is employed (x=0), we draw the individual wage observation w_1 out of $G(w|p)$. Given w_1 and p, we know the individual exit rate $\delta+\lambda_1\overline{F}(w_1|p)$. Thus we are in the position to draw job spells t_{1f}, as well as the destinations following these job spells, v.

References

Albæk, K. and E.S. Madsen (1987), "Wage distributions and employment effects of the Danish minimum wage", in R. Lund and P. Pederson, editors, *Studies in unemployment*, Samfundslitteratur.

Bontemps, C., J.M. Robin, and G.J. van den Berg (1997), "Equilibrium search with productivity dispersion", Discussion Paper TI 97–081/3, Tinbergen Institute.

Burdett, K. and D.T. Mortensen (1998), "Wage differentials, employer size, and unemployment", *International Economic Review*, forthcoming.

Card, D. and A.B. Krueger (1995), *Myth and Measurement: The New Economics of the Minimum Wage*, Princeton University Press, Princeton.

Dickens, R., S. Machin, and A. Manning (1994), "Estimating the effect of minimum wages on employment from the distribution of wages: A critical review", Technical report, Centre for Economic Performance.

Dolado, J., F. Kramarz, S. Machin, A. Manning, D.N. Margolis, and C.N. Teulings (1996), "Minimum wages: The European experience", *Economic Policy*, 319–372.

Flinn, C.J. and J.J. Heckman (1982), "New methods for analyzing structural models of labor force dynamics", *Journal of Econometrics*, 18, 115–168.

Gabler, S., F. Laisney, and M. Lechner (1993), "Seminonparametric estimation of binary–choice models with an application to labor–force participation", *Journal of Business and Economic Statistics*, 11, 61–80.

Gallant, A.R. and D.W. Nychka (1987), "Semi–nonparametric maximum likelihood estimation", *Econometrica*, 55, 363–390.

Koning, P., G. Ridder, and G.J. van den Berg (1995), "Structural and frictional unemployment in an equilibrium search model with heterogeneous agents", *Journal of Applied Econometrics*, 10, S133–S151.

Lindeboom, M. and J. Theeuwes (1991), "Job duration in the Netherlands", *Oxford Bulletin of Economics and Statistics*, 53, 243–264.

Melenberg, B. and A. van Soest (1993), "Semi–parametric estimation of the sample selection model", Technical report, CentER for Economic Research.

Meyer, R.H. and D.A. Wise (1983a), "Discontinuous distributions and missing persons: The minimum wage and unemployed youth", *Journal of Labor Economics*, 1, 66–100.

Meyer, R.H. and D.A. Wise (1983b), "The effect of the minimum wage on the employment and earnings of youth", *Econometrica*, 51, 1677–1698.

Mortensen, D.T. (1986), "Job search and labor market analysis", in O. Ashenfelter and R. Layard, editors, *Handbook of Labor Economics*, volume II, 849–919, Elsevier Science Publishers BV.

Narendranathan, W. and S. Nickell (1985), "Modelling the process of job search", *Journal of Econometrics*, 28, 28–49.

Ridder, G. (1984), "The distribution of single–spell duration data", in G.R. Neumann and N. Westergård-Nielsen, editors, *Studies in Labor Market Analysis*, Springer–Verlag.

Van den Berg, G.J. (1990), "Search behaviour, transitions to nonparticipation and the duration of unemployment", *The Economic Journal*, 100, 842–865.

Van den Berg, G.J. (1992), "A structural dynamic analysis of job turnover and the costs associated with moving to another jobs", *Economic Journal*, 100, 1116–1133.

Van den Berg, G.J. and G. Ridder (1993), "Estimating an equilibrium search model from wage data", in H. Bunzel, editor, *Panel Data and Labor Market Analysis*, North Holland, Amsterdam.

Van den Berg, G.J. and G. Ridder (1998), "An empirical equilibrium search model of the labour market", *Econometrica*, forthcoming.

Van der Klaauw, B. and R.H. Koning (1996), "Some applications of semi–nonparametric maximum likelihood estimation", Discussion paper 96–161/7, Tinbergen Institute.

Van Soest, A. (1989), "Minimum wage rates and unemployment in the Netherlands", *De Economist*, 137, 279–308.