Using Firm Data to assess the Performance of Equilibrium Search Models of the Labor Market

Gerard J. van den Berg
Aico van Vuuren

Using firm data to assess the performance of equilibrium search models of the labor market

Gerard J. van den Berg [*]   Aico van Vuuren [†]

June 1, 2000

Abstract

Equilibrium search models are useful tools for the evaluation of labor market policies. Recently developed equilibrium search models of the labor market are able to fit the wage distribution perfectly with longitudinal labor supply data, by estimating an appropriate distribution of labor productivity across firms. This paper formally compares such structural estimates to their directly observed counterparts in firm data. More generally, we investigate the extent to which these models are able to explain the observed distributions of wages, productivities and firm sizes across firms, as well as the extent to which they are able to explain the observed relationships between these variables across firms. The parameters that capture search frictions are estimated with worker data that are matched to the firm data.

1 Introduction

During the past two decades, equilibrium search models have become important tools for the analysis of labor market issues. These models deal with wage determination in the presence of informational frictions or search frictions (see Mortensen and Pissarides, 1999, and Van den Berg, 1999, for overviews of the theoretical and empirical literature, respectively). As such they explain wages, individual transitions between different labor market states, the durations spent in those states, and, last but not least, aggregate numbers like the unemployment rate and worker reallocation rates (Ridder and Van den Berg, 2000). Equilibrium search models are particularly useful to study policy changes, like changes in the minimum wage or the unemployment benefits level. Structural estimates of these models enables policy inference that is not subject to the Lucas critique.

Unfortunately, most existing equilibrium search models have difficulty explaining the cross-sectional distribution of wage data in a given labor market (see the above surveys; see also the discussions in Eckstein and Wolpin, 1990, and Van den Berg and Ridder, 1998). Recently, however, Bontemps, Robin and Van den Berg (2000) (BRVdB) developed and estimated an equilibrium search model that is able to provide a perfect fit to the wage distribution. The model extends the Burdett and Mortensen (1998) model with homogeneous workers and firms by allowing firms acting in a given market to have different labor productivities (or revenue products). Equilibrium search models are typically estimated with longitudinal labor supply data, covering unemployment and job durations, transitions from one job to another or to unemployment, and wages of employed workers. With such data, firm heterogeneity in the BRVdB model is unobserved heterogeneity, and BRVdB show that any wage distribution can be supported by an appropriate underlying continuous productivity distribution, provided that the former is in the set of wage distributions that can be generated by the model. The set of wage distributions that can be generated is a subset of the set of all distributions of nonnnegative random variables. This means that the empirical wage distribution can be used for a nonparametric specification test. However, if the test accepts the specification then there are basically no overidentifying restrictions on the wage distribution anymore (unless the data contain sufficiently rich additional wage data, for example on wages that are accepted by employed workers moving to another job in due course).

The BRVdB model has become rather popular in the literature. Moreover, most other recent empirical analyses also adopt heterogeneity of firms' productivity in order to obtain an acceptable fit to the wage data (see e.g. Bowlus, Kiefer

and Neumann, 2001). In this paper we examine the performance of the BRVdB model when confronted to *firm data*. In firm data, the firm's labor productivity is observed, and the distribution of these observations can be compared to the estimated distribution from more traditional data (on wages and individual labor market spells). More generally, we provide a comprehensive formal investigation of the extent to which the BRVdB model is able to explain the observed distributions of wages, productivities and firm sizes across firms, as well as the extent to which the model is able to explain the observed relationships between these variables across firms.

This paper thus provides the first formal analysis of the performance of such models when confronted to firm data.[1] Given the lack of overidentifying information in the data that are traditionally used, firm data may be useful to assess the performance of these models. A close inspection of the fit to firm data may generate research questions for future theoretical research. For example, the theoretical relation between the productivity and the wage reflects the first-order condition of the optimization problem of the firm, given the behavior of all other firms and all workers. If the observed relation differs from the predicted relation in a certain data interval then this may indicate that important aspects of the optimization problem of the corresponding firms are ignored in the model.

We use data from the Netherlands concerning three different industries: (1) textile, (2) publishing and printing, and (3) electrical machinery. As we shall see, these industries have quite different characteristics. We basically use two administrative data sets. The Production Survey covers the firm data, and these are matched to the Wage and Employment Survey, which contains information on job mobility and wages of individual workers. Analogously to BRVdB, the parameters that capture search frictions are estimated with the latter type of data. These parameters serve as inputs in the main empirical analyses. We show that these parameters suffice to characterize the distribution of firm sizes. BRVdB show that the frictional parameters together with data on the wage distribution enable one to back out the underlying productivity distribution. More precisely, if the frictional parameters and one of these distributions are known, then for each given wage (productivity) level, the corresponding productivity (wage) level and firm size can be calculated. In this paper we use all of this to compare

---

[1]It should be noted that BRVdB make a graphical comparison between the distributions of wages, productivities and firm sizes as estimated from labor force survey data on the one hand, and the observed distribution of these variables in firm data on the other hand. They also graphically compare the corresponding relations between the variables. We briefly discuss these results below. Robin and Roux (1998) and Postel-Vinay and Robin (2000) estimate particular equilibrium search models with joined datasets on workers and firms.

predictions to observations.

The paper is organized as follows. The next section summarizes the model framework as introduced by BRVdB. Section 3 discusses the data used in this paper. The empirical implementation is discussed in section 4. Results follow in section 5. Finally, we give our conclusions in section 6.

2 Equilibrium search theory

Since the model we use in this paper is already discussed in detail by BRVdB, the exposition of this model will be brief.

Workers seek to maximize their expected steady-state discounted future income. The opportunity cost of employment is denoted by b and is assumed to be constant across individuals. Job offers arrive according to a Poisson process with a constant rate λ_0 when unemployed. When employed, job offers arrive at the constant rate λ_1. Whenever a job offer arrives, the decision has to be made whether to accept it or to reject it and search further for a better offer. Layoffs arrive at the constant rate δ. The distribution of wage offers F is independent of the current state of the job searcher (employed or unemployed). Let supp(F) denote the support of F. We denote $\underline{w} = \inf[\text{supp}(F)]$ and $\overline{w} = \sup[\text{supp}(F)]$. The legal minimum wage, not necessarily positive, is equal to w_{min}. Necessarily then, $\underline{w} \geq w_{min}$. A well-known result is that the optimal strategy when unemployed is to accept any wage offer w greater or equal to the reservation wage (Burdett and Mortensen, 1998). For convenience, we assume the reservation wage to be lower than the minimum wage, which means that any wage offer is acceptable for the unemployed. This assumption seems to be reasonable in light of other studies on the Dutch labor market. Employed workers accept any job that pays them above their present wage.

The measure of individual workers is M and the measure of unemployed is equal to U. The measure of active firms is equal to N. Let $G(w)$ be the fraction of individuals with a wage lower than or equal to w. In a steady-state equilibrium, the flow of layoffs in an interval $(t, t + dt]$, $\delta(M - U)G(w)dt$ plus the flow of individuals who find a better job $\lambda_1 \overline{F}(w)(M - U)G(w)dt$, is equal to the flow of unemployed individuals accepting a wage smaller than or equal to w, $\lambda_0 F(w)dt$, if $w \geq w_{min}$. Thus

$$G(w) = \frac{F(w)}{1 + \kappa_1 \overline{F}(w)} \qquad (1)$$

3

where $\kappa_1 = \lambda_1/\delta$ denotes the expected number of job offers during a spell of employment. Likewise, we use the notation $\kappa_0 = \lambda_0/\delta$. We assume that workers draw job offers by randomly picking firms using a uniform sampling scheme. Consider a firm that offers only one wage w. The rate at which workers flow into this firm is equal to $(\delta + \lambda_1 G(w))\frac{M-U}{N}$. The rate in which workers flow out of the firm is equal to $\delta + \lambda_1 \overline{F}(w)$. Mortensen (1999) shows that the steady-state distribution of workers employed in a firm is Poisson where the single parameter is equal to the inflow divided by the outflow. By using equation (1), we find

$$l(w) = \frac{A}{(1 + \kappa_1 \overline{F}(w))^2}$$

where $A = \frac{1+\kappa_1}{1+\kappa_0}\kappa_0 \frac{M}{N}$. This parameter can be interpreted as the steady-state number of workers employed within a firm, conditional on the wage being offered. We denote the number of individuals employed within a firm by L. The distribution of L conditional the wage offer is given by

$$\mathbb{P}(L \leq l | W = w) = \sum_{k=0}^{l} \frac{l(w)^k}{k!} e^{-l(w)}$$

Firms may differ in their labor productivity p, which is not dependent on the number of workers. The distribution of p of the participating firms is $\Gamma(p)$, with $\underline{p} \geq 0$ as the infimum point of its support and $\overline{p}$ as the supremum. We assume that p has a finite mean, i.e. $\mathbb{E}_\Gamma(p) < \infty$. The assumption that the productivity level is a firm characteristic is maintained throughout the paper. Workers are homogenous in their productivity levels as well as their opportunity costs of employment.

We assume that firms seek to maximize their long-run average profits, defined as

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

It follows that firms participating in the labor market must have marginal productivity levels that are at least as high as the mandatory minimum wage. This means that the productivity distribution is actually a truncated distribution of the distribution of productivity levels of firms that can potentially enter the market. Note that a firm that sets a wage w smaller than its productivity p is

4

always willing to expand, and so will accept all job applicants and will always have a vacancy.

We make the assumption that $\Gamma(p)$ is continuous. BRVdB show that for this continuous case, the equilibrium strategies of firms are pure strategies: only one wage can be profit maximizing given the firm type. We use the notation $w = K(p)$ for this profit maximizing wage, where K is an increasing and continuous function on $[\underline{p}, \overline{p}]$. It thus follows that the distribution of wage offers can be derived from this mapping by the relationship $F(w) = \Gamma(K^{-1}(p))$

The optimal $w = K(p)$, given p and F, follows from the first order conditions: by taking $\partial \pi(p, w)/\partial w = 0$, we obtain

$$-(1 + \kappa_1 \overline{F}(w)) + 2\kappa_1 f(w)(p - w) = 0 \tag{2}$$

under the restriction that $w \geq w_{min}$, where $f(w)$ is the density associated with $F(w)$. This is an important equation since it derives an implicit function of a firm's wage offer given the productivity and the distribution of wage offers. We use this assumption tot test the restriction that high productivity firms offer higher wages. Firms with the lowest possible p (i.e. with $\underline{p}$) offer a wage equal to w_{min}. BRVdB show that the equilibrium profits for firms with productivity levels p are equal to

$$\Pi(p) = \pi(p, K(p)) = \frac{M - U}{N}(1 + \kappa_1) \left[\frac{\underline{p} - w_{min}}{(1 + \kappa_1)^2} + \int_{w_{min}}^p \frac{dx}{(1 + \kappa_1 \overline{\Gamma}(x))^2} \right] \tag{3}$$

From $\Pi(p) = (p - K(p))l(K(p))$, it follows that $K(p) = p - \Pi(p)/l(K(p))$. Thus, we find that the wage offer $w \equiv K(p)$ of a firm with productivity p equals

$$K(p) = p - \left[1 + \kappa_1 \overline{\Gamma}(p)\right]^2 \int_{w_{min}}^p \frac{dx}{(1 + \kappa_1 \overline{\Gamma}(x))^2} \tag{4}$$

This equation is the central equation of the model. In our empirical implementation, this equation is used to non-parametrically estimate the wage offers from the distribution of productivity levels.

Clearly, job durations are exponentially distributed where the single parameter equals $\delta + \lambda_1 \overline{F}(w)$. This is explained intuitively as follows: the hazard from leaving the job can be decomposed into two components. One component is the hazard of being fired, which is equal to δ and the other is the hazard of a quit. This contribution to the total hazard is equal to the rate in which job offers

5

arrive, multiplied by the probability that a job offer is acceptable (see for further details Burdett and Mortensen, 1998, and Van den Berg and Ridder, 1998). Hence, individual workers with higher wage levels stay longer in their job than those with lower wages. This result is used in the empirical implementation of the paper.

3 The data set

We use two different surveys of Statistics Netherlands for our research. These are the production survey (PS) and the wage and employment survey (WES). The data sets are based on two consecutive years of observation: 1993 and 1994. Firms are legally obliged to respond and give sufficient answers to the different questions of these surveys when sampled.

In the production survey, all firms with 20 or more employees are surveyed and from the firms with less than 20 employees a sample is drawn. We use the following variables of this survey for our analysis

1. *Registration number of the firm.* This is a unique identifier, which makes it possible to match this data set with the wage and employment survey

2. *Sector classification of the firm based on Standard Industry Classification (SIC).* Our sample contains firms in the manufacturing industry, who have 2-digit SIC numbers, ranging between 15 and 37. As stated in the introduction, we focus on three industries: the textiles industry (SIC number 17), the publishing industry (SIC number 22) and the electrical machinery industry (SIC number 31).

3. *Total employment in the firm.* This is the number of individuals who were working in the firm at the end of September of the year of observation.

4. *Total sales of the firm.* This is the monetary value of all sales during the year. Correction are made for fluctuations in the stock of unsold goods.

5. *Total value added of the firm.* This is the total sales of the firm minus the monetary value of all purchases. Corrections are made for fluctuations in the stock of primary goods.

6. *Total wage costs of the firm.* This is the total of the wage bill of the firm. It includes taxes and social security payments for both employers and employees.

6

7. *Depreciation costs of the firm.* These are the depreciation costs as they appear on the firms' balance sheets. They are roughly equal to a fixed percentage of the historical price of the long lived assets.

We assume that the total wage costs of firms divided by total employment levels are equal the wage offers of vacancies at firms. Hence, the distribution of F can be derived from these wage costs per worker. The productivity level is defined as the total value added of the firm divided by total employment.

The wage survey is based on a two-stage sample design. Statistics Netherlands takes a sample of firms to ask questions about total employment of the firm in the first stage. The sample is not a random draw but is dependent on firm size. All firms with 20 or more employees are drawn, while a sample is taken from those firms with less than 20 employees. This final sample is purely random and does not depend on sample size. Additional questions are asked for individual employees within the firm in the second stage. As before, the sample that is taken is not completely random: all firms with 100 or more employees are drawn and there is taken a sample of the firms with less than 100 employees. Additionally, the number of employees being sampled within the firm is dependent on firm size as well. Firms with less than 20 employees have to provide data related to all employees, while firms with 20 or more employees have to provide data of only a fraction of their employees. This fraction depends again on the firm size and varies from 7 to 50 percent. The actual percentage decreasing in the number of employees.

To obtain the distribution of wages among workers (i.e. the distribution G as presented in the section 2), it is important to know the probability of an individual worker being sampled within the wage survey. This probability is equal to the probability that the firm is sampled, multiplied by the probability that the individual worker is sampled, given that the firm is sampled. The first term of the product is decreasing with firm size and the second is increasing with firm size. We take account of this sampling device in our empirical implementation. We use the following variables of the wage and employment survey

1. *Registration number of the firm.* This is the same unique identifier as is used for the production survey.

2. *The sector classification of the firm based on the standard classification of industries (SIC).* Again, we use data of the three industries of analysis.

3. *The total number of workers within the firm.* This is the number of individuals working in the firm at September, 30th.

4. *The total number of employees in the sample.* This is the number of workers from whom the firm has provided answers to the questions of the survey.

5. *Hourly wage including extra payments for overtime hours.*

6. *Occupational classification of the particular worker.*

We use data of three different industries that are classified by their 2-digit SIC code. We note that the firms in the industries are still heterogeneous in the characteristics of the workers they employ. However, based on the number of observations that we have and the difficulties related to stratification with respect to other variables of the firm (for example regional differences), we decided not to stratify the firm data further.

Table 1 summarizes some descriptive statistics of the industries that we use in our analysis: the textiles industry, the publishing and printing industry and the industry for electrical machinery. First, we summarize the productivity levels per worker. The publishing industry has the highest productivity level among the industries and the productivity levels of the other industries are quite similar. We summarize the average wage bill per worker in the second row of table 1. The same patterns as with the productivity levels are found, where the publishing industry pays the highest wages. We report the mandatory minimum wages in the third row. The textiles industry and the publishing industry have quite similar mandatory minimum wages, while the mandatory minimum wage of the electrical machinery industry is much lower. Productivity differences might be the result of differences in the use of physical capital. As such, we would expect the publishing industry to have the highest depreciation costs. Looking at the fourth line of table 1, we see that this is indeed the case. The fifth row summarizes the total sales per worker for the different industries. The highest sales per worker are found in the publishing industry, which was also found to have the highest productivity levels per worker.

The next four rows of table 1 summarize the number of firms and the differences in firm size. First, we have the total number of firms within an industry. It is found that the publishing industry has by far the highest number of firms. Second, we find that most firms in the publishing industry are quite small. Only 3.5% of the firms has more than 20 employees. Additionally, we find that although the number of firms with more than 20 employees in the publishing industry is quite small, the number of very big firms is relatively high. This is also found when we look at the average employer size, which we find to be the lowest in the textiles industry and the highest in the electrical machinery industry.

	Textiles	Publishing and printing	Electrical machinery
Business statistics			
Productivity level[a]	6808	8106	6635
Wage offer[a]	5335	5991	5135
Mandatory minimum wage[a]	2516	2510	2164
Depreciation costs per worker[a]	559	1087	519
Sales per worker[a]	16137	20820	18838
Number of firms and employer size			
Total number of firms	900	4512	728
Percentage of firms > 20 employees	24.7%	3.5%	19.2%
Percentage of firms > 100 employees	6.0%	2.1%	3.7%
Employer size when > 20 employees	85	94	130
Characteristics of workers			
Job durations[b]	123.1	126.7	126.5
Percentage of skilled workers	25.3	61.4	37.9

[a]In Dutch guilders per month per worker
[b]In months

Table 1: Descriptive statistics of the three industries of analysis

The final two rows of table 1 summarize some of the characteristics of workers. First, we find job durations to be quite similar between the different industries. The textiles industry has the lowest elapsed job durations. Additionally, we find that there are important differences between the skills decomposition of the industries. The highest percentage of skilled workers is found in the publishing industry, while the lowest percentage is found in the textiles industry. This fact could explain a part of the observed differences in the productivity levels. We do not look at this effect in the present paper.

To sum up the results from the paragraphs above, we conclude that the industries are quite different from each other. The traditional textiles industry is an industry with a lot of low skilled labor, a high mandatory minimum wage, low capital intensity and large firms. The publishing and printing industry employs a lot of high skilled workers, has a high mandatory minimum wage while the production process is quite capital intensive in terms of the depreciation costs in this industry. Additionally, there are a lot of small firms in this industry. The electrical industry employs a lot of low skilled workers, has a low mandatory minimum wage and low depreciation costs. The percentage of large firms is smaller than in the textiles industry but is larger than in the publishing and printing industry. These different characteristics of the industries motivates the choice of these industries and therefore these can be seen as being representative for a lot of other industries, i.e. industries that are traditional, industries that employ a lot of high skilled workers and uses a lot of new technologies and industries in which workers are not very well protected by trade unions.

4 Preliminary issues in the empirical implementation

The distribution of G can be estimated by using the empirical distribution of the wage data

$$\widehat{G}(w) = \sum_i s_i 1(w_i \leq w) \qquad \text{with } \sum_i s_i = 1$$

where the w_i's are the observed wages and the s_i's are the weights for the observations. These weights are obtained by using information about the exact way in which the data is sampled. More precisely, the relative weights should be equal

to 1 divided by the probability that an individual employee is sampled[2]. From the discussion of the previous paragraph, we know that the probability that an individual employee is sampled is equal to the probability that the firm is sampled multiplied by the probability that the worker is sampled from the file of the firm.

An estimate of the wage offer distribution is obtained by using the relationship between F and G in equation (1)

$$\widehat{\overline{F}}(w) = \frac{\widehat{\overline{G}}(w)}{1 + \kappa_1 \widehat{G}(w)} \tag{5}$$

Although we have a stock instead of a flow sample, it can be argued that the elapsed unemployment durations are still exponentially distributed with the same parameter (see also Ridder and Van den Berg, 2000).

We have to take account of the fact that we only have observations of the year in which the employee started to work for the particular firm. An individual who started his job in the same year as the year of the interview, and thus appears to have an elapsed job duration of 0 months, is actually employed within the firm between 0 and 10 months. Similarly, an individual with a reported elapsed job duration of 10 months is actually employed within the firm somewhere between 10 and 22 months. We denote the observed job duration by t and the actual job duration by $\widetilde{t}$. The following relationship between t and $\widetilde{t}$ holds

$$\widetilde{t} = \begin{cases} \left[\frac{t-10}{12}\right] 12 - 10 & \text{if } t > 10 \\ 0 & \text{if } t \leq 10 \end{cases}$$

Where the argument between $[\cdot]$ is truncated to integers. Integration over the possible states of a particular value of $\widetilde{t}$ leads to the following log-likelihood function

$$\log L = \sum_{\widetilde{t}_i \geq 10} \left\{ (\delta + \lambda_1 \overline{F}(w_i))\widetilde{t}_i + \log(1 - e^{-12(\delta + \lambda_1 \overline{F}(w_i))}) \right\} +$$

$$\sum_{\widetilde{t}_i < 10} \left\{ \log(1 - e^{-10(\delta + \lambda_1 \overline{F}(w_i))}) \right\}$$

---

[2]We have to divide by the total sum of these weights to guarantee that the weights sum up to one.

	λ_1	δ	κ_1
Textiles industry	0.0141	0.0054	2.630
	(0.0017)	(0.0002)	(0.372)
Publishing industry	0.0157	0.0050	3.160
	(0.0013)	(0.0001)	(0.258)
Electrical machinery	0.0131	0.0045	2.908
	(0.0023)	(0.0001)	(0.592)

Table 2: Estimates of the labor market friction parameters for the different industries

Estimates of λ_1 and δ are obtained by maximization of this likelihood, conditional on equation (5). One might argue that the log-likelihood does not take account of the earnings distribution and the exact way in which workers and firms are sampled. However, it is possible to show that taking this into account only adds terms that are dependent on G, but not on λ_1 and δ. Hence, adding these terms in the log-likelihood does not alter the results.

We summarize the results of the estimates of λ_1, δ and κ_1 in table 2. The labor market frictions between the three different industries are rather comparable. They seem to be of less importance (i.e. higher κ_1) in the publishing industry, while the highest frictions are found in the textiles industry. We obtained standard errors by using bootstraps. These are found to be rather small.

5 The results

The estimated values of κ_1 are used in our analysis to estimate the equilibrium outcomes that are predicted by the model. We compare these predictions with the actual outcomes of the information of employers. We note that it might also be valuable to compare the wages earned by workers with the wages that are paid by firms. This comparison can be obtained with the derivation of equation (1). However, there are problems to compare the earned wages with the wage costs of firms. This is resulted by the social security payments that are paid by firms which are included in the wage costs but not in the gross wages earned by workers.

The original observations in the data set for employers are most likely to be affected by measurement error. This can be a big problem for the final estimates of the productivity and wage offer densities. We assume the following stochastic processes for the observed variables to be existent

$$p_{it} = p_i + \varepsilon_{it}$$

and

$$w_{it} = w_i + \eta_{it}$$

where i identifies the firm and t the year of observation. ε_{it} and η_{it} are random variables with unknown distributions. When we have time series data of these observed variables and when we assume that the ε_{it}'s and η_{it}'s are independently and identically distributed for all t, then we can estimate p_i and w_i with their respective averages.

5.1 The relationship between the productivity and wage level

Our first exercise in comparing the predictions with the observed data of firms is to derive the wage offers from the productivity levels and vice versa. Using the estimates of the labor market friction parameters, we can estimate the wage levels, given the productivity levels

$$\widehat{w}_i = K(p_i|\widehat{\kappa}_1) = p_i - \left[1 + \kappa_1 \widehat{\overline{\Gamma}}(p_i)\right]^2 \int_{\underline{w}}^{p_i} \frac{dx}{(1 + \kappa_1 \widehat{\overline{\Gamma}}(x))^2} \tag{6}$$

which is the counterpart of equation (4). We estimate the distribution of Γ by using a kernel density estimate. Figure 1 plots the estimated wage levels for the different productivity levels. We also plotted Nadayara-Watson kernel regression estimates of the productivity levels for comparison. The smoothing parameter is obtained by using least-squares cross validation (Silverman, 1986, and Blundell and Duncan, 1998). We find that the estimated relationship by using equation (6) is quite close to the relationship that is estimated by using kernel regression. For the textiles industry, it seems that the relationship between the wage level and the productivity level has a positive second order derivative. This means that high productivity firms pay relatively high wages. This case is not ruled out by the theory, but nevertheless seems quite unlikely (see BRVdB). The relationships of the other two industries seem to have negative second order derivatives.

We can derive the estimated productivity levels from the wage offers of firms by a reformulation of equation (2)
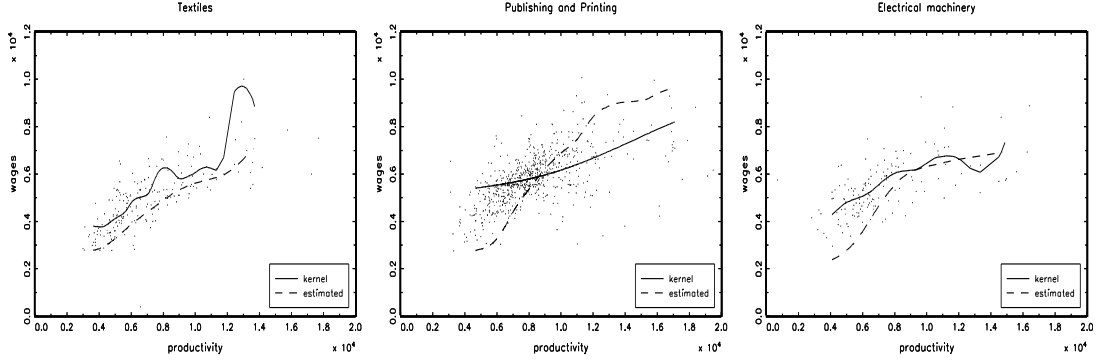
13

Figure 1: Estimated wage levels given the productivity levels by using kernel regression estimation (kernel) and by using equation (6) (estimated)

$$K^{-1}(w) = \widehat{p}_i = w_i + \frac{1 + \widehat{\kappa}_1 \widehat{\overline{F}}(w_i)}{2\widehat{\kappa}_1 \widehat{f}(w_i)} \tag{7}$$

As before, we can use the estimates of F, f and κ_1 to estimate this relationship. Results of these estimates are summarized in figure 2. We also plotted the results of a kernel regression, which are equal to the mirror image of the results presented in figure 1. The estimated relationship is downward sloping for very small productivity levels in the textiles and electrical machinery industry. This is not possible according to the theory, since it means that higher productivity firms offer lower wages. This means that some firms could increase their profits by decreasing their wages. Evidently this does not happen, so the theory fails to capture this behavior. There are a few explanations for this. First, workers are actually different in their marginal productivity and firms who lower their wages run the risk that they can only attract the low productive workers. The efficiency wage literature indicates that this is an important aspect. Second, we expect that there are differences in workers' benefit levels. In the discussion of the model, we explicitly assumed that the mandatory minimum wage is above the workers' reservation wage. This is not be a very plausible assumption for particular workers, while dropping this assumption affects firm behavior. In particular, firms do not lower their wages because they cannot attract all workers when their offered wages are too low. Bontemps, Robin and Van den Berg (1999) discuss the model with different unemployment benefits in detail. Although it should be possible to obtain somewhat comparable derivations as we did in the previous sections, we do not elaborate on this in our paper. Finally, the matching technology between workers and firms can also explain why firms do not reduce their wages. In particular, when workers meet firms according to the size of these firms, then the
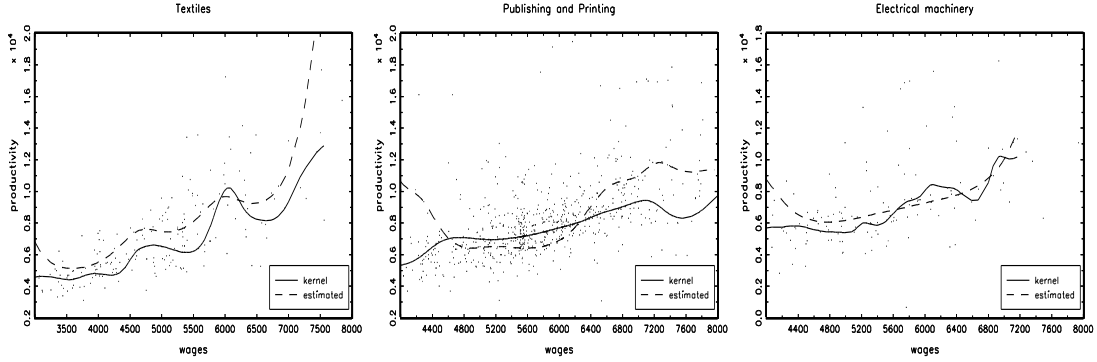
14

Figure 2: Estimated productivity levels given the wage levels by using kernel regression (kernel) and by using equation (7) (estimated)

wage offer distribution is degenerate even in the case when there are differences in the productivity levels among firms (see Burdett and Vishwanath, 1988) [3].

In section 5.3 we look at the issue of profit maximization by firms in more detail.

5.2 The predicted wage offer and productivity distribution

From the results of section 5.1, we can derive the estimated wage and productivity densities. We obtain these by calculation of the $\widehat{w}_i$'s and $\widehat{p}_i$'s and use these predicted values in kernel density estimates. These can be compared with the wage and productivity densities using kernel density estimates directly from the wage and productivity data.

The densities of the wage offers are plotted in figure 3. The estimated wage offer density of the textiles industry performs remarkably good. We find that the estimated wage offer density has the same spikes as are found in the kernel density estimates, although there are some problems with the estimation of the left tail of this distribution. The estimated wage offer distributions of the other two industries are less favorable. In both industries, the modus of the kernel density estimates are located to the right of the estimated distributions.

Figure 4 summarizes the estimated and observed density of the productivity distribution. We find some problems with the estimation of these densities for the different industries. The productivity density of the textiles industry has several intervals where there is a lot of probability mass. These intervals are also found

---

[3]Although they assume that there is only one productivity level, it is not hard to find that the distribution among productivity levels does not affect the results.
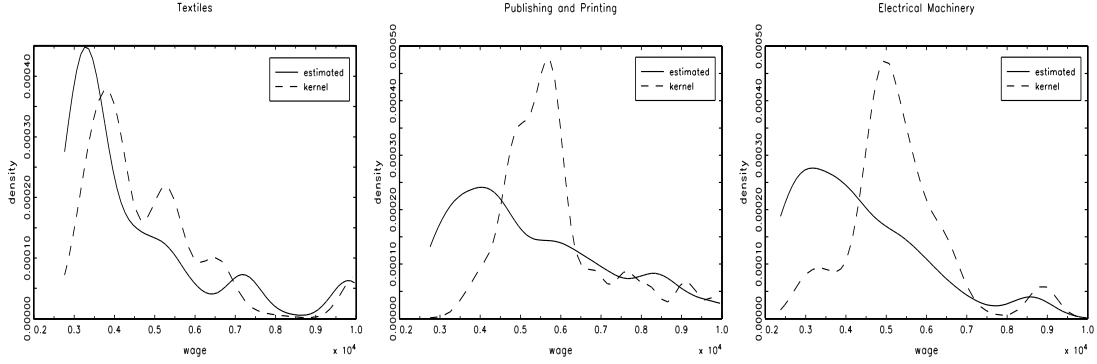
15

Figure 3: Estimates of the wage offer density by using kernel density estimates of the wage data (kernel) and by using the estimated wage offers from section 5.1 (estimated)
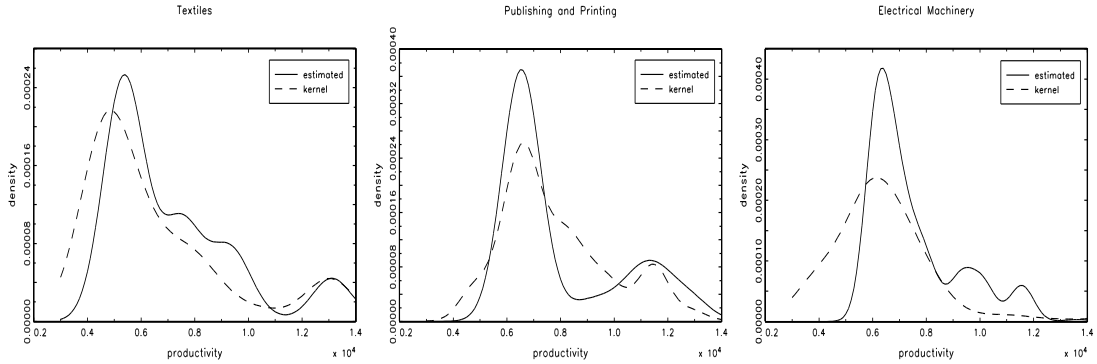

Figure 4: Estimates of the productivity density by using kernel density estimates of the productivity data (kernel) and the estimated productivity levels from section 5.1 (estimated)

by our estimated productivity levels, although the exact locations seem to be different. The problems that exist in the productivity distribution of the textiles industry do not seem to exist in the distribution of the publishing industry. We find that the modus of the estimated productivity distribution using wage offer data corresponds with that of the productivity distribution using a kernel density estimate directly from the productivity data. Additionally, the estimated right tail of the distribution performs quite well. On the other hand, the estimated probability mass just to the right of the modus is not very good. We find that the modus of the electrical machinery industry is at the good location of the density. However, the left tail of the productivity distribution of this industry does not fit the data very well.

The results from the figures above lead to very interesting insights of the performance of the equilibrium search models. From the data we know that the

16

variance of productivity levels is much higher that the variance of wage offers. This is actually an interesting result, since it is also predicted by the equilibrium search model. This means that from the data of the productivity levels, the calculation of the wage offers using equation (6) leads to wages that are much more concentrated than the original data. However, from the figures of the publishing and printing industry and the electrical machinery industry, it seems that the wages are still to far apart to predict the wage offer density very good. Additionally, it is predicted that the wage offer density is situated to the left of the density estimation by using kernel estimates directly from the data of wage offers. This means that firms pay higher wages than what the model expects given their productivity levels. On the other hand, we find the complete opposite picture from the productivity distribution. Especially, from the figure of the electrical machinery industry it is clear that the model cannot explain the low productivity levels that are observed from the data. Additionally, the observed variance of the productivity levels is much higher than that of the predicted productivity levels.

5.3 The restrictions on the wage offer distribution

Equation 7 can also be used to investigate whether the distribution of wage offers is in the set of possible distributions that can be derived from the model. From BRVdB, we have that the first order derivative of $K^{-1}(w)$ is equal to

$$
(K^{-1})'(w) = \frac{\kappa_1 f(w)^2 - f'(w)(1 + \kappa_1 \overline{F}(w))}{2\kappa_1 f(w)^2}
$$

$$
= \frac{3\kappa_1 g(w)^2 - g'(w)(1 + \kappa_1 G(w))}{2\kappa_1 g(w)^2} > 0
$$

(8)

Since high productivity firms pay higher wages, the right hand side has to be positive and hence the numerators in equation 8 have to be positive. This means that we have a restriction on both the wage offer data and the wage data sampled from a cross section of workers. Figure 5 illustrates the confidence intervals of the numerators for the wage offers from firms. Figure 6 illustrates confidence intervals for the wage data from employed workers. We use bootstrap samples to obtain these confidence intervals.

From figure 5, it is found that the restriction of a positive numerator is never rejected, although there are some values of the wage offers for which the point estimates of the numerator are quite small. Actually, the wide confidence interval result in the failure to reject the null hypothesis of a positive numerator.
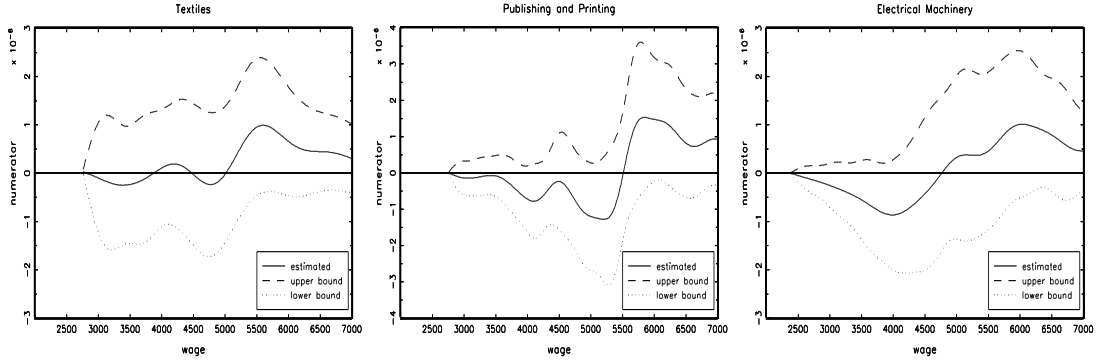
17

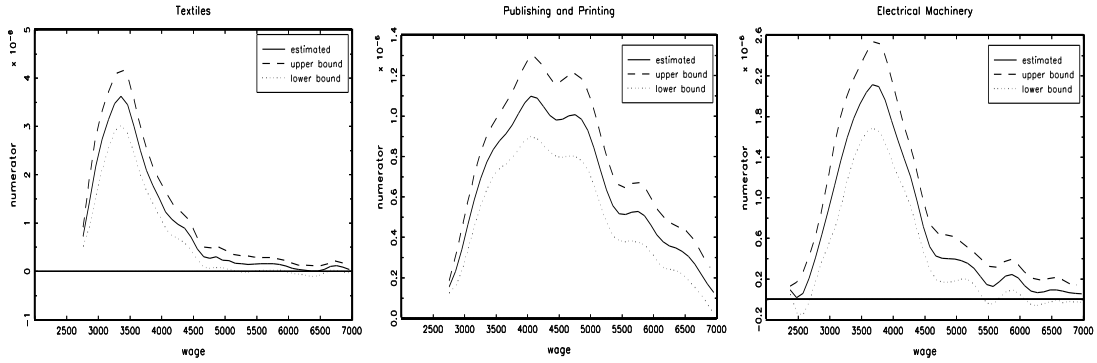Figure 5: Illustration of 95% confidence intervals for the numerator of the first line of equation (8)

Figure 6: Illustration of 95% confidence intervals for the numerator of the second line of equation (8)

The fact that the point estimates are negative for the numerator is another indication for the fact that firms can increase their profits by lowering their wages. As stated before, this might indicate that there are differences in productivity levels between individual workers.

Much narrower confidence intervals are found for the wages earned by employed workers in figure 6. Again, we are not able to reject the null hypothesis that the numerators are positive. Additionally, we find that the point estimates of the numerator are positive over the whole range of earned wages in our set of industries.

5.4   The distribution of firm size

The distribution of the number of individuals employed at a firm is obtained by integration over either the wages or the productivity levels

18

$$\mathbb{P}(L = l) = \int_{\underline{w}}^{\overline{w}} \mathbb{P}(L = l | W = w) f(w) dw$$

$$(9)$$

$$= \frac{\sqrt{A}}{2\kappa_1 l!} \int_{\frac{A}{(1+\kappa_1)^2}}^{A} y^{l-\frac{3}{2}} e^{-y} dy \qquad \forall l \in \mathbb{N}$$

The integral can be computed quite efficiently by noting that it is an incomplete gamma function. Note that this distribution does not depend on F and G and hence not on the strategy of firms. The parameter A can be estimated easily by

$$\widehat{A} = \bar{l}(1 + \widehat{\kappa}_1) \tag{10}$$

where $\bar{l}$ is a consistent estimator of the mean firm size in the industry (i.e. $\mathbb{E}(L)$). This can be seen by using

$$M - U = \frac{A}{1 + \kappa_1} N$$

The left hand side of this equation is equal to the total number of workers employed within the industry. Using the law of large numbers we find that this is equal to $N\mathbb{E}(L)$ for large N. The same result is obtained when we use

$$\mathbb{E}(L) = \int_{\underline{w}}^{\overline{w}} \mathbb{E}(L | w) f(w) dw$$

$$= A \int_{\underline{w}}^{\overline{w}} \frac{1}{(1 + \kappa_1 \overline{F}(w))^2} f(w) dw$$

$$= \frac{A}{1 + \kappa_1}$$

Table 3 gives estimates of the variable A for the three different industries. Note that based on these estimates, it is possible calculate the number of individuals that are working in the segment. These are 29281, 125701 and 27701 workers for the textiles, publishing and electrical machinery industry. Based on this,

19

	A
Textiles industry	118.1
Publishing industry	115.9
Electrical machinery	148.7

Table 3: Estimates of the variable $A = \frac{M-U}{N}(1 + \kappa_1)$
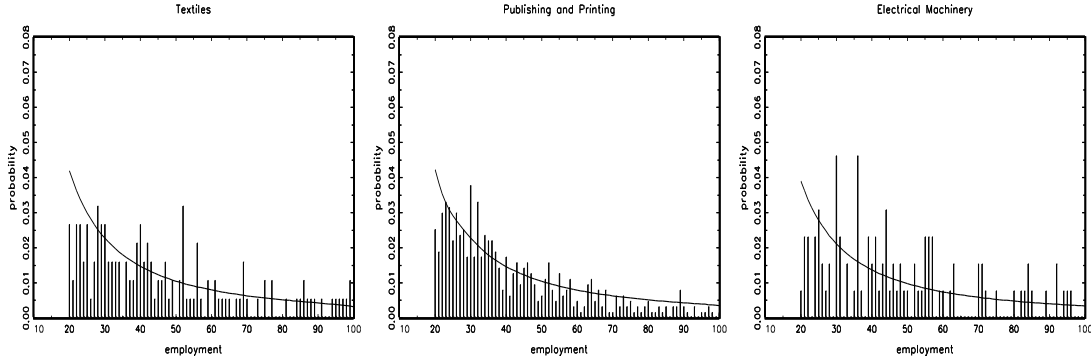

Figure 7: Estimates of the employment density (where the bars represent the number of observations)

the publishing industry seems by far the most important industry from a macro perspective.

We can make comparisons between the distribution in equation (9) and the actually observed distribution of employment levels within the industry. This is illustrated in figure 7. The global shape of the distribution is quite well estimated, although it seems not possible to explain the spikes in the distribution. These can be caused by measurement errors, since there are a lot of individual firms that report employment values dividable by 5.

We can also compare estimates of $\mathbb{E}(l|w)$ and $\mathbb{E}(l|p)$ with kernel regression estimates of l on the levels of the wage offers and productivity levels. Results of the estimated relationship between the employment level and the wage offers are illustrated in figure 8. The kernel regression estimates are quite messy, although we already used quite a large bandwidth. Nevertheless, the estimated relationship and the relationship estimated by kernel regression look quite the same.

We illustrated the relationship between the firms' productivity levels and the employment level in figure 9. These results are similar to the results found in figure 8.
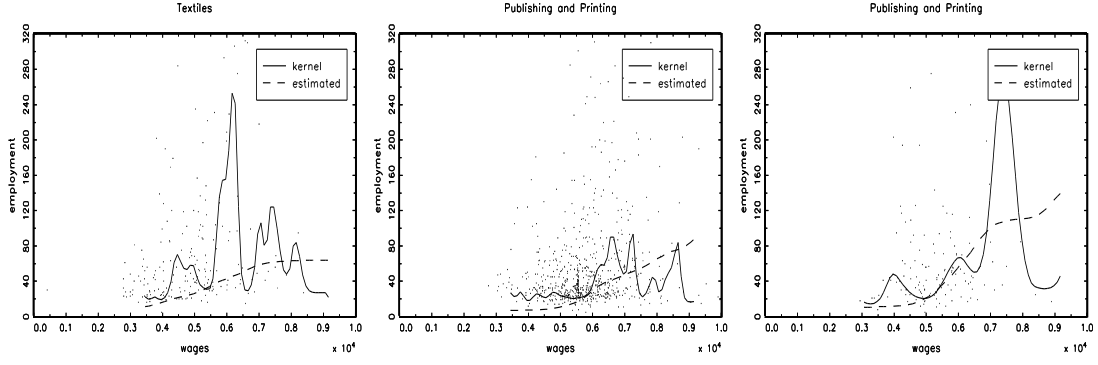
20

Figure 8: Estimates of the relationship between the wage level and the employment level by using kernel regression estimates (kernel) and the estimation of $\mathbb{E}(l \mid w)$ (estimated)
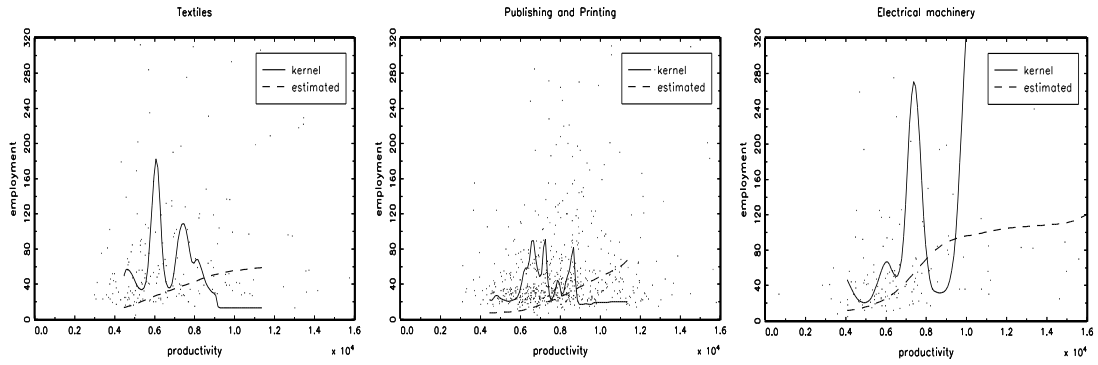

Figure 9: Estimates of the relationship between the productivity level and the employment level by using kernel regression estimates (kernel) and the estimation of $\mathbb{E}(l \mid p)$ (estimated)

5.5   Using χ^2 goodness of fit tests

We use χ^2-tests to test the goodness of fit of the productivity and wage offer distributions. For convenience, we only use the data from the firms that have 20 employees or more. The wage offer distribution, conditional on this observations is equal to

$$F(w|L \geq 20) = \frac{\mathbb{P}(L \geq 20|W \geq 20)F(w)}{\mathbb{P}(L \geq 20)}$$

$$= \frac{1 - \sum_{k=0}^{19} \frac{\sqrt{A}}{2^{\kappa_1}k!} \int_{\frac{A}{(1+\kappa_1)^2}}^{\frac{A}{(1+\kappa_1\overline{F}(w))^2}} y^{k-\frac{3}{2}}e^{-y}dy}{1 - \sum_{k=0}^{19} \frac{\sqrt{A}}{2^{\kappa_1}k!} \int_{\frac{A}{(1+\kappa_1)^2}}^{A} y^{k-\frac{3}{2}}e^{-y}dy}$$

The first step of this equation is obtained by using Bayes' rule and the second step can be found by substitution of equation (9). Likewise, the conditional productivity distribution is equal to

$$\Gamma(p|L \geq 20) = \frac{1 - \sum_{k=0}^{19} \frac{\sqrt{A}}{2^{\kappa_1}k!} \int_{\frac{A}{(1+\kappa_1)^2}}^{\frac{A}{(1+\kappa_1\overline{\Gamma}(p))^2}} y^{k-\frac{3}{2}}e^{-y}dy}{1 - \sum_{k=0}^{19} \frac{\sqrt{A}}{2^{\kappa_1}k!} \int_{\frac{A}{(1+\kappa_1)^2}}^{A} y^{k-\frac{3}{2}}e^{-y}dy}$$

The differences between the conditional and the unconditional productivity density are illustrated in figure 10. This illustration is based on the assumption that the productivity distribution is log normal. It is possible to see that the conditioning on big firms has remarkable effects on the left tail of the distribution.

The calculated probabilities are compared to the observed probabilities from the multinomial distribution

$$\widehat{P}_i = \sum_{i=1}^{n} \mathbf{1}(w_i \in A_i)$$

For $i = 1, \ldots k$ and where A_i is the set of classes with $\bigcup_{i=1}^{k} A_i \supseteq \mathrm{supp}(F), \mathrm{supp}(\Gamma)$ The estimated probabilities are equal to

$$P_i(\widehat{\kappa_1}) = \mathbb{P}_{\widehat{\kappa}_1}(W \in A_i|L \geq 20) = F_{\widehat{\kappa}_1}(\sup(A_i)|L \geq 20) - F_{\widehat{\kappa}_1}(\inf(A_i)|L \geq 20)$$
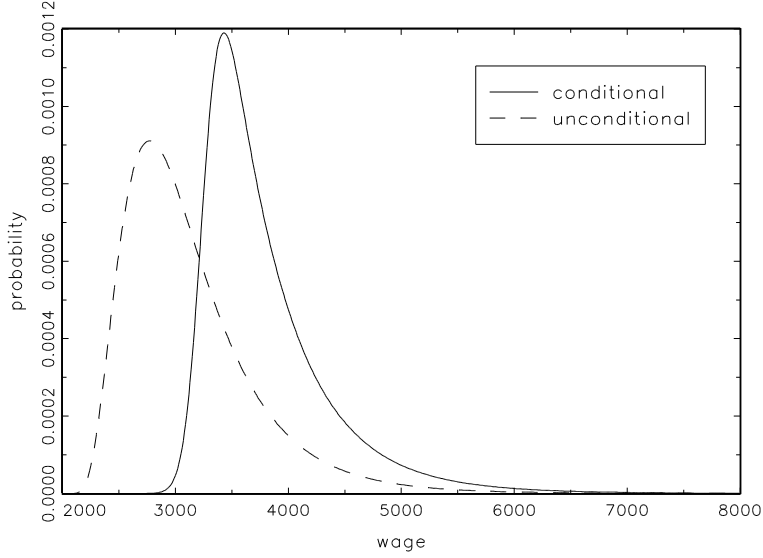
Figure 10: Illustration of the differences between the conditional and unconditional density when Γ is a log normal distribution with $\mu = 6.9$, $\sigma = 0.5$, $p_{min} = 2000$. The other parameters in the model are equal to $\lambda_0 = 0.1$, $\lambda_1 = 0.03$, $\delta = 0.01$, $M = 40000$, $N = 2000$ and $w_{min} = 1600$

for the wage offer distribution. These probabilities are similarly defined for the productivity distribution. Let $\widehat{P}$ and $P(\widehat{\kappa}_1)$ be the vectors containing the elements $\widehat{P}_i$ and $P_i(\widehat{\kappa}_1)$; $i = 1, \ldots k$. To correct for the fact that we estimate κ_1, we use the *Rao-Robson-Nikolin* instead of the conventional *Pearson* test statistic (Van der Vaart, 1998)

$$T = n \left(\frac{\widehat{P} - P(\widehat{\kappa}_1)}{\sqrt{P(\widehat{\kappa}_1)}} \right)^T \left(I_k - \frac{n}{m} \sigma^2(\widehat{\kappa}_1) C(\widehat{\kappa}_1)^T C(\widehat{\kappa}_1) \right)^{-1} \left(\frac{\widehat{P} - P(\widehat{\kappa}_1)}{\sqrt{P(\widehat{\kappa}_1)}} \right) \quad (11)$$

where

$$C(\widehat{\kappa}_1)_i = \frac{\frac{1}{n} \sum_{j=1}^n I_{A_i}(w_j) \frac{\partial \log f(w_j, \widehat{\kappa}_1)}{\partial \kappa_1}}{\sqrt{P(A_i; \widehat{\kappa}_1)}}$$

Here, m is the number of observations on which our maximum likelihood estimates are based and n is the number of observations that we use to compare the estimated distribution with. This statistic can be shown to follow a χ^2-distribution with $k - 1$ degrees of freedom under the null hypothesis that the data are drawn from the distribution being estimated. Actually, this statement

23

is based on the condition that $\widehat{\kappa}_1$ is an asymptotically efficient estimator of κ_1. Since we use a maximum likelihood procedure containing an empirical distribution function for G, asymptotic efficiency can not be guaranteed using the standard procedure. Therefore, the null hypothesis might be incorrectly rejected in some cases (*i.e.* the actual distribution of T is stochastically dominated by the χ^2 distribution with $k - 1$ degrees of freedom). However, the real value of the χ^2 goodness of fit test is never smaller than the *Pearson* statistic. Since the *Rao-Robson Nikolin* statistic is very small in our empirical analysis, we do not elaborate on this issue in this paper (see the discussion in Andrews, 1988 for the case where κ_1 is not efficiently estimated).

Another possibility to test the theory is by using minimum χ^2-estimates, i.e. the minimization of the statistic

$$T = \min_{\kappa_1 > 0} \left\{ \left(\frac{\widehat{P} - P(\widehat{\kappa}_1)}{\sqrt{P(\widehat{\kappa}_1)}} \right)^T \left(\frac{\widehat{P} - P(\widehat{\kappa}_1)}{\sqrt{P(\widehat{\kappa}_1)}} \right) \right\}$$

This statistic can be shown to follow a χ^2-distribution with $k - 2$ degrees of freedom (Van der Vaart, 1998). The problems with the asymptotic distribution of the *Rao-Robson Nikolin* test do not occur here. Note that both statistics do not have to result in the same conclusions, since the latter statistic is based on employer data only. This means that when the first method leads to a rejection of the theory, while the second does not lead to such a rejection, we expect problems to occur in the translation between employer and employee data.

Figures 11 and 12 illustrate the conditional distributions of the wage offers and productivity levels. The kernel density estimates are based on the firms with 20 employees or more. We find some differences with the results found from the unconditional distribution.

The results of the χ^2 goodness of fit tests are found in table 4. These tests are based on 10 classes, which means that the number of degrees of freedom of these tests is equal to 9. The levels of $\sigma^2(\widehat{\kappa}_1)$ can be found in table 2 and the levels of m in equation 11 are equal to 1878, 5613 and 1841 for the textiles, the publishing and the electrical machinery industry. It is possible to see that all tests are rejected under a 5% significance level. Using a 1% significance level leads to the failure to reject the null hypothesis for the electrical machinery industry.

Table 5 summarizes the results of the minimum χ^2-estimates of κ_1 using the classification based on 10 different intervals of the wage offer and productivity distribution. We find that when we minimize the χ^2-statistic based on the distribution of wage offers, then the resulting values of κ_1 decrease for the textiles and

24

	χ^2_F	χ^2_Γ	χ^2_L
Textiles industry	52.12	320.71	86.02
Publishing industry	348.99	584.32	157.37
Electrical machinery	21.10	206.97	54.37

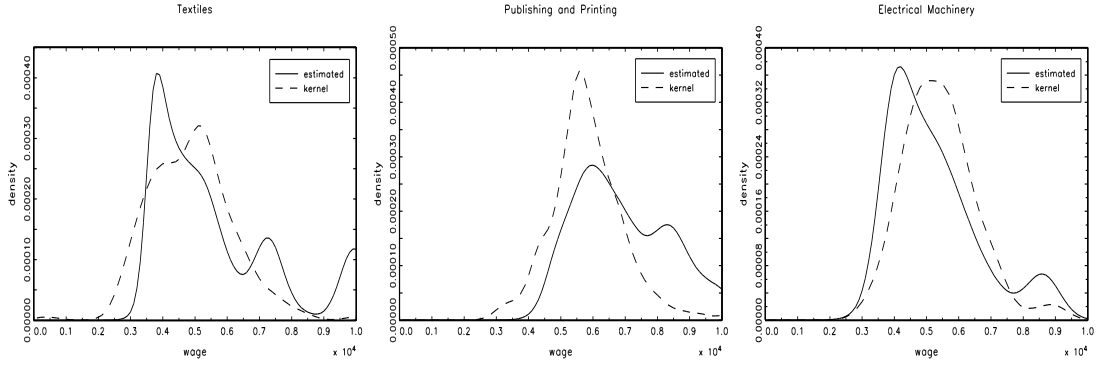Table 4: Results of the χ^2 goodness of fit tests

Figure 11: Estimates of the conditional wage offer density (estimated) together with the kernel density estimate of the firms with 20 or more employees (kernel)
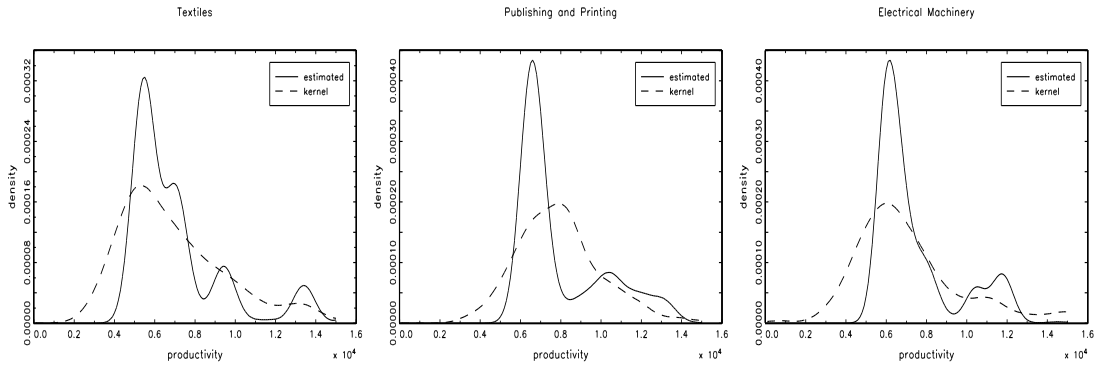

Figure 12: Estimates of the conditional productivity density (estimated) together with the kernel density estimate of the firms with 20 or more employees (kernel)

	Wage offers		Productivity levels		Employment levels	
	κ_1	χ^2	κ_1	χ^2	κ_1	χ^2
Textiles industry	2.423	45.54	1.410	51.08	2.507	68.36
Publishing industry	2.620	190.06	0.864	187.11	1.876	105.39
Electrical machinery	3.458	4.15	1.843	54.61	2.708	41.95

Table 5: Results of the minimum χ^2-estimates and the function value of the criterion function

publishing industry, while the value of this parameter increases for the electrical industry. Based on the statistic for the different productivity levels, the estimates of κ_1 decrease for the textiles and electrical machinery industry. It increases a little for the publishing industry. The κ_1's for the employment levels are similar to their original values, except for the publishing industry for which the value is roughly divided by 2.

The minimized χ^2-statistics resulting from the estimation method are presented in the second and fourth column of table 5. We see rather the same results for the χ^2-test statistics as we saw in table 4. The χ^2-statistic of the publishing industry is decreased quite a lot for the wage offer distribution, but is still far from a failure to reject the null hypothesis. Note that the degrees of freedom of this test is now decreased by 1 to a level of 8. The resulting value of the test statistic for the electrical machinery industry leads to a failure to reject the null hypothesis for any usual significance level. The values of the test statistic for the productivity distribution do not differ that much from the results found in table 4. The values of the χ^2-test statistic for the employment levels are all high enough to reject the null hypothesis.

5.6 The relationship between productivity levels and depreciation costs

Our final exercise focuses on data issues and can be seen as being suggestive or further research instead. As stated before, it is possible that there are serious problems to use the productivity levels of firms. Therefore, it might be worthwhile to use capital intensity. Although we do not have data on capital intensity we could use the depreciation costs of firms as an indication. Additionally, in a model where there are endogenous investment decisions by firms (as is the case in Robin and Roux, 1998), these decisions determine the different wage offer strategies
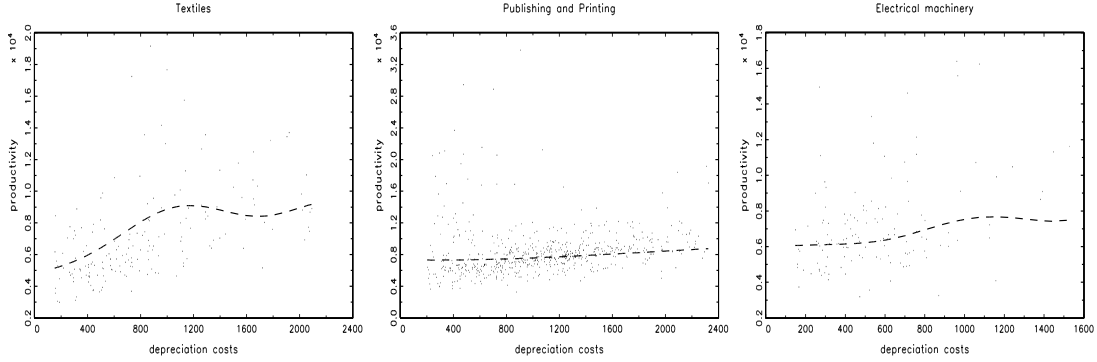
Figure 13: Illustration of the relationship between depreciation costs and pro-
ductivity levels of firms

	Textiles	Publishing and Printing	Electrical Machinery
constant	4091	7292	5213
	(392)	(315)	(544)
depreciation costs	4.440	0.795	4.424
	(0.631)	(0.363)	(1.296)
(depreciation costs)^2 (x1000)	-0.717	0.195	-0.068
	(0.195)	(0.085)	(0.561)
R^2	0.336	0.099	0.284

Table 6: Ordinary least squares estimates of the relationship between depreciation
costs and productivity levels

between firms. Figure 13 shows the relationship between productivity levels and
depreciation costs of firms by using Nadayara-Watson kernel regression estimates.
We find that the relationship is indeed upward sloping and that the relationship
is not linear. However, it seems possible to approximate the relationship by
ways of a second order polynomial. The estimation results using ordinary least
squares are summarized in table 6. We find that the depreciation costs as well
as the depreciation costs squared are highly significant (one exception is the
electrical machinery industry). The values of the R^2 differ remarkably between
the different segments, where it is quite large for the textiles industry and low
for the publishing and printing industry.

27

6 Conclusions

In this paper, we used employer data to test the predicted equilibrium outcomes of search models. For this purpose, we used the modelling framework developed in Bontemps, Robin and Van den Berg (2000) (BRVdB). We estimated the model using data on individually earned wages and elapsed unemployment durations in a similarly flexible way as they did. Based on the estimates of search frictions, we were able to obtain estimates of the productivity distribution, the wage offer distribution and the employment distribution among firms. Additionally, we analyzed the relationships between the wage offer, the productivity level and the employment levels. Formal tests were used to investigate whether the employer data could be sampled from the distributions being estimated.

In our empirical analysis, we used three different industries: the textiles industry, the publishing industry and the electrical machinery industry. The characteristics of these industries differ in some important respects. For example, productivity levels of the publishing industry and the percentage of high skilled workers in that industry are well above that of the other industries.

We found that there are problems with the estimation of the distribution of both the wage offers and productivity levels. It is especially hard to explain the low productivity levels of the larger firms. Additionally, it is found that the wages that are offered by some firms are sub-optimal, given the productivity levels of these firms. The standard deviations of the predicted productivity distributions are much smaller than those for the observations, while the opposite is true for the wage offer distribution. Formal tests for the goodness of fit of the distributions find that the null hypothesis is rejected in almost all cases, where the null hypothesis states that the observations from the data are sampled from the distribution that is predicted by the model.

The fundamental relation in the model of BRVdB is that between the wage offer and the productivity level. We tested the fit of this relationship in this paper and the problems we encountered have to be interpreted as a failure of the specific model to explain all phenomena of importance. This provides information about the locations of the possible problems and hence might give us suggestions for future research. More precisely, the problems with firm behavior that we found indicates that the differences between individual workers play an important role in the determination of the wage offers by firms. This is in line with the recently published results of Abowd, Kramarz and Maroglis (1999), who also find that the individual specific component is important in determining wage levels. Additionally, other aspects that we ignored in our analysis may play a

role as well, like the distribution of benefits among workers and the matching technology between workers and firms.

References

Abowd, J.M., F. Kramarz and D.N. Margolis, 1999, High wage workers and high wage firms, *Econometrica* **67**, 251–335.

Andrews, D.W.K., 1988, Chi-square diagnostic tests for econometric models: theory, *Econometrica* **56**, 1416–1452.

Blundell, R. and A. Duncan (1998), Kernel regression in empirical microeconomics, *Journal of Human Resources* **33**, 62–87.

Bontemps, C., J.M. Robin and G.J. van den Berg (1999), An empirical equilibrium job search model with search on the job and heterogeneous workers and firms, *International Economic Review* **40**, 1039–1075.

Bontemps, C., J.M. Robin and G.J. van den Berg (2000), Equilibrium search with continuous productivity dispersion: theory and non-parametric estimation, *International Economic Review* **41**, 305–358.

Bowlus, A.J., N.M. Kiefer and G.R. Neumann (2001), Equilibrium search models and the transition from school to work, *International Economic Review* **42**, forthcoming.

Burdett, K. and D.T. Mortensen (1998), Wage differentials, employer size and unemployment, *International Economic Review* **39**, 257–273.

Burdett, K. and T. Vishwanath (1988), Balanced matching and labor market equilibrium, *Journal of Political Economy* **96**, 1048–1065.

Eckstein, Z. and K.I. Wolpin (1990), Estimating a market equilibrium search model from panel data on individuals, *Econometrica* **58**, 783–808.

Mortensen, D.T. and C.A. Pissarides (1999), New developments in models of search in the labor market, *in:* O. Ashenfelter and D. Card, *Handbook of Labor Economics* **3**, North-Holland, Amsterdam.

Postel-Vinay, F., and J.M. Robin (2000), An equilibrium search model for matched employer-employee data, LEA-INRA, Paris.

Ridder, G., and G.J. van den Berg (2000), Estimating measures of labor market imperfection in five OECD countries, using aggregate data an equilibrium search framework, Department of Economics, Vrije Universiteit, Amsterdam.

Robin, J.M. and S. Roux (1998), Random or balanced matching: an equilibrium search model with endogenous capital and two-sided search, LEA-INRA and CREST-INSEE, Paris.

Silverman, B.W. (1986), *Density estimation*, Chapman and Hall, London.

Van den Berg, G.J. (1999), Empirical inference with equilibrium search models of the labor market, *Economic Journal* **109**, F283–F306.

Van den Berg, G.J., and G. Ridder (1998), An empirical equilibrium search model of the labor market, *Econometrica* **66**, 1183–1221.

Van der Vaart, A. (1998), *Asymptotic statistics*, Cambridge University Press, Cambridge.