Structural Estimation of Search Intensity: Do Non-Employed Workers Search Enough?*

Pieter A. Gautier†
Vrije Universiteit Amsterdam, Tinbergen Institute, CEPR

José Luis Moraga-González‡
University of Groningen

Ronald P. Wolthoff§
Vrije Universiteit Amsterdam, Tinbergen Institute

2 September 2007

Abstract

We present a structural framework for the evaluation of public policies intended to increase job search intensity. Most of the literature defines search intensity as a scalar that influences the arrival rate of job offers; here we treat it as the number of job applications that workers send out. The wage distribution and job search intensities are simultaneously determined in market equilibrium. We structurally estimate the search cost distribution, the implied matching probabilities, the productivity of a match, and the flow value of non-labor market time; the estimates are then used to derive the socially optimal distribution of job search intensities. From a social point of view, too few workers participate in the labor market while some unemployed search too much. The low participation rate reflects a standard hold-up problem and the excess number of applications result is due to rent seeking behavior. Sizable welfare gains (15% to 20%) can be realized by simultaneously opening more vacancies and increasing participation. A modest binding minimum wage or conditioning UI benefits on applying for at least one job per period, increases welfare.

Keywords: job search, search costs, labor market frictions, wage dispersion, welfare, structural estimation

JEL codes: J64, J31, J21, E24, C14

*We are grateful for useful comments received from Jaap Abbring, Gerard van den Berg, Charles Bos, Gregory Jolivet, Bas van der Klaauw, Aico van Vuuren, and Matthijs Wildenbeest. We also thank participants of the 62nd European Meeting of the Econometric Society in Budapest, the 2007 Annual Meeting of the Society for Economic Dynamics in Prague, the Workshop in Economic Theory in Rotterdam, the IZA Workshop Heterogeneity in Micro Econometric Models in Bonn, the Third Meeting of the Microdata RTN in Madrid, and seminar participants at the European University Institute in Florence and the Tinbergen Institute Amsterdam. Gautier acknowledges financial support from NWO through a VIDI grant.

†Department of Economics, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, pgautier@feweb.vu.nl.

‡Department of Economics, University of Groningen, PO Box 800, 9700 AV Groningen, j.l.moraga.gonzalez@rug.nl.

§Department of Economics, Vrije Universiteit Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam, wolthoff@tinbergen.nl.

1 Introduction

Many active labor market policies aim at increasing the search intensity of non-employed workers. Examples include (i) unemployment sanctions, like cuts in the benefits paid to the unemployed who do not engage in active job search (see Abbring, Van den Berg and Van Ours (2005)), (ii) counselling and monitoring, like advising long term unemployed workers on how to draft application letters (see Van den Berg and Van der Klaauw (2006)), (iii) financial aids, like subsidizing child care in order to increase the number of actively searching workers (see Heckman (1974) and Graham and Beller (1989)), or (iv) re-employment bonus schemes (see Meyer (1996)). The evaluation of policy programs of this kind is not easy because, on the one hand, it is difficult to measure job search intensity directly and, on the other hand, changes in search intensities affect the wage distribution and the matching rates for other workers. Hence, the general equilibrium effects can be large. In this paper we present a structural framework to estimate the primitive parameters of an equilibrium search model with endogenous search intensity and free entry of vacancies. Those primitives are the search cost distribution, the value of home production and the capital cost of vacancy creation. The estimates can then be used to calculate the socially optimal labor market allocation, as well as the desired search intensities and the desired level of labor market tightness.

We consider a discrete-time dynamic labor market with a continuum of identical, infinitely-lived workers and free entry of vacancies. At each point in time, workers are either employed at one of the firms or non-employed. Employed workers stay in their job until their match with the firm is destroyed by some exogenous shock and they become non-employed again. Non-employed workers apply for jobs at the beginning of a period. Since search intensity is the policy parameter of interest, we explicitly model it as the number of job applications workers send out per period. For each application a worker incurs a search cost. This search cost differs amongst workers and is drawn from a common non-degenerate cumulative distribution function (cdf). As in Gautier and Moraga-González (2004) (who consider a one period version of this model with identical workers), wages and the number of applications are jointly determined in a simultaneous-moves game. Firms enter the market and post wages to maximize profits. For the usual reasons, as explained in Burdett and Judd (1983) and Burdett and Mortensen (1998), firms play mixed strategies and offer wages from a continuous wage offer distribution.

Rather than assuming an exogenous specification for a matching function (see the summary of empirical studies in Petrongolo and Pissarides, 2004), this is to our knowledge the first paper that estimates a labor search model where the matching process is not only endogenously determined as a result of the firm and worker participation decisions, but also as a result of the search efforts of heterogeneous workers. Therefore, in our model, the primitive parameters are not the elasticities of an exogenously specified matching function but the quantiles of the search cost distribution. As in Albrecht, Gautier and Vroman (2006), our aggregate matching function is based on micro foundations and determined by the interplay between two coordination frictions: (i) workers do not know where other workers send their job applications and (ii) firms do not know to which workers other firms make employment offers to. These two frictions operate in different ways for different distributions of worker search intensities and have implications on wage determination. Working back from observed wages, we are able to estimate the quantiles of the search cost distribution and the implied matching rates by maximum likelihood.[1] To do this, we first derive an equilibrium relation between the accepted wage distribution and the wage offer distribution (which we do not directly observe). Then, we use the wage offer distribution together with data on labor market tightness and non-participation to estimate the distribution of search intensities and the matching probabilities. Since every worker continues to send applications till the marginal benefits of search equal the search cost, we can use this optimality condition to retrieve the magnitude of search costs for a given search intensity.

To illustrate the difference with models where either the wage distribution or search intensity is exogenous, consider the effects of increasing the minimum wage. This policy makes search more attractive on average. In our model, however, some workers may respond by increasing the number of job applications they send out from zero to one, others will increase their optimal number of job applications from one to two, and so on. Workers who already applied to many jobs could also potentially react by applying to fewer jobs because the wage distribution becomes more compressed. Each of these changes affects the aggregate matching rate in its own right. Consequently, the matching rate, the job offer arrival rate and the wage distribution are not policy invariant. Moreover, the way these endogenous variables respond to policy changes depends on the primitive search cost distribution.

We also derive the worker's reservation wage in each segment, which depends on the flow value of non-labor market time (i.e. home production and UI benefits), search costs and the wage distribution. Recently, Gautier and Teulings (2006) and Hornstein, Krusell and Violante (2006) argued that many search models cannot explain why reservation wages are substantially lower than the average or maximum wage, while at the same time unemployment or unemployment duration is low. In our model, unemployed workers who have a low search cost today realize that they can have a high search cost tomorrow. Therefore, they are willing to accept a low starting wage even though they have a large probability to receive one or more offers today.

The various policies mentioned above can be interpreted in this framework as aiming at either changing the shape of the search cost distribution or changing the marginal benefits of search. For example, one goal of subsidizing child care is to reduce the fraction of the labor force that does not search at all, while counselling unemployed workers is likely to lower the cost of writing effective application letters and increase the mean number of job applications. Besides policies that aim to directly affect search intensity, redistribution policies like UI insurance and minimum wages also affect search intensity. Without a suitable framework there is no way we can tell whether we should stimulate search intensity for all workers, only for particular groups or not at all. We find that

[1] Hong and Shum (2006) are the first to present structural methods to retrieve information on search costs in consumer markets for homogeneous goods. Hortaçsu and Syverson (2004) extend their approach to a richer setting where price variation is not only caused by search frictions but also by quality differences across products. The estimation method here is similar to that in the consumer search model of Moraga-González and Wildenbeest (2006). These models, however, are not directly applicable to the labor market since they do not capture the standard market frictions associated to rationing.

in the decentralized market equilibrium, too few workers participate in the labor market, while unemployed workers on average send too many applications. There are two reasons for the latter result. First, workers with very low search cost search too much relative to the social optimum because they do not internalize the fact that sending more applications increases the probability that multiple firms consider the same candidate. Second, search is partly a rent-seeking activity: more applications increase the expected maximum wage offer. We show that the market equilibrium output is about 15% to 20% lower than the socially optimal production. Since the planner increases participation, he also wants to increase the number of firms relative to the market equilibrium. However, given the optimal search strategies of the workers, entry is excessively large in the market because firms basically have monopsony power.

Interestingly, the introduction of a binding minimum wage can be desirable for 2 reasons: (i) it increases participation because the expected wage increases, (ii) it decreases rent seeking behavior, because it compresses the wage distribution. Besides this, it reduces vacancy supply which has ambiguous welfare effects. We model UI benefits to be conditional on searching at least once (as is the case for many OECD countries).[2] We argue that the advantage of this is that it increases the marginal benefits of sending one application rather than zero by a lot, but at the same time it does not give additional incentives to search more often, which keeps the negative congestion effects low. UI benefits therefore also increase participation without increasing rent seeking behavior. A final and important lesson from our analysis is that increasing participation or stimulating vacancy creation in isolation have very small welfare effects: the large welfare gains come from the interaction of both. It can therefore be restrictive to evaluate separate labor market policies in isolation as is commonly done.

The paper is organized as follows. Section 2 describes the theoretical model and section 3 shows how it can be estimated by maximum likelihood. Section 4 discusses our data and in section 5 we present our estimation results and discuss efficiency. Section 6 checks how robust the optimal policy is to relaxation of our simplifying assumptions. Section 7 discusses related literature and section 8 concludes.

2 Model

2.1 Setting

Consider a discrete-time labor market with a continuum of identical firms and identical, infinitely-lived workers. Both are risk neutral. We denote the measure of firms by N_F and we normalize the measure of workers to 1. We allow for free entry of firms, so N_F is endogenous. At each point in time, each worker is either employed at one of the firms or non-employed. The fractions of employed and non-employed workers at time t are denoted by e_t and $\overline{e}_t$ respectively, where $e_t + \overline{e}_t = 1$. Likewise, each job is either matched with a worker or vacant. The fraction of firms with vacancies is denoted

[2] Formally, in order to be eligible for UI benefits in the Netherlands, workers must apply four times per month where an application is defined broadly, i.e. making a phone call also often qualifies. In this paper we consider the much smaller set of serious applications that could potentially lead to a job offer.

by v_t. Employed workers stay in their job until their match with the firm gets destroyed by some exogenous shock; after this the workers in question flow into non-employment and the jobs become vacant. We assume that a fraction δ of the matches is destroyed every period.

In our model, non-employed workers can decide whether they want to search for a job or not. This provides us with a meaningful distinction between unemployment and non-participation. The non-participants are the non-employed workers who decide not to search because it is too costly, while the unemployed are workers who happen to search at least once but fail to get a job. We discuss this in more detail below. In each period a fraction m_W of the non-employed workers flows to employment and a fraction m_F of the vacancies gets filled. The fractions m_W and m_F are endogenous in our model and we will derive an expression for them in the next subsections. We make the usual assumption that the labor market is in steady state, meaning that the fraction of workers and firms in each state is constant over time, i.e. $e_t = e$ and $v_t = v\ \forall t$, where e and v are given by

$$e = \frac{m_W}{m_W + \delta} \tag{1}$$

and

$$v = \frac{\delta}{m_F + \delta}. \tag{2}$$

A worker who is employed in a given period receives a wage w. The payoff to the firm giving employment to the worker equals $y - k - w$, i.e. the difference between the value of the output produced, y, a capital cost k and the wage paid to the worker. Non-participants have a payoff that is determined by two components: the value of their home production and the value of their leisure. We assume that, together, these amount to a quantity denoted by h. An unemployed worker additionally receives unemployment (UI) benefits denoted by b. These benefits along with the option value of search determine the worker's reservation wage w_R. Firms with an unfilled vacancy do not produce, but still have to pay the capital cost. Their payoff therefore equals $-k$. All agents discount future payoffs at rate $1/(1+r)$.

We assume that a worker applies to jobs at the beginning of a period, but only learns whether she is accepted or not at the end of the period. Consequently, workers might want to send several applications simultaneously in order to reduce the risk of remaining unmatched. We allow workers to choose the number of applications, a, that they want to send. Because of computational considerations, we impose a maximum S on the number of jobs to which a worker can apply in a given period. Since S can be any finite number, this assumption is hardly restrictive.

Next, we turn to the search cost. We assume that for each application the worker incurs a search cost $c > 0$. The total cost of sending a applications therefore equals $C(a) = ca$.[3] The search cost per application, c, differs amongst workers, but is drawn from a common, non-degenerate distribution $F_c(c)$, defined over the set $[0, \infty)$. One very useful simplification we make is that workers draw a new search cost parameter every period. This captures the idea that the opportunity cost of job search

[3] In section 6.3 we consider the general class of search cost functions $C(a)$ that are weakly increasing in a and we show that the main conclusions in terms of the difference between the desired and the actual distribution of applications per worker do not change much from the linear benchmark case that we consider here.

is a random variable that is affected by things like having kids, health status, etc. The benefit of this assumption is that the reservation wage is the same for all workers. If we, alternatively did assume search cost to be worker specific, we would have to calculate search-cost-dependent reservation wages and this would make the model a lot more complicated. Since we only need the cross-sectional search cost distribution we choose the simple option.[4]

In related models of consumer search, e.g. Burdett and Judd (1983) and Moraga-González and Wildenbeest (2006), there usually is no rationing and each buyer is served. In a labor market model, the assumption of no rationing is unrealistic: firms typically hire only one or a few of the applicants for a certain job. To allow for rationing we assume an urn-ball matching function, that allows for multiple applications as in Albrecht, Tan, Gautier and Vroman (2004). The matching and wage determination process is as in Gautier and Moraga-González (2004):

1. Workers draw a search cost c, decide to how many jobs they wish to apply for and send their job application letters to random vacancies.

2. Each vacancy receiving at least one application selects a candidate randomly and offers her a job with a wage w. Applications that are not selected are returned as rejections.

3. Workers that receive one or more wage offers accept the highest one as long as it is higher than the reservation wage. Other wage offers are rejected.

The number of job applications workers send out and the wages firms set are determined in a simultaneous-moves game. In the next subsections we discuss the workers' and firms' optimal strategies. We focus on symmetric equilibria, i.e. where identical firms have similar strategies. In the estimation procedure we use a sample of the flow from non-employment to employment. This allows us to focus on the wage distribution for newly hired workers and to ignore the job-to-job transitions which are an additional source for wage dispersion (see Burdett and Mortensen (1998)). This way we can isolate the search intensity contribution to wage dispersion and keep the model tractable.

2.2 The workers' problem

The strategy of a worker with search cost c consists of a reservation wage w_R and a number of job applications $a(c)$ that she sends out to the firms. Since workers are ex ante identical, the reservation wage w_R is the same for all workers. However, workers learn their search cost c before they start applying to the vacancies, so different workers may send out different numbers of applications. Denote the fraction of non-employed workers sending a applications by p_a. For some of these workers (a fraction p_0), the search cost might be so high that it is not profitable for them to search

[4] We do not want to claim that worker specific search cost are irrelevant. For example, some workers are permanently in a position to contact many employers, just because they have a good network of contacts, or because they live in a location with many job opportunities, or because they possess the desirable social skills and working abilities.

even once in this period. These workers become non-participants. The other workers (fraction $1-p_0$) send at least one application and are therefore considered to be unemployed. Let u and n be the steady state fractions of unemployed and non-participating workers in the population, then:

$$n = p_0(1-e) \tag{3}$$

and

$$u = (1-p_0)(1-e).$$

Since search is random, all firms are equally likely to receive applications from the unemployed workers. This implies that the expected number of applications per vacancy is equal to the total number of applications divided by the number of vacancies:

$$\phi = \frac{(1-e)\sum_{a=1}^{S} ap_a}{vN_F} = \frac{\sum_{a=1}^{S} ap_a}{\theta}, \tag{4}$$

where v is the fraction of firms with vacancies and $\theta = vN_F/(1-e)$ denotes labor market tightness. Due to the infinite size of the labor market, the actual number of applicants at a specific vacancy follows a Poisson distribution with mean ϕ.[5] Likewise, the number of competitors that a worker faces at a given firm also follows a Poisson distribution with mean ϕ. In case an individual worker competes with i other applicants for the job of a firm, the probability that the individual in question will get the job equals $1/(1+i)$. Therefore, the probability ψ that an application results in a job offer equals

$$\psi = \sum_{i=0}^{\infty} \frac{1}{i+1} \frac{\exp(-\phi)\,\phi^i}{i!} = \frac{1}{\phi}\left(1-\exp(-\phi)\right). \tag{5}$$

We assume that if two or more firms compete for the same worker, the worker picks the highest wage and the other firms have to open a new vacancy in the next period as in Albrecht et al. (2004).

Given the assumptions above, the number of wage offers that a worker receives follows a binomial distribution.[6] More precisely, for a worker who sends a applications the probability $\chi(j|a)$ to get j job offers equals

$$\chi(j|a) = \begin{cases} \binom{a}{j}\psi^j (1-\psi)^{a-j} & \text{if } j \in \{0,1,...,a\} \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

We denote the fraction of non-employed workers that receive j job offers by q_j. This fraction is equal to the product of p_a (i.e. the fraction of non-employed workers sending a applications) and

[5] This is not completely obvious because in a finite labor market more matches are realized for a given mean search intensity when the variance is zero. The key intuition why the number of applicants follows a Poisson distribution in the limit and why all that matters is the average search intensity is that the probability that any two workers compete for the same job more than once is zero when workers apply to a finite number of jobs. Consequently, the event that application i results in a job offer only depends on labor market tightness and the total number of applications, and is independent of the event that application j results in a job offer.

[6] See Albrecht et al. (2006)

the probability that these a applications result in exactly j job offers, summed over all possible a:

$$q_j = \sum_{a=j}^{S} \chi\left(j|a\right) p_a. \tag{7}$$

This notation allows us to give a simple expression for the matching probability m_W that a non-employed worker flows into employment in the next period:

$$m_W = 1 - q_0 = 1 - \sum_{a=0}^{S} p_a (1-\psi)^a. \tag{8}$$

In order to derive an expression for the reservation wage we specify two discrete time Bellman equations. The first one defines $V_E\left(w\right)$, i.e. the expected discounted lifetime income of a worker who is currently employed at a wage w:

$$V_E\left(w\right) = w + \frac{1}{1+r}\left(\left(1-\delta\right) V_E\left(w\right) + \delta V_{NE}\right), \tag{9}$$

where V_{NE} denotes the expected value of being non-employed. Hence, the value of employment equals the sum of the wage w and the discounted value of employment if the worker stays in the job next period (probability $1-\delta$) or the discounted value of non-employment if the match with the firm gets destroyed (probability δ).

Non-employed workers face a trade-off when deciding how many job applications to send out. Applying to one more job is costly but it brings two sorts of benefits: one, it reduces the probability of remaining unmatched and two, it increases the likelihood to get a better paid job. Therefore, a non-employed worker with search cost c chooses the number of applications a in such a way that she maximizes her expected discounted lifetime payoff $V_{NE}(c)$:

$$V_{NE}(c) = h + \max_a \left(I_{a>0} b + \frac{1}{1+r}\left(\sum_{j=1}^{a} \chi\left(j|a\right) \int_0^\infty \max\left\{V_{NE}, V_E(w)\right\} dF_w^j\left(w\right) + \chi\left(0|a\right) V_{NE}\right) - ca\right). \tag{10}$$

This expression describes the value of non-employment for a worker with search cost c, which equals the sum of home production h and the expected discounted payoff of her optimal search strategy. If the worker sends a applications, then she receives j wage offers with probability $\chi\left(j|a\right)$. Each wage offer w is a random draw from a wage offer distribution F_w with corresponding density f_w.[7] In case the worker receives multiple job offers, she accepts the best one as long as that offer gives her a higher payoff than remaining non-employed. If the worker fails to find a job, she remains non-employed again in the next period. A necessary condition to receive UI benefits b is to actively search for a job (represented by the indicator function $I_{a>0}$), as is the case in most OECD countries. The total cost of sending a applications equals ca.

Ex ante, the non-employed workers do not know the value of the search cost that they will

[7] We derive this wage offer distribution in the next subsection.

draw. Their expected value of non-employment is therefore equal to

$$V_{NE} = \int_0^\infty V_{NE}(c) dF_c(c) . \tag{11}$$

By evaluating equation (9) in w_R and using the reservation wage property $V_E(w_R) = V_{NE}$, it follows that

$$V_{NE} = \frac{1+r}{r} w_R . \tag{12}$$

Substituting this expression back in (9) and rewriting gives

$$V_E(w) = \frac{1+r}{r+\delta}\left(w + \frac{\delta w_R}{r}\right) . \tag{13}$$

Next, we combine equations (10) to (13) to obtain an implicit expression for the worker's reservation wage.

$$w_R = h + \int_0^\infty \max_a \left(I_{a>0} b + \frac{1}{r+\delta} \sum_{j=1}^{a} \chi(j|a) \int_{w_R}^\infty (w - w_R) \, dF_w^j(w) - ca \right) dF_c(c) . \tag{14}$$

The reservation wage depends on the value of home production and the option value of search. One can easily show that this expression for the reservation wage satisfies Blackwell's (1965) sufficient conditions for a contraction mapping. Therefore, a unique value for the reservation wage w_R exists.

Let $\chi(j|a)$ be the probability to receive j offers conditional on sending a applications, given by (6). Then, the expected benefits of applying to a jobs is:

$$\zeta_a = \sum_{j=1}^{a} \chi(j|a) \int_{w_R}^\infty (w - w_R) \, dF_w^j(w) . \tag{15}$$

A worker with j offers receives the expected maximum of j draws from the wage offer distribution $F(w)$, but she has to give up the value of being non-employed, i.e. w_R. In order to derive which fraction of the workers applies to $0, 1, ..., S$ jobs, we first determine the expected gains from searching one additional time. We denote this increment by Γ_a. From the expressions above, it follows that Γ_a is equal to

$$\Gamma_1 = \frac{1}{r+\delta}\zeta_1 + b \tag{16a}$$

$$\Gamma_a = \frac{1}{r+\delta}\left(\zeta_a - \zeta_{a-1}\right), \ a = 2, ..., S, \tag{16b}$$

where the first terms on the right-hand sides of (16a) and (16b) reflect the expected wage increase of an additional job application. The second term of equation (16a) reflects that the worker becomes eligible to unemployment benefits b, when she searches once instead of being a non-participant. It is straightforward to show that Γ_a is a decreasing function of a. This implies that workers continue searching as long as Γ_a is larger than their search cost c. Hence, the fractions p_a satisfy the following conditions:

$$p_0 = 1 - F_c(\Gamma_1) \tag{17a}$$

$$p_a = F_c(\Gamma_a) - F_c(\Gamma_{a+1}), \; a = 1, 2, ..., S-1 \tag{17b}$$

$$p_S = F_c(\Gamma_S) \tag{17c}$$

2.3 The firms' problem

In this subsection we derive the wage offer distribution for newly hired workers. A firm with a vacancy offers one randomly picked applicant (if present) a wage w. In order to be attractive to both the firm and the applicant, this wage should be higher than the worker's reservation wage w_R, but lower than the value of the output that will be produced in case of a match net of capital cost, $\widehat{y} = y - k$. Moreover, the wage has to be higher than the legal minimum wage $w_{\min}$. Define $\underline{w} = \max\{w_R, w_{\min}\}$. The firm faces a trade-off within the interval $[\underline{w}, y-k]$: posting a lower wage increases its payoff $y - k - w$ conditional on the worker accepting the offer, but it also increases the probability that the worker rejects the offer and chooses to work for another firm.

For reasons similar to those in Burdett and Judd (1983) and Burdett and Mortensen (1998), there exists no symmetric pure strategy wage equilibrium.[8] However, there exists a mixed strategy equilibrium in wage offers to newly hired workers. Let F_w denote the equilibrium wage distribution. A firm that offers the lower bound of this distribution only attracts workers without other offers. Firms that offer wages below $\underline{w}$ will never hire workers. Therefore, the lower bound of the wage distribution must be equal to $\underline{w}$. The expected payoff for a firm offering $\underline{w}$, denoted by $\pi(\underline{w})$, equals the product of $y - k - \underline{w}$ and the probability that the firm offers the job to a worker who does not receive job offers from other firms:

$$\pi(\underline{w}) = (y - k - \underline{w}) \frac{1}{\theta} q_1. \tag{18}$$

In general, the expected payoff $\pi(w)$ to a firm offering a wage w to an applicant is equal to the product of $y - k - w$ and the probability that all other offers the applicant receives, if any, quote a lower wage[9]:

$$\pi(w) = (y - k - w) \frac{1}{\theta} \sum_{j=1}^{S} j q_j F_w^{j-1}(w). \tag{19}$$

In equilibrium, each wage in the support of F_w must yield the same level of expected profits to a firm. Therefore, equating (18) and (19) gives an equal profit condition that implicitly defines the

[8] To be precise, in Burdett and Judd's (1983) nonsequential search model there is a pure-strategy equilibrium where workers are offered the minimum wage. This type of equilibrium is non-generic in the sense that it can only exist for particular search cost distributions.

[9] See appendix A.1 for a derivation.

equilibrium wage distribution F_w:

$$\sum_{j=1}^{S} jq_j F_w^{j-1}(w) = q_1 \frac{y-k-\underline{w}}{y-k-w}. \tag{20}$$

Evaluating (20) at the upper bound $\overline{w}$, where $F_w(\overline{w}) = 1$, gives :

$$\overline{w} = y - k - \frac{(y-k-\underline{w})\, q_1}{\sum_{j=1}^{S} jq_j}, \tag{21}$$

which is strictly smaller than y, since $q_1 > 0$. Hence, firms never post wages up to the productivity level. After all, that would give the firm a payoff of zero with probability one, while posting a lower wage gives a strictly positive expected payoff, since some applicants do not compare wages. In Appendix A.2 we give an expression for the density function of posted wages f_w.

As noted above, the matching probability of a firm depends on its wage offer to the applicant. Given a wage offer w, the matching rate is equal to

$$m_F(w) = \frac{1}{\theta} \sum_{j=1}^{S} jq_j F_w^{j-1}(w). \tag{22}$$

Therefore, the probability to hire a worker at wage w equals the probability of offering the job to a worker who happens to get no job offers or only offers paying less than w. An expression for the *ex ante* matching probability m_F (before drawing a wage from the wage offer distribution) can be derived by integrating equation (22) over the support of $F_w(w)$. This yields

$$m_F = \frac{m_w}{\theta},$$

where $m_W = 1 - q_0$ which is given by (8). These equations show that the matching rates are completely endogenous and impose more structure on the model than most other search models that consider an exogenously specified matching function or job offer arrival rate.

Finally, we can derive the firm's value functions, which determine the extent of entry in the market. A firm that is matched to a worker produces output y and has to pay a capital cost k and a wage w. In the next period, the firm is still active with probability $1-\delta$; otherwise it has a vacancy again. Hence, the firm's value $V_F(w)$ of being matched with a worker earning a wage w is given by

$$V_F(w) = y - w - k + \frac{1}{1+r}\left((1-\delta) V_F(w) + \delta V_V\right). \tag{23}$$

where V_V denotes the lifetime payoff of a vacancy. A firm that has a vacancy incurs the capital cost k, but does not produce. If the firm offers a wage w, it matches with probability $m_F(w)$, resulting in a value $V_F(w)$ in the next period. If the firm does not match (probability $1 - m_F(w)$) it gets V_V again. Hence, V_V equals

$$V_V = -k + \frac{1}{1+r}\left(\int_{\underline{w}}^{\overline{w}} \left(m_F(w) V_F(w) + (1 - m_F(w)) V_V\right) dF_w(w)\right). \tag{24}$$

We assume free entry of vacancies, i.e. unmatched firms enter the market as long as the expected payoff is positive. Hence, in equilibrium it must be the case that $V_V = 0$. Substituting this condition into equation (23) and solving for $V_F(w)$ gives

$$V_F(w) = \frac{1+r}{r+\delta}(y - w - k).$$

We now note that as a result of the constancy-of-profits condition we can write $m_F(w)V_F(w)$ as

$$m_F(w)V_F(w) = \frac{1+r}{r+\delta}\pi(w) = \frac{1+r}{r+\delta}\frac{1}{\theta}q_1(y - k - \underline{w}). \tag{25}$$

Substituting (25) in (24) and using the free entry condition $V_V = 0$ yields

$$0 = -k + \frac{1}{r+\delta}\frac{1}{\theta}q_1(y - k - \underline{w}). \tag{26}$$

This expression implicitly determines the free-entry equilibrium number of firms in the market.

2.4 Efficiency

An interesting policy question is whether the market is efficient. To answer this question we consider a social planner that can decide how many firms enter the market and how many vacancies a worker should apply for. We distinguish two different cases: (i) a social planner that is constrained in the sense that he cannot solve the coordination frictions in the market and (ii) an unconstrained planner who can match any specific worker and firm. So, the constrained planner faces the same matching function as the market, which was given in equation (8) and an unconstrained planner generates a number of matches that equals the minimum of the number of unemployed and the number of vacancies.

$$m_W^U = \frac{\min\{u, vN_F\}}{1-e}.$$

This allows us to decompose the efficiency loss in the economy into a part that is directly due to frictions and a part that is due to distorted incentives. The aim of the social planner is to maximize total discounted future output, net of application and entry costs. Suppose that in period t, $N_{F,t}$ firms are present in the market and that the workers apply according to $\{p_{0,t}, p_{1,t}, ..., p_{S,t}\}$. These parameters imply values for ϕ_t, ψ_t, $m_{W,t}$ analogous to the equations derived in section 2.2. Then, given employment e_t at time t, the matching technology implies that the the employment level in the next period is given by:

$$e_{t+1} = e_t(1-\delta) + m_{W,t}(1-e_t).$$

In each period the employed workers produce y, while the non-employed produce h and incur the search cost c for each application they send. Note that if the planner wants a group of workers to search S times, it is optimal to let this group consist of the individuals with the lowest search costs. Similarly, the workers that should not search at all are the ones with the highest search costs. Therefore, there is a one-to-one relationship between the set $\{p_{0,t}, p_{1,t}, ..., p_{S,t}\}$ and the function $a_t(c)$ which assigns the optimal number of applications to an individual with search cost c. Each of the $N_{F,t}$ firms present in the market has to pay the capital cost k. Hence, the Lagrangian is given by

$$L\left(\{p_{0,t},...,p_{S,t},N_{F,t}\}_{t=0}^{\infty}\right)=\sum_{t=0}^{\infty}\left(\frac{1}{1+r}\right)^{t}\left(e_t y+(1-e_t)\left(h-\int_0^{\infty}a_t\left(c\right)cdF_c(c)\right)-N_{F,t}k\right)$$
$$+\lambda_t\left(e_{t+1}-e_t+\delta e_t-m_{W,t}\left(1-e_t\right)\right),$$

where λ_t denotes the Lagrange multiplier for period t. If we restrict ourselves to the steady state, this problem reduces to maximizing per-period output:

$$\Pi^*=\max_{p_0,...,p_S,N_F}\left(\frac{m_W}{\delta+m_W}y+\frac{\delta}{\delta+m_W}\left(h-\int a\left(c\right)cdF_c\left(c\right)\right)-N_Fk\right), \tag{27}$$

subject to the conditions $\sum_{a=0}^{S}p_a=1$, $p_a\in[0,1]\ \forall a$, and $N_F\geq 0$.[10] Since our estimation method provides us with estimates of y, h, and $F_c(c)$, we can numerically solve this maximization problem and confront the market outcome with the social optimum. We do this in section 5.3.

3 Maximum likelihood estimation

3.1 Likelihood

The model described in the previous section can be estimated by maximum likelihood. In this section we provide a short sketch of the estimation procedure. For more details we refer to appendix A.2. We show how we can obtain estimates for $\{p_0,...,p_S\}$, which in turn provide us with the cutoff points $\{\Gamma_1,...,\Gamma_S\}$ of the search cost distribution. We start by discussing the data that are required to estimate the model. First of all, we need cross-sectional wage data for newly hired workers who enter from unemployment. Our source for this sort of information is the Dutch AVO data set, which contains information on the Dutch Labor market. We discuss this data set in more detail in section 4. Secondly, we need some aggregate statistics on the labor market. In particular, we need information about the number of vacancies vN_F in the market and about the fractions of employed (e), unemployed (u) and non-participating ($n=1-e-u$) individuals. Accurate data to estimate these variables are available from most statistical agencies and they are usually not only available for the labor market as a whole, but also for submarkets. Third, we need information about the level of the unemployment benefits to calculate household production. The unemployment benefits, b, can, without loss of generality, be defined as the product of a replacement rate ρ and the average wage. An estimate for ρ can easily be obtained from macro-data. Note that the replacement rate ρ only changes the decomposition of the reservation wage into b and h. Therefore, the estimates for $\{p_0,...,p_S\}$ do not depend on the value of ρ.

Two other parameters have to be fixed exogenously: the maximum number of applications per period S and the discount factor $1/(1+r)$. One can easily test whether the estimation results are

[10] Note that total discounted future output differs from per-period output by a factor $\frac{1+r}{r}$. Since the factor is a constant, omitting it does not affect the optimal values for $p_0,...,p_S$, and N_F.

sensitive to the values chosen for these parameters, but in general this does not seem to be the case. For example, $S = 30$ and $S = 40$ give very similar results, because the difference in expected payoff between searching 30 or 40 times is negligible. Likewise, note that choosing a different value for the interest rate only affects the scale of the search cost distribution $F_c(c)$.[11] It does not change the estimates of the search fractions p_a, the job offer probability ψ, the job offer fractions q_j, or the net productivity $\widehat{y} = y - k$.

These parameters, the data, the structure of the model and the steady state assumption provide us with all the information we need to estimate the search cost distribution. The first step is to use (3) to identify p_0 as the ratio of the fraction of non-participants in the population to the fraction of non-employed. The other fractions p_a are estimated by maximizing the likelihood of the observed wages.

Note that, as in many models with on-the-job search, cross-sectional wages are not representative for the wages that are offered by the firms, but only for the wages that are accepted by the workers. High wage offers are more likely to be accepted than low wage offers, so the distributions of the offered wages and the accepted wages differ from each other. We denote the distribution of the accepted wages by $G_w(w)$ with associated density $g_w(w)$. Conditional on receiving at least one job offer, a worker will only accept a wage that is lower than some value w if all the j offers that she receives after sending a applications are lower than w. This means that $G_w(w)$ simply follows from $F_w(w)$ (see appendix A.2 for a derivation).

Flinn and Heckman (1982) and Kiefer and Neumann (1993) suggest to use the lowest wage and the highest wage in the sample to estimate the lower end and the upper end of the support of the wage offer distribution.[12] Although this approach gives superconsistent estimates, we do not follow this suggestion, since these order statistics are quite sensitive to outliers. Instead, we estimate the net productivity $\widehat{y}$ and the lower bound $\underline{w}$ as parameters in our maximum likelihood problem. Together they imply a value for the upper bound $\overline{w}$ as was shown in equation (21). A small fraction of the observations in our data set lies outside the interval $[\underline{w}, \overline{w}]$. We consider these observations to be the result of measurement error. We incorporate this measurement error in our model in the standard way (see e.g. Wolpin (1987)): the observed wage $\widetilde{w}$ depends on the true wage w and a multiplicative random error term ε with a log-normal distribution with parameters $\mu = 0$ and $\sigma^2 = \text{var}(\log(\varepsilon))$. We will estimate the value of σ.

The density of the observed wages can now be obtained by integrating over all possible values of the error term. Let $g_{\widetilde{w}}(\widetilde{w})$ denote this density, then the likelihood of the sample is equal to the

[11] Equation (16) shows that $\Gamma_{a>1} = \dfrac{1}{r+\delta}(\zeta_a - \zeta_{a-1})$, where ζ_i does not depend on r or δ. Doubling $r+\delta$ therefore halves these cutoff points. The effect on Γ_1 is slightly smaller, since Γ_1 includes the constant b.

[12] See also Donald and Paarsch (1993) for a discussion of the use of order statistics to estimate the bounds of distributions.

product of $g_{\widetilde{w}}(\widetilde{w}_i)$ for each individual i.[13] So, the maximum likelihood problem is given by

$$\max_{p_1,...,p_S,\sigma,\underline{w},\widehat{y}} \frac{1}{N}\sum_{i=1}^{N}\log g_{\widetilde{w}}(\widetilde{w}_i), \tag{28}$$

subject to the conditions $\sum_{a=0}^{S} p_a = 1$, $p_a \in [0,1]$ $\forall a$ and $w_{\min} \leq \underline{w} \leq \widehat{y}$. The derivations in the appendix show that the productivity and the capital cost only enter the expression for $g_{\widetilde{w}}(\widetilde{w}_i)$ as the difference $y - k$. Hence, the productivity and the capital cost are not separately identified: we can only obtain an estimate for the net productivity $\widehat{y}$. Ex post however, we can retrieve the value for k from equation (26). Subsequently, the productivity y simply equals the sum of $\widehat{y}$ and k.

As is common in these kind of models, the reservation wage is only identified if it exceeds the minimum wage.[14] In that case $w_R = \underline{w}$. Otherwise, we can only derive some bounds on w_R (or alternatively one has to make parametric assumptions). The upper bound in that situation is given by $\underline{w}$, while the lower bound is defined by equation (14) and the restriction $h = 0$. We discuss these bounds in section 6.1.

The covariance matrix of the estimates is calculated by taking the inverse of the negative Hessian matrix evaluated at the optimum. The standard errors of the other variables, e.g. q_j, can be calculated by using the delta method.

3.2 Goodness of fit

In our model, the density of accepted wages $g_w(w)$ has a flexible form: it can be strictly upward sloping, but also hump-shaped. This is shown in figure 1, which displays the wage density for two different search profiles $\{p_0, p_1, ..., p_S\}$ while keeping the other parameters fixed to some arbitrarily chosen values (in particular, $\underline{w} = 0$, $y - k = 20$, $\theta = 1$). If many workers search very little, then a given firm with an applicant does not face much competition from other firms. As a result, a large part of the probability mass is at low wages. Conversely, if enough workers send many applications, then firms have an incentive to post relatively high wages as well. Hence, by choosing the right values $\{p_0, p_1, ..., p_S\}$ we get a hump-shaped wage distribution. This flexibility is an important advantage compared to existing search models with identical workers and jobs, like for example Burdett-Mortensen (1998). These models are unable to generate wage distributions similar to the hump-shaped ones observed in real-world markets.

At first it may seem that our model is so flexible that we can fit any distribution because we basically have a polynomial of degree S. However, our model places a lot of structure on this polynomial. First, fractions p_a are probabilities and must therefore be non-negative and sum to 1. Secondly, our explicit modelling of the contact process imposes restrictions on the relation between $\{p_0, p_1, ..., p_S\}$ and $\{q_0, q_1, ..., q_S\}$, see equation (7). For example, $q_j > 0$ implies that $q_i > 0$ for all $i < j$. This imposes structure on the expected payoffs of searching a times, see (15), which in turn affects the shape of $F_w(w)$. Thirdly, the shape of $F_w(w)$ is further restricted by the equal profit

[13] See the appendix for an expression for $g_{\widetilde{w}}(\widetilde{w})$.
[14] Flinn and Heckman (1982) label this the recovarability problem.

condition of firms, given in equation (20). Fourth, workers with multiple offers choose the highest wage, which imposes conditions on the relation between $F_w(w)$ and $G_w(w)$, see (30) in appendix A.2.

Because of these restrictions, it is not obvious that our model can generate a good fit. In the empirical analysis, we assess the fit of the model in three different ways. First of all, we compare the wage distribution implied by the model to a kernel estimate and check whether they are close to each other. This test alone is however not sufficient, since the maximum likelihood estimation is designed to match the wage distribution. Therefore, we also evaluate the fit by (i) comparing the matching and the separation probabilities of the firms and workers predicted by the model to the actual ones and (ii) by considering the value of σ. Judging the matching and separation probabilities is relevant since it gives an indication whether the model can fit the average durations of unemployment and employment spells. The value of σ provides a natural test on the fit of the model, because it indicates which degree of measurement error is required to get a good fit of the wage distribution. If we find a very large value for σ, then a large part of the variation in the data cannot be explained by the model, implying that the model performs relatively poor. If we however find a small value for the standard deviation, this can be seen as supporting evidence for the model.

3.3 Search cost distribution

From the maximum likelihood estimates, we can derive cutoff points of the search cost distribution. In appendix A.3 we show how we apply a change of variables to get a simplified expression of (15) for ζ_a. The marginal gains from an additional application can then be calculated from (16), where the equilibrium value for the separation rate δ follows from the steady state condition given in equation (1). The values of Γ_a serve as cutoff points of the search cost distribution $F_c(c)$, as is shown in equation (17).

Figure 2 illustrates how the search cost distribution can be estimated from the observed wage distribution for the simplifying case that $w_R = 0$. Note first that the expected maximum wage offer a worker may receive when applying for a jobs, ζ_a, corresponds to a point on $G_w(w)$ (panel 1). The shape of $G_w(w)$ determines the marginal benefits of search, Γ_a. For example, in a close-to-competitive economy where workers are the scarce factor, most job applications result in an offer so wages are close to net productivity. As a result, one should expect Γ_1 to be very large and $\Gamma_{a>1}$ to be close to zero. Figure 2 shows that the marginal benefits of applying to more than 1 job are positive but decreasing. A worker, realizing that her applications do not affect the wage distribution, takes $\Gamma_1, \Gamma_2, ..., \Gamma_S$ as given and chooses her search intensity such that the marginal gains of an additional application equal the marginal cost (panel 2). For example, if her search cost equals c_{low} in the second panel, then she should apply four times. If her search cost is c_{high} instead, she should not search at all since the marginal gain of the first application is already smaller than the marginal cost.

An econometrician proceeds in exactly the opposite way. When he observes (or estimates) that a fraction p_0 of the workers does not search at all, he concludes that the search cost c of each of these individuals must have exceeded Γ_1. This provides him with one point of the search cost distribution $F_c(c)$, i.e. $p_0 = 1 - F_c(\Gamma_1)$ (panel 3). Similarly, by taking for example $F_c(\Gamma_4) - F_c(\Gamma_5)$ one obtains the fraction of workers with search costs such that if they search 4 times or less, the marginal benefits exceed the marginal cost but if they search 5 times, the marginal cost of search exceeds the marginal benefits. So, this determines the fraction of workers who search 4 times, p_4. In the estimation procedure, we start with the wage distribution which gives information on the fraction of workers who received j offers, q_j. The structure of the model relates $\{q_0, ..., q_S\}$ to $\{p_0, ..., p_S\}$ and implies values for the marginal benefits of searching a times. The set $\{q_0, ..., q_S\}$ is chosen such that the wage offer distribution implied by the model is as similar as possible to the observed one. Together, this procedure determines S points located on the search cost distribution. Hence, if for example the difference between Γ_1 and Γ_2 is large but p_1 is nevertheless small, it suggests that $F_c(c)$ is flat in that region.

In section 5.3, where we solve the social planner's problem, we need the full distribution $F_c(c)$. An approximation of this distribution can be obtained by interpolating the S cutoff points. In this paper, we use linear interpolation. Note that we also have to extrapolate the distribution, because we do not know the distribution of the search costs among the non-participants. We only know that for this group the search cost c is larger than Γ_1, because otherwise they would have searched at least once. However, for the social planner it makes a difference whether the search cost of a specific non-participant is only slightly higher than Γ_1 or much higher. We start by using linear extrapolation: we assume that the search cost distribution keeps increasing linearly for $c > \Gamma_1$, with the same slope as just before Γ_1, until it reaches 1. In section 6.2 we relax this assumption by considering bounds on the search cost of non-participants (also shown in panel 3).

Finally, solving the social planner's problem requires values for unemployment benefits b and household production h. The value for b can be calculated from the replacement rate ρ and the average wage. To be precise,

$$b = \rho \int_{\underline{w}}^{\overline{w}} w dG_w(w).$$

This result together with (14) can be used to derive an estimate for h, which simply follows from the difference between the reservation wage and the option value of search.

4 Data and empirical issues

4.1 Parameters

We apply the model developed in the previous section to the Dutch labor market. The wage data that we use for estimation are described in detail in the next subsection. Here, we first explain how we obtain estimates for the exogenous parameters. We start by setting the maximum number of applications S equal to 30. As mentioned before, the estimation results are not sensitive to this specific value.

We use data from Statistics Netherlands to get a value for the replacement rate ρ. The Dutch government spent 4075.5 million euros on UI benefits in 2005. The stock of unemployed contained on average 305140 individuals in that year. Hence, on average 13350 euros were paid per individual. Since the average income amounted to 33000 euros, we set the average replacement rate ρ equal to 0.40. This is exactly the same value that Hornstein et al. (2006) use for the US.

In order to determine a reasonable value for the discount factor r, we first have to fix the length of a period in our model. For this, we rely on Van Ours and Ridder (1993), who study vacancy durations. They find that the time that elapses between posting and filling a vacancy conditional on having candidates is about four months. Given an annual interest rate of 5%, this implies that $r = 0.0164$. It is worth stressing that the length of a period does not affect our estimates of the search intensities p_a, the probability of getting a job offer ψ, the fractions q_j of workers getting j offers, or the net productivity $\widehat{y}$. It only affects the discount rate, which in turn only rescales the search cost distribution.

Values for the labor market statistics e, u, and n are also obtained from Statistics Netherlands. Data for these parameters are available for each combination of calendar year, gender, education, and age cohort. We use that information to calculate the values of p_0 for our sample, taking into account the composition of the sample. The number of vacancies vN_F is calculated indirectly: it equals the product of the average labor market tightness $\theta/(1-p_0)$ (0.70 in our sample) and the unemployment rate u. In section 4.4 we present the numerical values for these parameters.

4.2 AVO data set

The source for the wage data that we use in the empirical application is the AVO data set[15] of the Dutch Labor Inspectorate, which is part of the Ministry of Social Affairs and Employment. The data are collected annually from the administrative wage records of a sample of firms. The sample period spans from 1992 to 2002. The sampling procedure consists of two stages. In October of each year, first a stratified sample of firms in the private sector is drawn from the Ministry's firm register. The strata are based on industrial sector and firm size (measured by the number of employees). In the second stage, workers are sampled from the administrative records of the firms. Information is collected at two points in time: one year before the sampling date and the sampling date itself. The number of workers sampled depends on the firm size, the number of workers who are newly hired, who have stayed in their job or who have quit the firm, and the number of workers covered by collective labor agreements. The data set contains sampling weights for both the firm strata and the employees. For the firm the weight is equal to the inverse of the probability that the firm is sampled, while for the workers it corresponds to the inverse of the probability that the worker was selected from all employees at the firm. Multiplying these values gives the weight that can be used to calculate sample statistics for the workers.

A consequence of the sampling design is that we do not observe flows that occur between the two sampling dates. Our assumption that the length of one period in the model equals four months implies that we only observe the time points $t = 0, 3, 6...$ in the AVO data set. Workers may have experienced other employment and unemployment spells between these moments of observation.

[15] AVO is the Dutch acronym for *Terms of Employment Study.*

Note that this is not a major problem for our analysis. The only assumption we must make is that the exact moment at which a newly hired worker entered his job does not affect his wage at the sampling date. This assumption seems very reasonable, given the fact that workers in general get salary increases only once a year.

A big advantage of this administrative data set compared to survey data is its precision. Missing values are rare and some variables are observed in great detail.[16] For example, the data set distinguishes seven different wage components, which together add up to the total compensation for the worker. These components include, besides the basic wage, also personal bonuses, commuting allowances and the monetary value of wage in kind. In our study, we do not need the exact division of the wage, but the fact that all these components are reported allows us to determine the total payoff of each worker.

Besides the wage information the data set also contains background characteristics of both workers and jobs. For example, we know the gender of the individual and her age and educational level. For the educational level of a worker, the survey reports the highest diploma obtained. We transform this into years of completed education using the design of the Dutch schooling system. With respect to the nature of the jobs, information is available on sector of industry, firm size and type of occupation. Sectors are coded according to the Dutch National Industrial Activity Classification (SBI'93) of Statistics Netherlands which is roughly similar to the SIC classification. Firm size is measured by the number of employees of the firm at the sampling moment and the occupation variable distinguishes between seven types of jobs.

Furthermore, we observe what type of contract a worker has. Most workers are covered by a collective employment agreement (CAO), which is bargained over at the sector level, or by some leading firms within the sector. The Minister of Social Affairs and Employment can declare this agreement legally binding for all other firms in the same sector, implying that these firms must offer the same terms of employment to its employees. This is labelled AVV. Some large companies have their own collective employment agreement. Finally, there are also workers who have a bilateral bargained wage contract. These workers are typically employed at higher positions in the firm. It is important to note that the existence of collective labor agreements does not rule out wage dispersion. A typical collective labor agreement provides many different salary scales and to a large extent firms can determine themselves according to which salary scale they will pay the newly hired worker. Furthermore, firms can also use bonuses and allowances to pay a worker a salary that exceeds the CAO wages.

A last important variable measures the job level on an eight-point scale. The lowest value (1) corresponds to jobs that consist of "very simple, continuously repeating activities, for which no education and only a little experience is required and which are performed under direct supervision". At the other end of the spectrum, the highest job level (8) implies "managing large companies or comparable departments or organizations " (Venema, Faas and Samadhan, 2003). The number of

[16] Nevertheless, some measurement error seems present in the data. We discuss this topic in more detail in subsection 4.4.

observations in job level 7 and 8 is relatively small. Therefore, we combine these workers with the ones in job level 6.

For our analysis, we select the workers who flow from unemployment to employment. As argued before, we can isolate the contribution of search frictions to wage dispersion in this way. We further restrict the sample by focussing on workers with an almost full-time job. More precisely, we select all workers who work for at least 32 hours per week, which corresponds to 80% of a typical working week of 40 hours. The rationale for this selection is that the behavior of part-time workers might differ substantially from the behavior that we try to describe in our model. Moreover, the labor markets for part-time and full-time jobs are almost completely separated in the Netherlands, implying that these two groups of workers hardly compete with each other for a job. We also exclude individuals below 23 years of age and above 65 years. Individuals above 65 face mandatory retirement and a lower minimum wage applies to workers below 23 years of age. Hence, both groups cannot be considered to be identical to the rest of the workers.

Because of missing variables, we cannot use the samples of 1992 to 1995 and 1999. Hence, we use data from six waves (1996 to 1998 and 2000 to 2002). We correct the wage data for inflation by using a wage index and calculate the hourly wage for each worker by dividing her monthly wage by the number of hours worked. In section 4.4 we give some descriptive statistics of the sample, but first we describe in the next subsection how we partition the labor market into five segments.

4.3 Segments

In the theoretical model we make two important assumptions about the labor market. First, we assume that, apart from measurement error, differences in search cost are the only source of wage dispersion amongst individuals. Secondly, we consider a labor market in which no new workers can enter and in which the matching probability only depends on the strategy of the agents that are present in the market. In reality, workers obviously earn different wages for many reasons. Therefore, we first create approximately homogeneous segments correcting for observed heterogeneity and then we assume that our model suits each of those segments. The more segments one creates, the more homogeneous the workers in a given segment will be; however, at the same time, the assumption that we do not allow the best worker in segment i to compete with the worst worker in segment $i+1$ becomes more restrictive. As a compromise, we construct five segments. We assume that these segments constitute separate labor markets within the economy and that each worker and each firm is active in exactly one of the five submarkets. Further, we assume that within a segment all workers and all firms are homogeneous.

In order to create the segments, we construct a worker skill index L_s and a job-complexity index L_c, as in Gautier and Teulings (2006). We create the skill index for the workers by regressing the logarithm of an individual's wage w_i, denoted by ω_i, on all his/her observable characteristics: gender, years of education, years of working experience[17] (also squared and cubed), interaction

[17]As common in literature, we define work experience as a function of age and the years of schooling. To be precise, we assume the following relation: experience = (age - years of education - 6) / 50, where rescaling is applied for

terms, and year dummies. Similarly, the job-complexity index is created by regressing ω_i on dummy variables for the sector, the type of contract for the job, the job level, the occupation type, and year dummies. Appendix B provides details.

The estimation results of these regressions are displayed in table 1 and 2. The fit is good and most coefficients are in line with what is usually found in Mincerian type wage regressions. For example, an extra year of education increases log(wage) by 0.075 for school-leavers, but this effect is smaller for more experienced workers. In the job complexity regression, log(wage) is increasing in the job level. The correlation between the skill level and the complexity level is 0.58. Hence, there is positive assortative matching in the labor market: better skilled workers have more complex jobs. We create the segments accordingly. A straightforward way of achieving this, is by defining:

$$\Upsilon\left(L_s, L_c\right) = L_s L_c.$$

Next, we define the five segments as the quantiles of $\Upsilon\left(L_s, L_c\right)$.[18]

If we repeat the skill and the complexity regression for each of the segments separately, we observe indeed that the segments are much more homogeneous than the labor market as a whole. For example, performing the skill regression on the first segment gives an R^2 of only 0.048 while for the whole sample it is 0.358. This means that only a negligible fraction of the wage dispersion in this segment can be attributed to differences in human capital factors like education and experience. The complexity regression can explain a slightly larger part of the wage variation ($R^2 = 0.188$), but again considerably less than for the entire labor market ($R^2 = 0.475$). The same conclusion holds for the other segments. The only segment that calls for some circumspection in the interpretation of the results is the fifth. There the R^2 values of the skill and the complexity regression are 0.222 and 0.256 respectively, implying that a larger part of the heterogeneity is not filtered out.

4.4 Descriptive statistics

In this subsection we present the labor market statistics that we use in the estimation of the model as well as some descriptive statistics of the AVO data set. A first issue is that we discard some observations in order to prevent that our estimate of σ is determined by outliers. Inspection of the data shows that in each segment some individuals are earning either very low or very high wages compared to other workers. It seems likely that the workers earning these wages are not representative for the rest of the segment. They probably earn these very high or low wages for reasons other than the intensity of search. For example, they might possess very specific skills or we might just have misclassified them and put them in the wrong segment. Therefore, we calculate the 10th percentile $w_{0.1}$, the median $w_{0.5}$ and the 90th percentile $w_{0.9}$ of the wage distribution in each segment and we delete observations that are smaller than $w_{0.5} - \frac{3}{2}\left(w_{0.5} - w_{0.1}\right)$ or larger than

reasons of computational convenience.

[18] We have experimented with several other definitions of the segments as well. This did not change any of the main conclusions. The advantage of this one above, for example, defining Υ as $E[w|s,c]$ is that our measure is more conservative in the sense that less wage variation within segments can be explained by observable characteristics.

$w_{0.5} + \frac{3}{2}(w_{0.9} - w_{0.5})$.[19]

After discarding the outliers we still observe that a small but strictly positive fraction (between 0.7% and 1.9%) of the workers in the lowest three segments earns a wage that is lower than the legal minimum wage, which equals 7.51 euros per hour. Given the strict enforcement of labor laws in the Netherlands, it seems highly unlikely that these workers actually earn such a low wage. Therefore, we interpret this phenomenon as evidence of reporting mistakes in either the monthly wage or in the number of worked hours. Our model can easily deal with this, since we explicitly allow for measurement error in the data.

The descriptive statistics of the Dutch labor market are displayed in table 3. As the table shows, the labor market conditions are clearly increasing in the segment number. Compared to workers in a lower segment, workers in a given segment are (i) more likely to be employed and (ii) more likely to search for a job when non-employed. The table also presents the number of vacancies and some characteristics of the wage distribution for each of the segments. As one would expect, the average wage is strictly increasing in the segment index. In the next subsection we explore how these stylized facts affect the search strategy of the various types of workers and we estimate their search cost distributions.

Table 4 presents some descriptive statistics of the AVO data set. The table lists the mean and the standard deviation for several worker and job characteristics. A first observation is that workers in higher segments are better educated. The workers in segment 5 have on average completed almost seven years of education more than the workers in segment 1. This difference corresponds to more than 80% and is strongly significant. Higher segments contain relatively more men than women. There are no large differences in the average age across segments. Another observation is that workers in the higher segments work more often in the service sector and less often in trade or industry. The type of contract and the firm size seem relatively comparable across the segments.

5 Results

5.1 Market equilibrium

We estimate the model for each of the five segments separately. The estimation results are shown in table 5. To ease the reading, the fractions p_a and q_j that appear in the table are reported conditional on searching at least once. A first interesting observation with respect to the search intensity probabilities is that on the one hand the first and the second segment give similar results and on the other hand segment 3, 4, and 5 are much alike. In the first two segments, the majority of the searchers sends out one job application per period. The remaining workers search almost always twice. In the three highest segments, this pattern is reversed. Most individuals searching for a job send out two applications, while a smaller group only searches once. In all five segments, a small fraction of the workers applies to many (i.e. 30) vacancies. The average number of applicants per

[19] If observed wages were normally distributed, this procedure would lead to deleting 2.7% of the observations at both the top and the bottom of the distribution. However, the wage distributions are skewed to the right and this results in a removal of slightly more observations in the right tail (on average 3.8%) than in the left tail (1.4%).

vacancy varies between 2.3 and 3.3. This results in a job offer probability between 29% and 39%, implying that most workers get either zero or one job offer. Between 4% and 8% of the unemployed receive 2 offers.

We find that both the productivity of a match y and the capital cost k are monotonically increasing across segments. The net productivity $y-k$ is also increasing, except between segments 2 and 3, but the difference is only 0.8 and not statistically significant. The net output produced by a filled vacancy is 17.68 euros per hour in segment 1 and increases to 39.54 euros per hour in segment 5. This is approximately 2 to 2.5 times the average wage in each segment, implying that firms capture a considerable part of the total output. The estimate for the unemployment benefits b ranges from 3.60 euros per hour in the lowest segment to 6.26 euros per hour in the highest segment. We find that the legal minimum wage is binding in the two lowest segments, but not in the other three. Hence, in these latter three segments we can identify the reservation wage and obtain an estimate for h, the combined value of home production and utility derived from leisure. It turns out that h is an important component of the reservation wage. The estimates are between 6.11 and 6.52 euros per hour, which corresponds to 60%-80% of the reservation wage. For segments 1 and 2, the minimum wage is binding and we can only identify an upper bound on the value of home production.

Maximization of the likelihood also provides us with an estimate for the density of accepted wages $g_w(w)$. This estimate can be used to calculate the expected wage $E_{G_w(w)}[w]$. The values obtained in this way are also displayed in table 5. They are very close to the values found in the data, which were presented in table 3. The expected wage offer $E_{F_w(w)}[w]$ is always slightly lower, reflecting the fact that lower wage offers are less likely to be accepted than higher ones.

Figure 3 provides a closer look at the fit of the model. There we compare the estimate for $g_{\widetilde{w}}(\widetilde{w})$ to a kernel estimate of the wage density.[20] The figure shows that our model indeed matches the wage distribution very well. We also formally test the fit of the model by performing a Kolmogorov-Smirnov (KS) test.[21] The values of the test statistic are shown in the last row of table 5. It turns out that for segment 2 and 5 the test statistic is below the critical value 1.36. Hence, the test does not reject the null hypothesis that the empirical cdf and the estimated cdf have the same distribution in these two segments. In the other segments, the test statistic is significant. This is however a common finding in the estimation of search models with many observations (see e.g. Postel-Vinay and Robin (2004)).

The estimated match probability m_W for a non employed worker is lowest in segment 1 (3.6%) and highest in segment 5 (7.5%). For the job destruction rate δ, we find an opposite pattern. It is highest in segment 1, where a fraction 3.9% of the matches is destroyed in each period, and monotonically decreases to 1.1% in segment 5. Note that these are probabilities per period. In order to check whether they match the actual probabilities, we convert them to annual values and we average over the segments. Appendix C gives the details. We find an annual aggregate matching

[20] We use a standard normal kernel with bandwith $1.06 s_w n^{-1/5}$, where s_w denotes the standard deviation of w.

[21] We calculate the KS statistic as $\Delta\sqrt{N}$, where N is the number of observations and Δ is the maximum absolute difference between the estimated and the empirical distribution of the observed wages.

probability for the workers that equals 14.1%. The annual matching probability conditional on search in the current period, implied by the model, is 50.1%. The annual aggregate firing probability implied by the model is equal to 5.2%.

These values are close to values given by other sources. Data of Statistics Netherlands[22] indicates that of the non-employed workers at the end of 1999, 11.0% was employed one year later. Broersma, den Butter and Kock (1998) report that over the period 1970-1995, the flow from unemployment to employment divided by the stock of unemployment in the previous period is 55%.[23] Van den Berg and Van der Klaauw (2001) report three-month-unemployment-exit probabilities of 37%-45% for 1982-1994, which is roughly in line with our aggregate four-month-unemployment-exit probability $m_{W|U}$ of 46.5%. The model also performs well for the job loss rate δ. Using again data of Statistics Netherlands, we find that of the employed workers at the end of 1999, 5.2% was non-employed after a year. Our estimate matches this figure perfectly.

For firms we estimate matching probabilities, m_F, to be between 61.3% and 71.5%. The weighted average over the segments equals 66.9%. This value is in line with the matching probability given by Van Ours and Ridder (1992). Using Dutch survey data, they find that 7% of the reported vacancies had been filled four months later. Den Haan, Ramey and Watson (2000) find exactly the same value (but on a quarterly basis) for the US labor market. The matching rates for firms and workers are variables that do not enter the likelihood function and consequently our estimation procedure is not designed to match them. So it is encouraging to see that the predicted values are close to the actual ones.

In order to determine to what extent the good fit of the wage distribution depends on the presence of measurement error, we judge the estimates for the standard deviation σ. We still find values between 0.007 (segment 2) and 0.074 (segment 5). The higher estimate of σ in segment 5 is in line with the fact that we still found considerable heterogeneity there. However, in general we can conclude that the degree of measurement error is small. The estimates for σ are of the same order of magnitude as the values found by Van den Berg and Ridder (1998), who find standard deviations of 0.022 and 0.045. Dey and Flinn (2005) argue that the degree of measurement error that is required to provide a good fit of the model to the data can be considered to be an index of the degree of model misspecification. Such being the case, we conclude that our model gives an adequate description of the labor market.

The good fit is partly caused by the fact that our model allows the densities of accepted wages to be hump-shaped (also if we estimate the model without measurement error). This feature distinguishes our model from those described by, for example, Burdett and Mortensen (1998) and Gautier and Moraga-González (2004) which imply increasing densities. Another interesting result is that the wage distribution in each segment first-order stochastically dominates the distributions in all lower segments. Note that this was not directly visible in the raw data, where the lowest wages in the third segment were lower than in the second segment. By retrieving the true wage

[22] http://statline.cbs.nl.

[23] This number is calculated from different tables in their paper as: $(UO/U) * (F_{ue}/U0) = 0.79 * 290/418 = 0.55$, where UO is total unemployment outflow and F_{ue} is the flow from unemployment to employment.

distribution, our model reveals that this is only the result of measurement error. In figure 4 we present the estimates for the search cost distributions. We find that search costs (measured in the same unit as the wages and the productivity) are in general higher in higher segments.

It is difficult to obtain direct information on the number of applications that workers send out. Van der Klaauw, Van Vuuren and Berkhout (2003) have information on this variable for university graduates. The median number of applications per 4 months (one period in our model) in the period before a job was found is between 4 and 5 while our model predicts that it is close to 2 in the highest segment (which is the relevant one for university graduates).[24] So for this particular group we either underestimate the number of applications or overestimate the length of a period. Remember however that we make the simplifying assumption in our model that each application has an equal probability to be accepted while in reality workers may send some applications to jobs that are far above or below their league. Those applications may have a much lower acceptance probability, implying that the relevant number of applications is less than 4 per period.

5.2 Mean-min ratio

In a recent paper Hornstein, Krusell and Violante (2006, hereafter HKV) argue that a large class of search and matching models is not able to explain the degree of wage dispersion that is observed in reality. Since our model belongs to the same class, we check to what extent our findings are susceptible to this critique.

HKV discuss a specific measure of wage dispersion, which is defined as the ratio between the average wage and the lowest wage paid to employed workers. They show that a closed form expression for this mean-min ratio (Mm) can be obtained in a general class of search and matching models, without making any parametric assumption on the wage offer distribution. As defined before, ρ denotes the replacement rate, i.e. the ratio between b and the average wage. Then Mm is given by

$$Mm = \frac{1 + \frac{m_W}{r+\delta}}{\rho + \frac{m_W}{r+\delta}}.$$

They calibrate their model with US data on m_W and δ, which results in $Mm = 1.036$. However, at the same time they find that in US data sets with wage information the ratio between the average wage and the reservation wage is typically about 1.70 or larger. From this, they conclude that standard search models are not able to explain the observed combination of a low reservation wage and a high matching rate for unemployed workers. A similar point was made in Gautier and Teulings (2006) who focus on the ratio of the competitive and the reservation wage.[25] They argue that low unemployment rates imply small search frictions while substantial wage dispersion implies large search frictions.

In order to check the performance of our model in jointly explaining observed unemployment and wage dispersion, we set m_W and δ at the estimated values that we obtained in the previous

[24] We thank Aico van Vuuren for kindly giving us this information.

[25] They report substantially less wage dispersion than Hornstein et al. (2006). The difference is due to the fact that Gautier and Teulings correct for measurement error and unobserved heterogeneity.

subsection, while keeping ρ at 0.4 and r at 0.016. Then, we calculate the mean-min ratio as predicted by the market equilibrium and we compare this to what we observe in the data. The results of this are given in table 6. If we follow HKV by taking the fifth percentile of the wage distribution ($w_{5\%}$) in each segment as the reservation wage, then the mean-min ratio in the data varies between 1.215 (segment 1) and 1.598 (segment 5).[26] Note that our model differs in one important respect from the standard search models that they discuss. We do not only consider unemployed and employed workers, but we also allow workers to draw a high search cost such that it is optimal to become a non-participant. This matters for the matching probability. If we ignore the non-participants, we find that the matching probability $m_{W|U} = \frac{m_W}{1-p_0}$ is between 0.4 and 0.5 per period. This implies a mean-min ratio Mm_U that varies between 1.037 (segment 5) and 1.073 (segment 1). Those values are very close to the one found by HKV.

In our model however, the unemployed workers are a selective subsample of the total group of non-employed workers, namely the ones who happened to draw a low search cost in the current period and therefore have a large probability to receive a job offer in this period. They realize that in the next period they may draw a high search cost and they take this into account when they determine their reservation wage. We find that the mean-min ratio for the entire group of non-employed workers, Mm_{NE}, is between 1.190 (segment 5) and 1.572 (segment 1). Hence, in our model where workers have a positive probability to become a non-participant in the next period, a much larger part of wage dispersion can be explained by search frictions. The possibility of becoming non-participant and consequently obtaining a very low matching rate in the next period is consistent with a low reservation wage and a high transition rate from unemployment to employment.[27]

5.3 Efficiency

To check whether the Dutch labor market is constrained efficient, we solve the planner's problem for each of the five segments. We use the estimates of the search cost distribution $F_c(c)$, the productivity y, the capital cost k and the home productivity h that we obtained above and maximize equation (27). Note that in the lowest two segments, we cannot identify the exact value of h. In those segments we set $h = 6.11$, the same value as in segment 3. The fact that h is almost the same in segments 3, 4 and 5 makes this defendable. Nevertheless, in section 6 we drop this assumption and calculate the planner's solution for the bounds on h in segments 1 and 2. There, we also check how sensitive our welfare analysis is to different search cost functions, and assumptions about the search cost of the non-employed.

A priori there is no trivial answer to the question whether the number of applications sent by workers in the market equilibrium is too high or too low from a social planner's point of view.

[26] For optimal comparison with HKV, we do not throw away outliers here.

[27] In Albrecht and Vroman (2006), UI benefits fall over time. Their model is therefore also consistent with a low unemployment rate and substantial wage dispersiont. In Albrecht and Axell (1984) workers are assumed to have different values of leisure. If there is enough heterogeneity in reservation values but at the same time, most workers accept most offers, there can be low unemployment together with substantial wage dispersion.

Workers might underinvest in search since they face a standard hold-up problem. They only receive a part of the social benefits of their investments in search and therefore too many workers with high search cost may decide not to send applications. On the other hand, workers could also send too many applications, since they only take into account their own expected payoff and ignore the congestion effects their applications cause in the market: if workers send multiple applications, several firms might offer the job to the same worker, implying that all except one remain unmatched.

What about firm behavior? Albrecht et al. (2006) show that when all workers search two or more times, efficient entry requires full ex ante and full ex post (i.e. Bertrand) competition for workers. This is not the case in our model. There is no full ex ante competition, since the firm that offers the lowest wage in the market receives as many applications as the other firms, and there is no full ex post competition, because a firm that offers the job to a worker with (an) other offer(s) still has a positive expected payoff. So, firms have too much market power making wages too low and entry excessive. If, most workers search only once, as is the case here, vacancy supply can be either too high or too low because of the standard congestion externalities. Opening a vacancy is good for the workers but bad for the other firms. Below we show that if the planner can jointly determine $a(c)$ and the number of firms, he typically increases both vacancy supply and the number of participants. But for the reasons mentioned above, we find that if we impose the planner's participation level to the market, there would be too many vacancies opened in the market.

Table 7 presents the key parameters of both the constrained and the unconstrained planner's solution for each of the segments. We observe three important differences between the market equilibrium and the constrained planner's strategy. First, the planner wants a higher number of firms to enter the labor market. The increase in entry is relatively small (8%) in segment 5, but amounts to more than 50% in segment 1. Second, the planner increases participation since he internalizes the hold-up problem: a considerable group of non-employed workers (15%-30%) must search once rather than zero times. Finally, the planner decreases the number of workers sending two or more applications. These workers (3%-10% of the non-employed) have low search costs, which makes it profitable for them to send so many applications. As described above however, they do not to take into account that their large number of applications increases the probability that multiple firms consider the same candidate, which is socially wasteful. Therefore, it is better if they apply only once per period. For the unconstrained planner we find optimal strategies that are quite similar to the constrained planner's solution, except that the unconstrained planner wants to increase the participation of workers and the entry of firms even more.

Figure 6 displays the effects of these changes in strategy on steady state employment, non-participation, and unemployment. It shows that implementation of the constrained planner's solution would typically increase the employment rate, decrease non-participation and increase unemployment. The increase in unemployment is a direct effect of a higher participation rate: when more workers search for a job, also a larger number will fail to find one. The unconstrained planner lowers both non-participation and unemployment.

Finally, table 7 reports the output that obtains every period in the market equilibrium as well as in the planner's solution. We define the efficiency of the labor market as the ratio between these two values. It turns out that the labor market is not fully efficient. The constrained planner generates a 15% to 20% higher output than the market, while the unconstrained planner does on average about 40% better. This result allows us to divide the total efficiency loss into a part caused by wrong incentives and a part caused by coordination frictions. We find that in the lowest 4 segments wrong incentives and coordination frictions respectively comprise about 45% and 55% of the total inefficiency. In the highest segment, wrong incentives explain a little bit more than half (about 60%) of the inefficiency.

It is important to stress that the inefficiency results largely depend on the interaction between participation of workers and entry of firms. If we did keep either search intensity or the number of firms constant and let the constrained planner optimize over the other variable, then it would turn out that the planner can hardly do better than the market (differences are in the order of 1%). The reason for this is that participation and firm's entry are complementary. A larger number of active firms encourages workers' participation, which, in turn, makes it more attractive for firms to enter. This implies that when one wants to estimate the effect of, for example, an active labor market program, it is crucial to take both factors into account. Otherwise, one will severely underestimate the true effects of the program. This also suggests that the common practice in the program evaluation literature to study the effects of policy changes in isolation can be very restrictive.

Our efficiency results also shed new light on the desirability of a binding but moderate minimum wage. It has the potential to reduce all three externalities that are present in our model. First, by increasing the average wage it makes more workers search once rather than zero times, increasing participation. Second, by compressing the wage distribution it reduces the incentives to search more than once, decreasing rent seeking behavior and coordination frictions. Third, it reduces monopsony power and excessive entry of vacancies. Unemployment benefits that are conditional on searching at least once have similar effects. Simulations show that a marginal increase in either the minimum wage of the UI benefits is welfare improving.[28]

6 Robustness

In our main analysis we made some simplifying and arbitrary assumptions, namely: (i) the value of household production in segments 1 and 2 that could not be identified (because the minimum wage is binding there) is equal to the value of household production in segment 3, (ii) the irrecoverable part of the search cost distribution for the non-participants can be obtained by linear extrapolation, and (iii) the search cost functions $C(a)$ are linear. To what extent do those arbitrary assumptions affect our main results? In order to investigate this we relax (i)-(iii) in the subsequent subsections.

[28] Due to numerical constraints we are not able to derive the socially optimal minimum wage or benefit level. Another reason to only consider marginal changes in the minimum wage or UI benefits is that we cannot rule out multiple equilibria. If we would consider large changes in those parameters, the model could jump to a different equilibrium.

6.1 Value of home production

In table 7 we fixed household production h in the lowest two segments, where the minimum wage was binding, to 6.11, i.e. the same value as in the third segment. Instead, we could use bounds for h, where the lower bound is zero and the upper bound would be the value of h for which the reservation wage is equal to the minimum wage, i.e. $h = 7.08$ for segment 1 and $h = 6.74$ for segment 2. The different values for the household production influence the estimated search cost distributions, as is shown in figure 5. Not surprisingly, we find that the estimated search costs are higher for lower values of h. After all, lower values of h imply larger benefits of search. In order to have the same values $p_0, ..., p_S$ (that maximize the likelihood) in equilibrium, the costs of search must be higher as well in that case.

Hence, the value of h affects the planner's solution in two ways: directly by changing the contribution of a non-employed worker to total output, and indirectly via the estimated search cost distribution. Table 8 shows that the latter effect dominates. The constrained planner sets p_0 at a higher value for the lower bound of h than for the upper bound. The main conclusion that participation should be increased and that a small fraction of workers sends too many applications however remains. Furthermore, the efficiency loss due to wrong incentives is similar to what was found in table 7.[29] We can conclude that the assumption about home production in the lowest two segments does not affect our main conclusions.

6.2 Search cost for non-participants

Since the search cost of workers who decide not to search are in principle irrecoverable, we made a parametric assumption, namely that their search cost could be obtained by linear extrapolation. In this subsection we relax this assumption by considering bounds for the search cost of non-participants. The planner generates a higher output as the search cost of the non-participants is lower. The upper bound to the planner's solution is therefore obtained when all non-participants have a search cost that is equal to Γ_1, i.e. the marginal benefit of the first application. They cannot have a lower search cost, since otherwise becoming a non-participant would not have been an utility maximizing choice. On the other hand, the lower bound to the planner's solution arises when all non-participants have an infinitely large search cost. Table 9 presents both bounds to the constrained planner's solution. In order to ease comparison, the table also displays again the market equilibrium and the planner's solution in case of linear extrapolation, as they were given in table 7.

Not surprisingly, the planner keeps p_0 equal to the value in the market equilibrium, when all non-participants have infinitely large search costs. Moreover, he changes the number of firms only marginally in that case. Like in the linear case however, the planner does not want workers to send more than one application. The average gain relative to the market equilibrium is about 10%, which is not very different from the value found with linear extrapolation. If we assume that all

[29] Note that a change in h also affects the calculated value for the market output. The relative difference between market and planner's output is therefore the most informative measure.

non-participants have search cost Γ_1, then it is optimal for the planner to let everybody search once ($p_1 = 1$). He increases the number of firms to almost 1 in each of the segments. The gain compared to the market now varies between 20% (segment 5) and almost 40% (segment 1 and 2).

In most segments, especially the higher ones, the minimum and maximum search cost case that we consider provide reasonably narrow bounds on the planner's solution. Moreover, both the minimum and maximum case are actually unrealistically extreme. We know that in the labor market non-participants with low and high search costs coexist. A large fraction of the non-participants is unable to work for various reasons, like disability or because they follow an education. At the same time however, certain non-participants can be considered to be marginally attached (MA) to the labor force. Jones and Riddell (1999) show for example that a small fraction of the Canadian non-participants has a positive probability to flow into employment[30]. Hence, the linear extrapolation case, in which some non-participants have relatively low search cost, while others have high search costs, describes the labor market a lot better than the two bounds. Nevertheless, it is encouraging to see that the bounds, despite being unrealistically extreme, are reasonably close to this linear case.

6.3 Search cost function

In section 2 we assumed that different workers can have different search costs but that the search cost function is linear for all workers. So, we considered functions $C(a) = ca$, with $c > 0$. In reality, this assumption might not hold. $C(a)$ will be concave in a if workers invest a lot of time in drafting the first application letter but spend less time on the subsequent ones. On the other hand, $C(a)$ will be convex in a if workers have easy access to a small number of vacancies (e.g. via their network of friends and colleagues), but have to search really hard to find other job openings. Because of this, we relax the linearity assumption in this subsection and consider a very general class of search cost functions.

We allow different workers to have different shapes of the search cost function: in every period each worker i draws a search cost function $C_i(a)$ from a fixed collection of search cost functions. We only make two very weak assumptions about the shape of $C_i(a)$: $C_i(a)$ is (i) equal to zero for $a = 0$ and (ii) weakly increasing in a. Although the workers know the collection of search cost functions from which they draw, the econometrician does not observe this. Therefore, we have to make one more assumption. Note that the collection of search cost functions does not directly enter the maximum likelihood problem, but determines which sets of $\{p_0, ..., p_S\}$ are feasible. Suppose for example that each search cost function $C_i(a)$ consists of a (stochastic) fixed cost for the first application and that all other applications can be send for free. This would imply that workers either send 0 or S applications and never $1, 2, ..., S-1$ applications. Hence, only solutions in which $p_0 + p_S = 1$ would be possible in that case. Since this solution generates a lower likelihood than

[30] If those workers are included in unemployment they would consist of 25-30% of the unemployed. For the Netherlands, an upper bound estimate of the number of marginally attached workers is 7.5% of the labor force (calculated as all non-disabled workers in 1999 below age 54 who are available, but not necessarily immediately, for 12 hours or more, including school leavers, see Bijsterbosch and Nahuis (2000)).

the solution that we obtain, this collection of search cost functions is not feasible.

In fact, no collection of search cost functions could generate a higher likelihood value than we obtain, since our linear search cost functions did not restrict the set $\{p_0, ..., p_S\}$ in any way. Hence, we impose that the collection of search cost functions supports the ML estimates found in section 5. This implies that we allow some workers to have a search cost function that consists of a fixed cost only, but not too many. Some workers must have a different search cost function (e.g. moderately concave, linear or even convex). Otherwise, no worker would send one or two applications, while the ML estimates indicate that we need a substantial amount of such workers in order to fit the data well.

This condition on the collection of search cost functions implies that the solution that we find for the market equilibrium does not change, because it maximizes (28) and remains feasible. The planner's solution will change, but it can be bounded. First, note that we can identify for a worker who applied to $\widehat{a} > 0$ jobs that the total cost she makes is below a certain threshold, namely the sum of the marginal benefits of searching $\widehat{a}$ times: $C_i(\widehat{a}) < x_{\widehat{a}} = \sum_{l=1}^{\widehat{a}} \Gamma_l$. Since the total search costs are weakly increasing in a, the cost for this worker of searching $a \in \{1, ..., \widehat{a} - 1\}$ times is also at least 0 and at most $x_{\widehat{a}}$. Hence, unlike in the linear case, we cannot rule out that a worker who sends 20 applications can pay more for sending 5 applications than a worker who actually sent 5 applications. Similarly for $a > \widehat{a}$, the total search cost are at least $\sum_{l=\widehat{a}+1}^{a} \Gamma_l$. For example, consider a worker who applied once and assume that the marginal benefits of sending 2 applications is 15. The total cost of sending 1 application must have been at least 0 and therefore the total cost of sending 2 applications for this worker must be at least 15 (otherwise the worker would have sent two applications).

For each segment we can now calculate a 31 by 31 matrix where cell ij contains the minimum amount of search cost of sending j applications for a worker who has actually sent i applications. We can do the same for the maximum search cost. Using those matrices, the planner can then determine his solutions. Note that all workers now have different search costs than in the baseline linear case. As a consequence, the option value of search changes, which results in a different estimate for the value of home production.[31] This does not affect the market output, since the change in h is exactly offset by the change in expected search costs in equilibrium. However, this is not true for the planner's solution. Hence, we let the planner take the new estimate for h into account when he determines his solution. Table 10 gives the results and what strikes is that the bounds are quite tight. In fact, they are very similar to the bounds found in table 9, where we only relaxed the search cost of the non-participants.[32]

Obviously, we again find that the lower the search costs are, the higher output is and the smaller the desired fraction of non-participants p_0 is. For the high search cost case, the planner's solution is close to the market equilibrium. The planner lets only a small fraction of the workers search and it is optimal that some workers send two applications. The intuition for the latter result

[31] This was not the case in the previous subsection. There only the search costs of non-participants changed. Since they do not search, the option value remains unchanged.

[32] The bounds are not necessarily wider than in table 9, because the value of h has changed.

is the following: since there are few searchers, the matching process makes the socially marginal benefits of sending two applications positive. Moreover, for the workers who send two applications in equilibrium, the marginal cost of their second application is zero in the high search cost case. The marginal social benefits of sending three or more applications remain negative because of congestion externalities, so the planner never lets workers search three or more times.[33] In the low search cost case, workers should not send two applications, because there are already many workers who send one application and the expected queue length is such that the marginal social benefits of two applications are negative. Finally, in the high search cost case the planner lets the workers who searched thirty times become non-participants. Letting these workers search once (as in the low search cost case) is not a good strategy, since the (maximum) cost for them of applying once is the same as the cost of applying thirty times.

We can conclude that if we consider a very general class of search frictions, the bounds on the planner's solution hardly change. Although the estimate of the value of home production changes and although we can no longer conclude that it is never desirable that workers send out more than one application, the main message remains that given our endogenous matching process, participation is generally too low and unemployed workers should not send too many applications.

7 Related literature

In this section we relate our paper to the existing literature. First, our model is very similar to the noisy search model of Burdett and Judd (1983) where (in labor market terminology) workers can receive multiple offers. As in Kandel and Simhorn (2002) we allow for the possibility that applications are rejected. We extend their model by endogenizing search intensity and the distribution of job offers. Since we allow for coordination frictions in the matching process, increasing the average search intensity does not make the model converge to the Walrasian equilibrium like in their model. Stern (1989) also estimates a simultaneous job search model but he has an exogenous wage offer distribution.

Albrecht and Axell (1984) also get wage dispersion due to worker heterogeneity. Their heterogeneity is in terms of reservation wages while ours is in terms of search costs which gives us a continuous rather than a discrete wage distribution. Bontemps, Robin and Van den Berg (2000) and Mortensen (2003) focus on heterogeneity on the firm side. Bontemps, Robin and Van den Berg (1999) have heterogeneity on both the worker and firm side. Introducing firm heterogeneity in the Burdett-Mortensen (1998) model of on-the-job search gives a good fit of the wage distribution. All the introduced heterogeneity in the above mentioned papers is motivated by the fact that wage data do not fit the mixed-strategy wage distributions implied by the models. Burdett and Mortensen (1998), Burdett and Judd (1983) and Gautier and Moraga-González (2004) all fail

[33] Albrecht et al. (2006) call this the second coordition problem: the more applications workers send, the larger the probability that multiple firms consider the same candinate. This negative effect of multiple applications can outweight the reduction of the first coordination problem (more applications make it less likely that a firm receives no applications).

to produce hump-shaped distributions. We show that simply allowing for ex post heterogeneity in search cost gives a very good fit of the wage distribution. Basically, the fat right tail within a segment suggests that there is a small fraction of workers with low search cost, receiving many offers.

There are various models with endogenous search intensity. Benhabib and Bull (1983) consider the optimal number of applications in a partial search model with an exogenous wage distribution where, as in our model, workers take the highest offer. In Mortensen (1986), workers can increase the job offer arrival rate by spending more time on search. Bloemen (2005) estimates this model and Van der Klaauw, Van Vuuren and Berkhout (2003) estimate an extension of this model on a sample of university graduates where they allow search intensity before graduation to be time-varying. Christensen et al. (2005) estimate a wage posting model where workers can make investments to increase the job offer arrival rate. The congestion externalities of multiple applications that are present in our model are absent in their model. Albrecht et al. (2004) derive a matching function with multiple applications. More applications make it less likely that a vacancy has no applicants but more likely that multiple firms consider the same candidate.[34] The matching rate is determined by the interaction between those coordination frictions. The aggregate matching function is typically first increasing and then decreasing in average search intensity. This paper extends this matching framework by allowing for heterogeneity in search cost and is the first one which estimates it simultaneously with the wage distribution. This is important for policy analysis because wage policies affect search intensity and policies that affect search intensity will also affect the wage distribution.

In principle, our model allows a non-employed worker to be in any of 30 different search states, each referring to the number of applications she sends. In the macro search literature, the focus has been more on the distinction between participation and non-participation. We have defined non-participation as workers who do not apply to any job but one could alternatively make a distinction between workers sending many or few applications. For simplicity we only consider heterogeneity in search cost to drive participation and search intensity but in Pries and Rogerson (2004) variations in market productivity drive the participation decision while in Pissarides (2000) and Garibaldi and Wasmer (2005) variations in home productivity determine participation. In Frijters and Van der Klaauw (2006), (true) duration dependence of unemployment can push the reservation wage below the value of home production.

There are many other structural estimates of search models, we mention just a few. Besides the ones mentioned above, Eckstein and Wolpin (1990) have estimated the Albrecht-Axell model, Van den Berg and Ridder (1998) estimate the Burdett-Mortensen model and Postel-Vinay and Robin (2004) estimate an on-the-job search model with Bertrand competition between the poaching and the incumbent firm. To our knowledge, there does not exist previous work estimating a labor market version of the Burdett and Judd (1983) model with rationing as in Gautier and Moraga-González

[34] In current work, it has become standard to define search intensity by the number of simultaneous job applications workers send out, see Albrecht, Gautier and Vroman (2003, 2006), Gautier and Moraga-González (2004), Gautier and Wolthoff (2006), Galenianos and Kircher (2005), Kircher (2007), Shimer (2004), Chade and Smith (2006).

(2004), which is what we do here.

Finally, there are a couple of other papers that study the general equilibrium effects of labor market policies that increase search intensity. Flinn (2006) estimates a matching model with Nash bargaining and finds potential positive welfare effects of a binding minimum wage. This is consistent with our findings. Davidson and Woodbury (1993) and Blundell, Costa Dias and Meghir (2003) study the general equilibrium effects of giving a subset of workers a wage bonus or subsidy. Both find huge offsetting equilibrium effects and the latter even find a sign reversal since jobs taken by the treatment group would in the absence of the treatment be filled by non-treated workers. Lise, Seitz and Smith (2003) calibrate their equilibrium search model to data from the control group and then simulate a Canadian income assistance program within the model. They show that the model mimics the transition rate of the treatment group but that the total welfare effects are reversed when the general equilibrium effects are taken into account. A similar methodology is applied in Todd and Wolpin (2003). We estimate the equilibrium model from the beginning and then compare the optimal search intensity distribution with the observed one and find that non-participation is too high and unemployed workers search to intensively. In our model, wage subsidy or counseling schemes for a subset of currently unemployed workers will increases their search intensity and individual employment probabilities but at the same time reduce the employment probabilities for other workers. Our results suggest that active labor market programs can best be targeted at the weakly attached workers, i.e. the ones who are non-participant but who are close to the margin of participating.

8 Final remarks

We have presented a structural estimation of the search cost distribution and the implied search intensity of workers. Unlike most of the literature, we have explicitly defined the search intensity as the number of applications that workers send out per period. The model is estimated by maximum likelihood using wages of newly hired workers. We find that in all segments most unemployed workers search once or twice per four months while a small fraction of the searchers (between 0.7 and 2.5%) applies to thirty jobs. We also show that the decentralized market outcome is about 15% to 20% below the constrained planner's outcome who takes the coordination frictions, value of home production, productivity and the search cost distribution as given. It appears that the planner would like some workers to search less. Especially, applying for two or thirty jobs is socially wasteful. Clearly, this conclusion could change if firms are heterogeneous. Then sending a large number of applications can improve the quality of the match. Another important policy lesson of our model is that job creation programs or programs to increase labor force participation have small positive or negative effects in isolation but policies that aim to simultaneously create more jobs and increase participation can be much more effective.

One important real world feature that we left out in the model is on-the-job search. A natural question is therefore how our results would change if we did allow for it. The answer to this question depends on how employed and unemployed workers compete. Since unemployed workers have an unattractive outside option firms would prefer to offer a wage conditional on previous labor market state if they could. In that case, allowing for on the job search would (i) decrease expected job duration and (ii) would increase our estimate for home production. (i) has a similar effect as increasing the match destruction rate δ, which does not affect the estimates for the search fractions p_a, the job offer probability ψ, the productivity y, etc. It only changes the scale of the search cost distribution. (ii) is caused by the fact that the option value of search during unemployment relative to search during employment goes down. In order for the reservation wage to be equal to the lowest wage in the segment, h would have to go up. Decreasing the definition of a period has an impact on the discount factor and only rescales the search cost distribution. Hence, if we consider a period to be two rather than 4 months, the same results would apply if the yearly discount rate would be about 10% rather than the 5% we assume now.

Another important assumption in our model was that firms that fail to hire their candidate cannot offer the job to the next candidate. Albrecht et al. (2006) show that allowing for firms to make shortlists of workers is very tedious. It does reduce coordination frictions and consequently increases the matching rates but recall does not eliminate the coordination frictions because a firm with four candidates can still loose all of them to competing firms. Allowing for recall will increase the social benefits of sending multiple applications.

Compared to the other empirical equilibrium search models in the literature we have modelled the matching process and search intensity with a lot more detail but in other respects our model is simpler. For example, since workers are ex ante identical, unemployment duration follows a geometrical distribution while in reality there typically is negative duration dependence. One way to get positive duration dependence in our framework is by assuming the heterogeneity in search cost to be worker specific such that high-search-cost workers receive fewer offers in each period in expectation. This is a research avenue we plan to pursue in future work.

References

[1] Abbring, J.H., G.J. van den Berg and J.C. van Ours, 2005, The effect of unemployment insurance sanctions on the transition rate from unemployment to employment, *Economic Journal*, vol. 115(505), pp. 602-630.

[2] Albrecht, J.W. and B. Axell, 1984, An equilibrium model of search unemployment, *Journal of Political Economy*, vol. 92(5), pp. 824-840.

[3] Albrecht, J.W., P.A. Gautier, S. Tan and S.B. Vroman, 2004, Matching with multiple applications revisited, *Economics Letters*, vol. 84(3), pp. 311-314.

[4] Albrecht J.W., P.A. Gautier and S.B. Vroman, 2003, Matching with multiple applications, *Economics Letters*, vol. 78(1), pp. 67-70.

[5] Albrecht J.W., P.A. Gautier and S.B. Vroman, 2006, Equilibrium directed search with multiple applications, *Review of Economic Studies*, vol. 73(4), pp. 869-891.

[6] Albrecht J.W. and S.B. Vroman, 2006, Equilibrium search with time-varying unemployment benefits, *Economic Journal*, vol. 115(505), pp. 631-648.

[7] Benhabib J. and C. Bull, 1983, Job search, the choice of intensity, *Journal of Political Economy*, vol. 91(5), pp. 747-764.

[8] Bijsterbosch M.G. and N.J. Nahuis, 2000, *Vergroting arbeidsaanbod een uitdaging voor het beleid* (in Dutch), DNB working paper, Dutch Central Bank (DNB).

[9] Blundell, R., M. Costa Dias and C. Meghir, 2003, *The impact of wage subsidies: a general equilibrium approach*, mimeo, Institute for Fiscal Studies.

[10] Blackwell, D., 1965, Discounted dynamic programming, *Annals of Mathematical Statistics*, vol. 36(1), pp. 226-235.

[11] Bloemen, H.G., 2005, Job search, search intensity, and labor market transitions: an empirical analysis, *Journal of Human Resources*, vol. 40(1), pp. 231-269.

[12] 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*, vol. 40(4), pp. 1039-1074.

[13] 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*, vol. 41(2), pp. 305-358.

[14] Broersma L., F.A.G. den Butter and U. Kock, 1998, A national accounting system for labour market flows : an application to the Netherlands, Research memorandum 59, Vrije Universiteit Amsterdam.

[15] Burdett, K. and K.L. Judd, 1983, Equilibrium price dispersion, *Econometrica*, vol. 51(4), pp. 955-969.

[16] Burdett, K. and D.T. Mortensen, 1998, Wage differential, employer size, and unemployment, *International Economic Review*, vol. 39(2), pp. 257-273.

[17] Christensen, B.J., R. Lentz, D.T. Mortensen, G.R. Neumann, and A. Wervatz, 2005, On-the-job search and the wage distribution, *Journal of Labor Economics*, vol. 23(1), pp. 31-58.

[18] Chade, H. and L. Smith, 2006, Simultaneous search, *Econometrica*, vol. 74(5), pp. 1293-1307.

[19] Davidson, C. and S.A. Woodbury, 1993, The displacement effect of reemployment bonus programs, *Journal of Labor Economics*, vol. 11(4), pp. 575-605.

[20] Den Haan, W.J., G. Ramey and J. Watson, 2000, Job destruction and propagation shocks, *American Economic Review*, vol. 90(3), pp. 482-498.

[21] Dey, M.S. and C.J. Flinn, 2005, An equilibrium model of health insurance provision and wage determination, *Econometrica*, vol. 73(2), pp. 571-627.

[22] Donald, S.G. and H.J. Paarsch, 1993, Piecewise pseudo-maximum likelihood estimation in empirical models of auctions, *International Economic Review*, vol. 34(1), pp. 121-148.

[23] Eckstein, Z. and K.I. Wolpin, 1990, Estimating an equilibrium search model using longitudinal data on individuals, *Econometrica*, vol. 58(4), pp. 783-808

[24] Flinn, C.J., 2006, Minimum wage effects on labor market outcomes under search, matching, and endogenous contact rates, *Econometrica*, vol. 74(4), pp. 1013-1062.

[25] Flinn, C.J. and J.J. Heckman, 1982, New methods for analyzing structural models of labor force dynamics, *Journal of Econometrics*, vol. 18(1), pp. 115-168.

[26] Frijters P. and B. van der Klaauw, 2006, Job search with nonparticipation, *Economic Journal*, vol. 116(508), pp. 239-277.

[27] Galenianos, M. and P. Kircher, 2005, *Directed search with multiple job applications*, PIER working paper 05-022, Penn Institute for Economic Research.

[28] Gautier, P.A. and J.L. Moraga-González, 2004, *Strategic wage setting and random search with multiple applications*, Tinbergen Institute discussion paper 04-063/1, Tinbergen Institute.

[29] Gautier, P.A. and C.N. Teulings, 2006, How large are search frictions?, *Journal of the European Economic Association*, vol. 4(6), pp. 1193-1225.

[30] Gautier, P.A. and R.P. Wolthoff, 2006, *Simultaneous search with heterogeneous firms and ex post competition*, Tinbergen Institute discussion paper 06-032/3, Tinbergen Institute.

[31] Garibaldi, P. and E. Wasmer, 2005, Equilibrium search unemployment, endogenous participation and labor market flows. *Journal of the European Economic Association,* vol. 3(4), pp. 851-882.

[32] Graham, J.W. and A.H. Beller, 1989, The effect of child support payments on the labor supply of female family heads, *Journal of Human Resources*, vol. 24(4), pp. 664-688.

[33] Heckman J.J., 1974, The effect of child-care programs on women's work effort, *Journal of Political Economy*, vol. 82(2), pp. 153-161.

[34] Jones S.R.G. and W.C. Riddell, 1999, The measurement of unemployment: An empirical approach, *Econometrica*, vol. 67(1), pp. 147-61.

[35] Hornstein, A., P. Krusell and G.L. Violante, 2006, *Frictional wage dispersion in search models: a quantitative approach*, CEPR discussion paper 5935, CEPR.

[36] Hortaçsu, A. and C. Syverson, 2004, Product differentiation, search costs, and competition in the mutual fund industry: a case study of S&P 500 index funds, *The Quarterly Journal of Economics*, vol. 119(2), pp. 403-456.

[37] Hong, H. and M. Shum, 2006, Using price distributions to estimate search costs, *Rand Journal of Economics*, vol. 37(2), pp. 257-275.

[38] Kandel E. and A. Simhon, 2002, Between search and Walras, *Journal of Labour Economics*, vol. 20(1), pp. 59-85.

[39] Kiefer, N.M. and G.R. Neumann, 1993, Wage dispersion with homogeneity: the empirical equilibrium search model, *Panel Data and Labour Market Dynamics*, H. Bunzel et al., eds., North Holland, Amsterdam.

[40] Kircher, P., 2007, *Simultaneous directed search with recall*, mimeo, University of Pennsylvania.

[41] Lise, J., S. Seitz and J. Smith, 2003, *Equilibrium policy experiments and the evaluation of social programs*, mimeo.

[42] Meyer, B.D. and, 1996, What have we learned from the Illinois reemployment bonus experiment? *Journal of Labour Economics*, vol. 14, pp. 26-51.

[43] Moraga-Gonzalez, J.L. and M.R. Wildenbeest, 2006, *Maximum likelihood estimation of search costs*, Tinbergen Institute discussion paper 06-019/1, Tinbergen Institute.

[44] Mortensen, D.T., 1986, Job search and labor market analysis, *Handbook of Labor Economics*, O. Ashenfelter and R. Layard, eds., North Holland, Amsterdam.

[45] Mortensen, D.T., 2003, *Wage dispersion: why are similar workers paid differently?*, MIT Press, Cambridge.

[46] Petrongolo, B. and C.A. Pissarides, 2001, Looking into the black box: A survey of the matching function, *Journal of Economic Literature* 39: 390-431, 2001.

[47] Pissarides, C.A., 2000, *Equilibrium unemployment theory*, 2nd edition, MIT Press, Cambridge.

[48] Postel-Vinay, F. and J.M. Robin, 2004, To match or not to match? Optimal wage policy with endogenous worker search intensity, *Review of Economic Dynamics*, vol. 7(2), pp. 297-331.

[49] Pries, M.J. and R. Rogerson, 2004, *Search frictions and labor market participation*, mimeo, Maryland University.

[50] Shimer, R., 2004, *Search intensity*, mimeo, University of Chicago.

[51] Stern, S., 1989, Estimating a simultaneous search model, *Journal of Labor Economics*, Vol. 7(3), pp. 348-369.

[52] Teulings, C.N. and P.A. Gautier, 2004, The right man for the job, *Review of Economic Studies*, vol. 71(2), pp. 553-580.

[53] Todd, P. and K. Wolpin, 2003, *Using experimental data to validate a dynamic behavior model of child schooling on fertility*, mimeo, University of Pennsylvania.

[54] Van den Berg, G.J. and G. Ridder, 1998, An empirical equilibrium search model of the labor market, *Econometrica*, vol. 66(5), pp. 1183-1221.

[55] Van den Berg, G.J. and B. van der Klaauw, 2001, Combining micro and macro unemployment duration data, *Journal of Econometrics*, vol. 102(2), pp. 271-309.

[56] Van den Berg, G.J. and B. van der Klaauw, 2006, Counseling and monitoring of unemployed workers: theory and evidence from a controlled social experiment, *International Economic Review*, vol. 47(3), pp. 895-936.

[57] Van der Klaauw, B., A van Vuuren and P. Berkhout, 2003, *Labor market prospects, search intensity and the transition from college to work*, CEPR discussion paper 4515, CEPR.

[58] Van Ours, J.C. and G. Ridder, 1992, Vacancies and the recruitment of new employees, *Journal of Labor Economics*, vol. 10(2), pp. 138-155.

[59] Van Ours, J.C. and G. Ridder, 1993, Vacancy durations: search or selection?, *Oxford Bulletin of Economics and Statistics*, vol. 55(2), pp. 187-198.

[60] Venema, P.M. and A. Faas and J.A. Samadhan, 2003, *Arbeidsvoorwaardenontwikkelingen in 2002* (in Dutch), Arbeidsinspectie, The Hague.

[61] Wolpin, K.I., 1987, Estimating a structural job search model: the transition from school to work, *Econometrica*, vol. 55(4), pp. 801-818.

Appendix

A Derivations

A.1 Equal profit condition

The payoff of a firm is $y-k-w$ in case of a match. A firm matches if it offers its candidate a higher wage than all other firms competing for the same worker. The probability for a firm to have at least one applicant is equal to $1-e^{-\phi}$. The conditional probability that the candidate has sent a applications is given by $\frac{ap_a}{\sum_{i=1}^{S} ip_i}$. The $a-1$ other applications of the candidate result in $j \in \{0,1,...,a-1\}$ other job offers with probability $\chi(j|a-1)$, which are all lower with probability $F_w^j(w)$. Hence, the expected payoff of a firm offering w is given by

$$\pi(w) = (y-k-w)\left(1-e^{-\phi}\right)\sum_{a=1}^{S}\frac{ap_a}{\sum_{i=1}^{S} ip_i}\sum_{j=0}^{a-1}\chi(j|a-1)\,F_w^j(w).$$

By using $\left(1-e^{-\phi}\right)=\phi\psi=\frac{\sum_{i=1}^{S} ip_i}{\theta}\psi$ and the definition of $\chi(j|a-1)$, we can simplify this expression as follows

$$\begin{aligned}
\pi(w) &= (y-k-w)\,\phi\psi\sum_{a=1}^{S}\frac{ap_a}{\sum_{i=1}^{S} ip_i}\sum_{j=0}^{a-1}\binom{a-1}{j}\psi^j(1-\psi)^{a-1-j}F_w^j(w)\\
&= (y-k-w)\frac{1}{\theta}\sum_{a=1}^{S}ap_a\sum_{j=0}^{a-1}\binom{a-1}{j}\psi^{j+1}(1-\psi)^{a-1-j}F_w^j(w)\\
&= (y-k-w)\frac{1}{\theta}\sum_{a=1}^{S}\sum_{j=1}^{a}ap_a\binom{a-1}{j-1}\psi^j(1-\psi)^{a-j}F_w^{j-1}(w)\\
&= (y-k-w)\frac{1}{\theta}\sum_{j=1}^{S}jq_jF_w^{j-1}(w).
\end{aligned}$$

A.2 Likelihood

The first step in the estimation of the model is to calculate the fraction of non-searchers p_0. From equation (3), it follows that it equals the ratio of the fraction of non-participants in the population to the fraction of non-employed:

$$p_0 = \frac{n}{1-e}. \tag{29}$$

The other fractions p_a are estimated by maximizing the likelihood that the observed wages are generated by our model. Note that a distribution for p_a, together with v, u and the urn-ball type of matching function that follows from (4) and (5), implies a job offer probability ψ. This job offer probability is the key parameter in the mapping from the number of applications p_a to the number of job offers q_j. Given estimates for the net productivity $\widehat{y}$ and the lower bound $\underline{w}$ of the support of the wage offer distribution $F_w(w)$, we can calculate the upper bound $\overline{w}$ of the support by using equation (21). Then, we can solve equation (20) to get the full wage offer distribution.

Note that in data one typically does not observe all wage offers, but only the wage offers that have been accepted by the workers. Since workers can compare wage offers, high wage offers are more likely to be accepted than low wage offers. This implies that the distributions of the offered wages and of the accepted wages differ from each other. Let $G_w(w)$ denote the distribution of the wages accepted by the non-employed workers. In order to derive an expression for $G_w(w)$, consider a worker who receives $j > 0$ job offers. She will only accept a wage that is lower than some value w if all her j job offers are lower than this value. As a result, the following relationship between $G_w(w)$ and $F_w(w)$ holds:

$$G_w(w) = \frac{\sum_{j=1}^{S} q_j F_w^j(w)}{1 - q_0}. \tag{30}$$

It is straightforward to show that $G_w(w)$ first-order stochastically dominates $F_w(w)$. Taking the first derivative of this expression with respect to w gives the density of the accepted wages:

$$g_w(w) = \frac{\sum_{j=1}^{S} j q_j F_w^{j-1}(w) f_w(w)}{1 - q_0}, \tag{31}$$

where $f_w(w)$ denotes the density function of the posted wages. An expression for this density can be derived by applying the implicit function theorem to equation (20). This yields

$$f_w(w) = \frac{\sum_{j=1}^{S} j q_j F_w^{j-1}(w)}{(y - k - w) \sum_{j=2}^{S} j(j-1) q_j F_w^{j-2}(w)}.$$

These equations show that $g_w(w)$ only depends on the productivity and the capital cost via the difference $y - k$. Hence, in the maximum likelihood estimation only the net productivity $\widehat{y}$ is identified. Ex post however, we can obtain estimates for y and k by using the equality $y = \widehat{y} + k$ and by rewriting the free entry condition (26) in the following way:

$$k = \frac{1}{r + \delta} \frac{1}{\theta} q_1 (\widehat{y} - \underline{w}).$$

As we explain in the main text, we estimate the lower bound $\underline{w}$ of the support of the wage distribution as a parameter in the maximum likelihood procedure. Together with the estimate for $\widehat{y}$, this implies a value for the upper bound $\overline{w}$. In order to explain observations outside the bounds of the support, we allow for measurement error. To be precise, we assume that the observed wage $\widetilde{w}$ depends on the true wage w and a random error term ε in a multiplicative way:

$$\widetilde{w} = w\varepsilon.$$

The error term ε has a log-normal distribution with parameters $\mu = 0$ and $\sigma^2 = \mathrm{var}(\log(\varepsilon))$. We estimate σ as a parameter in the maximum likelihood procedure. The density of the observed wages can then be obtained by integrating over all possible values of the error term. If a wage $\widetilde{w}$ is observed, the error term must have been in the interval $[\widetilde{w}/\overline{w}; \widetilde{w}/\underline{w}]$. Hence, the density of the observed wages $g_{\widetilde{w}}(\widetilde{w})$ is equal to

$$g_{\widetilde{w}}(\widetilde{w}) = \int_{\widetilde{w}/\overline{w}}^{\widetilde{w}/\underline{w}} g_w(\widetilde{w}/\varepsilon) \frac{1}{\varepsilon} \eta(\varepsilon) \, d\varepsilon, \tag{32}$$

where $1/\varepsilon$ is the Jacobian of the transformation, $\eta(\varepsilon)$ denotes the log-normal density and $g_w(w)$ is given by (31). The integral in this equation must be calculated numerically, since it depends on $F_w(w)$, for which no explicit expression exists. Assuming independence of the N observations, the maximum likelihood problem is then given by

$$\max_{p_1,\dots,p_S,\sigma,\underline{w},\widehat{y}} \frac{1}{N} \sum_{i=1}^{N} \log g_{\widetilde{w}}(\widetilde{w}_i),$$

subject to the conditions $\sum_{a=0}^{S} p_a = 1$, $p_a \in [0,1]$ $\forall a$ and $w_{\min} \leq \underline{w} \leq \widehat{y}$.

A.3 Search cost distribution

The maximum likelihood estimation provides us with estimates for $p_0, \dots, p_S, \underline{w}$ and y. Using these estimates, we can derive cutoff points of the search cost distribution according to equation (17). This requires the calculation of the marginal gains from search Γ_a as given by equation (16). Note that this variable depends on the integral $\int_{w_R}^{\infty} w dF_w(w)^j$. To simplify the calculation of this integral, we apply a change of variables.

First, invert equation (20) and denote the inverse function of $F_w(w)$ by $w(z)$

$$w(z) = y - k - \frac{(y - k - \underline{w}) q_1}{\sum_{j=1}^{S} j q_j z^{j-1}}. \tag{33}$$

Then, we can write:

$$\int_{w_R}^{\infty} w dF_w^j(w) = \int_{\underline{w}}^{\overline{w}} w dF_w^j(w) = \int_0^1 j w(z) z^{j-1} dz,$$

Substituting this in equation (15) gives

$$\zeta_a = \sum_{j=1}^{a} \chi(j|a) \int_0^1 j \left(\widehat{y} - w_R - \frac{(\widehat{y} - \underline{w}) q_1}{\sum_{i=1}^{S} i q_i z^{i-1}} \right) z^{j-1} dz$$

The marginal gains from an additional application can then be calculated from (16). The equilibrium value for the separation rate δ that we need in this calculation follows from the steady state condition given in equation (1):

$$\delta = \frac{(1 - q_0)(1 - e)}{e}.$$

This procedure gives us S cutoff points $(\Gamma_1, \dots, \Gamma_S)$ of the search cost distribution $F_c(c)$. For some purposes, e.g. for assessing the efficiency of the market equilibrium, we need an estimate of the full distribution (i.e. for every possible value of c). On the interval $[0, \Gamma_1]$ we obtain this estimate by using linear interpolation between the cutoff points:

$$F_c(c) = \ \sum_{j=i+1}^{S} p_j + \tfrac{p_i}{\Gamma_i - \Gamma_{i+1}} (c - \Gamma_{i+1}) \quad \forall c \in [\Gamma_{i+1}, \Gamma_i) \text{ and } i = \{1, ..., S\} \ ,$$

where we define $\Gamma_{S+1} = 0$.

For $c > \Gamma_1$ we assume that the search cost distribution keeps increasing linearly, with the same slope as just before Γ_1, until it reaches 1. Hence, on this interval $F_c(c)$ is given by

$$F_c(c) = \begin{cases} 1 - p_0 + \frac{p_1}{\Gamma_1 - \Gamma_2} (c - \Gamma_1) & \forall c \in [\Gamma_1, \Gamma_0) \\ 1 & \forall c \geq \Gamma_0 \end{cases}, \text{ where } \Gamma_0 = \Gamma_1 + \frac{p_0}{p_1} (\Gamma_1 - \Gamma_2).$$

For solving the planner's problem we also need estimates for the unemployment benefits b and the household production h. The value for b equals the product of the replacement rate ρ and the average wage:

$$b = \rho \int_{\underline{w}}^{\overline{w}} w dG_w(w) = \rho \left(w_R + \frac{1}{1 - q_0} \sum_{a=1}^{S} p_a \zeta_a \right).$$

Use this and (14) to derive an estimate for h. Using the same simplifications as above, we can rewrite this expression as

$$h = w_R - \int_0^\infty \max_a \left(I_{a>0} b + \frac{1}{r + \delta} \zeta_a - ca \right) dF_c(c) \tag{34}$$

Next, partition the support of $F_c(c)$ into the intervals $[\Gamma_{S+1}, \Gamma_S)$, $[\Gamma_S, \Gamma_{S-1})$, ..., $[\Gamma_2, \Gamma_1)$, $[\Gamma_1, \Gamma_0]$, where Γ_{S+1} and Γ_0 are the lower bound and the upper bound of the support of $F_c(c)$. Due to the linear interpolation, $f(c)$ is constant on each of these intervals. Let f_a denote the value of $f(c)$ on the interval $[\Gamma_{a+1}, \Gamma_a)$. Then the following expression holds:

$$f_a = \frac{F(\Gamma_a) - F(\Gamma_{a+1})}{\Gamma_a - \Gamma_{a+1}} = \frac{p_a}{\Gamma_a - \Gamma_{a+1}}.$$

Substituting this in (34), we can write

$$\begin{aligned} h &= w_R - \sum_{a=1}^{S} \int_{\Gamma_{a+1}}^{\Gamma_a} \left(I_{a>0} b + \frac{1}{r + \delta} \zeta_a - ca \right) \frac{p_a}{\Gamma_a - \Gamma_{a+1}} dc \\ &= w_R - \frac{1}{r + \delta} \sum_{a=1}^{S} p_a \zeta_a + \frac{1}{2} \sum_{a=1}^{S} a p_a (\Gamma_a + \Gamma_{a+1}) - b(1 - p_0) \\ &= w_R - \frac{1}{r + \delta} \sum_{a=1}^{S} p_a \left(\zeta_a + \frac{1}{2} a (\zeta_{a+1} - \zeta_{a-1}) \right) - b \left(1 - p_0 - \frac{1}{2} p_1 \right), \end{aligned}$$

where we define $\zeta_{S+1} = \zeta_S$ to simplify notation. This relates h to variables that we can estimate or directly observe.

B Labor market segments

In order to create the segments, we construct a worker skill index L_s and a job-complexity index L_c, as in Gautier and Teulings (2006). For the worker skills we assume the following linear relationship

$$\omega = X_s \beta_s + \varepsilon_s,$$

where β_s is a vector of coefficients and ε_s is an error term. The matrix X_s contains the explanatory variables: gender, years of education, years of working experience[35] (also squared and cubed), interaction terms, and year dummies. Next, we define the skill L_s of an individual as the predicted value following from this regression:

$$L_s = X_s \hat{\beta}_s,$$

where $\hat{\beta}_s = (X_s' X_s)^{-1} X_s' \omega$. Likewise, we construct a complexity measure for the jobs. We regress the logarithm of the wage paid by firm for this job on several job and firm characteristics:

$$\omega = X_c \beta_c + \varepsilon_c,$$

where X_c includes a constant, dummy variables for the sector, the type of contract for this job, the job level, occupation, and year dummies.[36] The complexity L_c of the job is defined as the predicted value of the regression:

$$\begin{aligned} L_c &= X_c \hat{\beta}_c \\ &= X_c \left(X_c' X_c\right)^{-1} X_c' \omega. \end{aligned}$$

C Flow probabilities

As discussed in section 4, we assume that a year consists of three periods. This implies that a worker can flow from employment in year τ (time t) to non-employment in year $\tau + 1$ (time $t + 3$) in four different ways. She can loose her job at the beginning of either period $t + 1$, $t + 2$ or $t + 3$, and remain non-employed after that. Alternatively, she can loose her job at $t + 1$, get a new job at $t + 2$ and loose it again at $t + 3$. Hence, the yearly separation rate $\delta_{W,3}$ for the workers is given by

$$\begin{aligned} \delta_{W,3} &= \delta (1 - m_W)^2 + (1 - \delta) \delta (1 - m_W) + (1 - \delta)^2 \delta + \delta m_W \delta \\ &= \delta \left(\delta^2 + m_W^2 + 2\delta m_W - 3\delta - 3m_W + 3\right). \end{aligned}$$

Expressions for the yearly separation rate of firms ($\delta_{F,3}$) and the annual matching probability for workers ($m_{W,3}$) and firms ($m_{F,3}$) can be derived in a similar way. The per-period matching

[35] As common in literature, we define work experience as a function of age and the years of schooling. To be precise, we assume the following relation: experience = (age - years of education - 6) / 50, where rescaling is applied for reasons of computational convenience.

[36] Although we also observe the size of the firm, we do not include this variable in the job complexity regression to avoid endogeneity problems.

probability conditional on search is given by

$$m_{W|U} = 1 - \frac{q_0 - p_0}{1 - p_0} = \frac{1 - q_0}{1 - p_0}.$$

The annual matching probability conditional on search in the current period therefore equals

$$m_{W|U,3} = m_{W|U} (1 - \delta)^2 + \left(1 - m_{W|U}\right) m_W (1 - \delta) + \left(1 - m_{W|U}\right) (1 - m_W) m_W + m_{W|U} \delta m_W.$$

In order to be able to compare the estimated probabilities to the actual ones, we aggregate over the segments. For this we need to know the relative size s_i (i.e. the total mass of workers) for each of the segments. Normalize the size of the first segment to 1. Note that we defined the segments in such a way that the expected number of people flowing from non-employment to employment is the same in each one of them. Hence, the relative size s_i of segment $i > 1$ is defined by

$$m_{W,3,i} (1 - e_i) s_i = m_{W,3,1} (1 - e_1),$$

where $m_{W,3,i}$ denotes the annual matching probability and e_i the employment rate in segment i. The total number of matches formed in the market is then equal to $\sum_{i=1}^{5} m_{W,3,i} (1 - e_i) s_i$ and the total number of non-employed workers equals $\sum_{i=1}^{5} (1 - e_i) s_i$. So, the aggregate annual matching probability $\overline{m_{W,3}}$ can be calculated as follows:

$$\overline{m_{W,3}} = \frac{\sum_{i=1}^{5} m_{W,3,i} (1 - e_i) s_i}{\sum_{i=1}^{5} (1 - e_i) s_i}.$$

Aggregating $m_{W|U,3}$ can be done in a similar way. Likewise, the following expression holds for the aggregate annual separation probability $\overline{\delta_{W,3}}$:

$$\overline{\delta_{W,3}} = \frac{\sum_{i=1}^{5} \delta_{W,3,i} e_i s_i}{\sum_{i=1}^{5} e_i s_i},$$

where $\delta_{W,3,i}$ is the yearly separation rate in segment i.

The aggregate annual matching rate for the firms can be calculated according to

$$\overline{m_{F,3}} = \frac{\sum_{i=1}^{5} m_{F,3,i} v_i N_{F,i} s_i}{\sum_{i=1}^{5} v_i N_{F,i} s_i},$$

where v_i and $N_{F,i}$ respectively denote the fraction of vacancies and the measure of firms in segment i.

D Tables and figures

Variable	Estimate	Std.err.		Variable	Estimate	Std.err.	
Constant	1.192	0.116	*	**Year**			
				1997	-0.012	0.015	
Education				1998	0.010	0.014	
Years of education	0.075	0.008	*	2000	0.054	0.014	*
				2001	0.036	0.015	*
Gender				2002	0.041	0.016	*
Male	-0.047	0.032					
				Interaction effects			
Experience				Educ. × Experience	-0.110	0.066	
Experience	3.792	0.930	*	Educ. × Experience^2	0.404	0.183	*
Experience^2	-9.955	2.283	*	Educ. × Experience^3	-0.378	0.156	*
Experience^3	7.805	1.733	*	Male × Experience	0.349	0.355	
				Male × Experience^2	0.790	1.016	
Statistics				Male × Experience^3	-1.343	0.817	
Observations	5801						
R^2	0.358						

* = significant at 5% level.
Reference groups: female, 1996.

Table 1: Estimation results of the skill regression

Variable	Estimate	Std.err.		Variable	Estimate	Std.err.	
Constant	2.251	0.023	*	**Job level**			
				Level 2	0.027	0.014	*
Sector				Level 3	0.166	0.014	*
Industry	-0.026	0.019		Level 4	0.327	0.016	*
Education	-0.013	0.028		Level 5	0.570	0.023	*
Construction	0.133	0.022	*	Level 6-8	0.743	0.046	*
Trade, reparation	-0.052	0.021	*				
Hotel, catering	0.055	0.031		**Occupation type**			
Transport, communic.	-0.049	0.023	*	Administrative	-0.031	0.012	*
Financial services	-0.001	0.031		Automation	-0.022	0.021	
Other services	-0.064	0.022	*	Commercial	-0.029	0.015	*
Health care	0.007	0.026		Service providing	-0.062	0.011	*
Culture, recreation	-0.052	0.036	*	Creative	-0.021	0.031	
				Management	0.193	0.039	*
Year							
1997	-0.025	0.013		**Coll. empl. agreement (CAO)**			
1998	0.001	0.013		AVV	0.081	0.015	*
2000	0.023	0.013		Company CAO	-0.047	0.019	*
2001	0.014	0.013		No CAO	-0.052	0.010	*
2002	0.051	0.015	*				
Statistics							
Observations	5801						
R^2	0.475						

* = significant at 5% level.
Reference groups: agriculture, 1996, industry CAO, 1-4 employees, level 1, simple technical activities.

Table 2: Estimation results of the complexity regression

	Segment				
	1	**2**	**3**	**4**	**5**
Labor market states					
Employment (e)	0.479	0.615	0.736	0.790	0.875
Unemployment (u)	0.043	0.033	0.030	0.027	0.022
Non-participation (n)	0.477	0.351	0.234	0.183	0.103
Vacancies					
Number of vacancies (vN_F)	0.030	0.023	0.021	0.019	0.015
Non-participation					
Number of non-participants (p_0)	0.917	0.914	0.885	0.870	0.826
Wage distribution					
Lowest observed wage	7.40	7.42	7.36	8.08	9.46
Highest observed wage	12.07	14.08	15.17	16.81	27.25
Mean observed wage	9.00	10.07	10.73	12.17	15.75

Table 3: Values of the exogenous parameters per segment

	Segment 1		Segment 2		Segment 3		Segment 4		Segment 5	
	mean	**s.d.**	**mean**	**s.d.**	**mean**	**s.d.**	**mean**	**s.d.**	**mean**	**s.d.**
Observations	1043		1179		1153		1070		1022	
Male	0.52	0.50	0.63	0.48	0.71	0.45	0.72	0.45	0.67	0.47
Education	8.05	1.88	10.42	1.88	11.72	2.16	12.93	2.23	14.75	1.38
Age	32.55	9.45	31.83	9.70	31.49	8.78	31.93	8.73	33.50	8.65
Sector										
Agriculture	0.02	0.12	0.03	0.18	0.03	0.18	0.02	0.13	0.02	0.14
Industry	0.30	0.46	0.24	0.43	0.25	0.44	0.18	0.39	0.15	0.36
Education	0.00	0.01	0.01	0.11	0.01	0.10	0.01	0.09	0.01	0.11
Construction	0.01	0.12	0.03	0.16	0.07	0.26	0.21	0.40	0.09	0.29
Trade, reparation	0.17	0.38	0.19	0.39	0.13	0.34	0.12	0.33	0.08	0.27
Hotel, catering	0.04	0.20	0.04	0.19	0.03	0.18	0.02	0.15	0.01	0.10
Transp., communic.	0.04	0.20	0.10	0.30	0.11	0.31	0.04	0.20	0.03	0.17
Financial services	0.01	0.08	0.02	0.14	0.03	0.17	0.06	0.23	0.06	0.24
Other services	0.27	0.45	0.19	0.39	0.23	0.42	0.22	0.42	0.37	0.48
Health care	0.10	0.30	0.13	0.34	0.09	0.28	0.09	0.29	0.13	0.34
Culture, recreation	0.03	0.18	0.02	0.13	0.01	0.11	0.02	0.14	0.04	0.19
Coll. empl. agr.										
Industry CAO	0.55	0.50	0.63	0.48	0.62	0.49	0.62	0.49	0.53	0.50
AVV	0.03	0.16	0.05	0.22	0.06	0.25	0.09	0.29	0.09	0.29
Company CAO	0.10	0.30	0.06	0.23	0.06	0.24	0.05	0.21	0.02	0.14
No CAO	0.32	0.47	0.27	0.44	0.25	0.44	0.25	0.43	0.36	0.48
Firm size										
1-4 employees	0.02	0.16	0.03	0.16	0.05	0.21	0.04	0.20	0.02	0.15
5-9 employees	0.08	0.27	0.06	0.24	0.11	0.31	0.09	0.29	0.09	0.29
10-19 employees	0.12	0.33	0.13	0.33	0.13	0.33	0.13	0.34	0.09	0.28
20-49 employees	0.22	0.42	0.18	0.38	0.17	0.37	0.21	0.41	0.12	0.33
50-99 employees	0.09	0.29	0.11	0.32	0.11	0.31	0.10	0.29	0.12	0.33
100-199 employees	0.10	0.31	0.13	0.34	0.12	0.33	0.09	0.28	0.09	0.28
200-499 employees	0.14	0.35	0.16	0.36	0.11	0.31	0.14	0.35	0.14	0.34
$\geq$ 500 employees	0.22	0.41	0.21	0.41	0.22	0.41	0.20	0.40	0.33	0.47
Job level										
Level 1	0.21	0.41	0.02	0.14	0.01	0.11	0.00	0.02	0.00	0.00
Level 2	0.57	0.49	0.20	0.40	0.07	0.25	0.01	0.12	0.00	0.02
Level 3	0.21	0.41	0.76	0.43	0.75	0.43	0.45	0.50	0.09	0.29
Level 4	0.00	0.00	0.01	0.11	0.17	0.38	0.51	0.50	0.33	0.47
Level 5	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.15	0.44	0.50
Level 6-8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.14	0.35
Nature of occupation										
Simple techn. act.	0.39	0.49	0.33	0.47	0.42	0.49	0.39	0.49	0.23	0.42
Administrative	0.08	0.28	0.20	0.40	0.15	0.36	0.19	0.40	0.13	0.34
Automation	0.01	0.08	0.01	0.10	0.03	0.16	0.05	0.22	0.08	0.27
Commercial	0.06	0.24	0.08	0.27	0.11	0.31	0.12	0.33	0.10	0.31
Service providing	0.46	0.50	0.38	0.48	0.29	0.45	0.19	0.40	0.25	0.43
Creative	0.00	0.05	0.01	0.08	0.01	0.11	0.03	0.18	0.13	0.33
Management	0.00	0.00	0.00	0.00	0.00	0.03	0.01	0.11	0.09	0.29

Table 4: Descriptive statistics per segment

	Segment 1		Segment 2		Segment 3		Segment 4		Segment 5	
	est.	**s.e.**	**est.**	**s.e.**	**est.**	**s.e.**	**est.**	**s.e.**	**est.**	**s.e.**
Applications (in %)										
p_1	57.9	8.4	63.7	5.4	39.3	18.2	37.9	1.18	39.1	10.9
p_2	40.1	8.1	35.4	5.3	59.7	17.9	61.4	11.7	58.4	10.5
...	...	...	...	...	...	...	...	...	...	...
p_{30}	2.0	0.4	0.9	0.2	1.0	0.3	0.7	0.2	2.5	0.5
Other ML parameters										
$\widehat{y}$	17.68	0.89	26.96	1.50	26.15	2.04	30.70	1.72	39.54	2.72
$\underline{w}$	7.51	0.00	7.51	0.00	7.56	0.12	8.29	0.13	10.61	0.29
σ	0.012	0.005	0.007	0.003	0.031	0.012	0.026	0.013	0.074	0.016
Job offers (in %)										
q_0	56.5	1.2	52.0	0.6	51.2	0.6	49.9	0.4	57.1	1.2
q_1	37.1	1.7	41.7	1.0	40.6	1.2	41.6	0.7	35.5	1.4
q_2	4.4	0.3	5.4	0.4	7.2	0.6	7.7	0.5	4.9	0.4
...	...	...	...	...	...	...	...	...	...	...
q_4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1
q_5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1
q_6	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.2	0.1
q_7	0.2	0.1	0.0	0.0	0.1	0.1	0.0	0.0	0.3	0.1
q_8	0.2	0.1	0.1	0.0	0.1	0.1	0.1	0.0	0.4	0.1
q_9	0.3	0.1	0.1	0.0	0.1	0.1	0.1	0.0	0.4	0.1
q_{10}	0.3	0.1	0.1	0.0	0.2	0.1	0.1	0.0	0.3	0.0
q_{11}	0.3	0.0	0.1	0.0	0.1	0.0	0.1	0.0	0.3	0.0
q_{12}	0.2	0.0	0.1	0.0	0.1	0.0	0.1	0.0	0.2	0.0
q_{13}	0.2	0.0	0.1	0.0	0.1	0.0	0.1	0.0	0.1	0.0
q_{14}	0.1	0.0	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0
q_{15}	0.0	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Transition probabilities (in %)										
δ	3.9		2.6		2.0		1.7		1.1	
m_W	3.6		4.1		5.6		6.5		7.5	
m_F	62.2		68.6		69.7		71.5		61.3	
Other variables										
ϕ	2.824		2.308		2.692		2.607		3.313	
ψ (in %)	33.3		39.0		34.6		35.5		29.1	
b	3.60		4.02		4.27		4.83		6.26	
h	[0,7.08]		[0,6.74]		6.11		6.15		6.52	
k	96.66		274.06		295.20		395.23		541.31	
$E_{F_w}[w]$	8.87		9.91		10.46		11.86		15.11	
$E_{G_w}[w]$	9.01		10.06		10.67		12.09		15.65	
Statistics										
Obs.	1043		1179		1153		1070		1022	
LogL.	-1.352		-1.776		-1.940		-2.077		-2.641	
KS	3.25		1.26		1.59		1.53		1.29	

The presented fractions (in %) are conditional on searching at least once. The fraction of non-searchers (p_0) is displayed in table 3. The not reported fractions are equal to (or rounded down to) zero for all segments.

Table 5: Estimation results

	Segment				
	1	**2**	**3**	**4**	**5**
Data					
w_{avg}	9.12	10.21	10.97	12.35	16.33
$w_{1\%}$	7.12	7.50	7.40	7.14	8.31
$w_{2\%}$	7.50	7.65	7.51	7.98	9.06
$w_{5\%}$	7.51	7.90	7.92	8.60	10.22
$Mp1$	1.281	1.361	1.481	1.729	1.965
$Mp2$	1.216	1.336	1.461	1.548	1.803
$Mp5$	1.215	1.293	1.385	1.435	1.598
Model					
m_W	0.036	0.041	0.056	0.065	0.075
$m_{W\|U}$	0.435	0.480	0.488	0.501	0.429
δ	0.039	0.026	0.020	0.017	0.011
r	0.016	0.016	0.016	0.016	0.016
ρ	0.400	0.400	0.400	0.400	0.400
Mm_{NE}	1.572	1.435	1.310	1.258	1.190
Mm_U	1.073	1.051	1.044	1.039	1.037

Table 6: Mean-min ratio in data and model

	Market	Planner			Market	Planner	
		cstr.	**uncstr.**			**cstr.**	**uncstr.**
Segment 1				**Segment 2**			
p_0	91.7	76.3	70.5	p_0	91.4	71.0	65.8
p_1	4.8	23.7	29.5	p_1	5.5	29.0	34.2
p_2	3.3	0.0	0.0	p_2	3.1	0.0	0.0
...	...	...	...	...	...	...	...
p_{30}	0.2	0.0	0.0	p_{30}	0.1	0.0	0.0
N_F	0.510	0.782	0.917	N_F	0.639	0.858	0.954
Output	8.28	9.96	12.08	Output	11.77	14.52	17.77
Gain		20%	46%	Gain		23%	51%
Segment 3				**Segment 4**			
p_0	88.5	70.6	66.8	p_0	87.0	69.0	65.6
p_1	4.5	29.4	33.2	p_1	4.9	31.0	34.4
p_2	6.9	0.0	0.0	p_2	8.0	0.0	0.0
...	...	...	...	...	...	...	...
p_{30}	0.1	0.0	0.0	p_{30}	0.1	0.0	0.0
N_F	0.757	0.886	0.962	N_F	0.809	0.904	0.969
Output	13.56	15.83	18.54	Output	16.78	19.07	22.10
Gain		17%	37%	Gain		14%	32%
Segment 5							
p_0	82.6	52.5	50.0				
p_1	6.8	47.5	50.0				
p_2	10.1	0.0	0.0				
...	...	...	...				
p_{30}	0.4	0.0	0.0				
N_F	0.890	0.962	0.989				
Output	25.81	29.93	32.45				
Gain		16%	26%				

The fractions p_a are percentages. Omitted values are equal to (or rounded down to) zero.

Table 7: Comparison of the market equilibrium with the constrained (cstr) and unconstrained (uncstr) planner's solution

	Market	Planner		Market	Planner
Segment 1 ($h = 0$)			**Segment 1** ($h = 6.99$)		
p_0	91.7	80.7	p_0	91.7	74.4
p_1	4.8	19.3	p_1	4.8	25.6
p_2	3.3	0.0	p_2	3.3	0.0
...	...	...	...	...	...
p_{30}	0.2	0.0	p_{30}	0.2	0.0
N_F	0.510	0.781	N_F	0.510	0.787
Output	5.10	6.47	Output	8.74	10.55
Gain		27%	Gain		20%
Segment 2 ($h = 0$)			**Segment 2** ($h = 6.63$)		
p_0	91.4	079.2	p_0	91.4	69.3
p_1	5.5	20.8	p_1	5.5	30.7
p_2	3.1	0.0	p_2	3.1	0.0
...	...	...	...	...	...
p_{30}	0.1	0.0	p_{30}	0.1	0.0
N_F	0.639	0.833	N_F	0.639	0.863
Output	9.42	11.05	Output	11.97	14.89
Gain		17%	Gain		24%

The fractions p_a are percentages. Omitted values are equal to (or rounded down to) zero.

Table 8: Effect of bounds for household production on the planner's solution

	Market	Planner				Market	Planner		
		max. cost	lin. cost	min. cost			max. cost	lin. cost	min. cost
Segment 1					**Segment 2**				
p_0	91.7	91.7	76.3	0.0	p_0	91.4	91.4	71.0	0.0
p_1	4.8	8.3	23.7	100.0	p_1	5.5	8.6	29.0	100.0
p_2	3.3	0.0	0.0	0.0	p_2	3.1	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...
p_{30}	0.2	0.0	0.0	0.0	p_{30}	0.1	0.0	0.0	0.0
N_F	0.510	0.547	0.782	0.957	N_F	0.639	0.636	0.858	0.964
Output	8.28	9.18	9.96	11.37	Output	11.77	12.91	14.52	16.46
Gain		11%	20%	37%	Gain		10%	23%	40%
Segment 3					**Segment 4**				
p_0	88.5	88.5	70.6	0.0	p_0	87.0	87.0	69.0	0.0
p_1	4.5	11.5	29.4	100.0	p_1	4.9	13.0	31.0	100.0
p_2	6.9	0.0	0.0	0.0	p_2	8.0	0.0	0.0	0.0
...	...	...	...	...	...	...	...	...	...
p_{30}	0.1	0.0	0.0	0.0	p_{30}	0.1	0.0	0.0	0.0
N_F	0.757	0.747	0.886	0.972	N_F	0.809	0.793	0.904	0.975
Output	13.56	14.79	15.83	17.37	Output	16.78	18.02	19.07	20.67
Gain		9%	17%	28%	Gain		7%	14%	23%
Segment 5									
p_0	82.6	82.6	52.5	0.0					
p_1	6.8	17.4	47.5	100.0					
p_2	10.1	0.0	0.0	0.0					
...	...	...	...	...					
p_{30}	0.4	0.0	0.0	0.0					
N_F	0.890	0.896	0.962	0.985					
Output	25.81	28.76	29.93	31.04					
Gain		11%	16%	20%					

The fractions p_a are percentages. Omitted values are equal to (or rounded down to) zero.

Table 9: Effect of bounds on non-participant's search cost on the planner's solution

	Market		Planner	
		max. cost	lin. cost	min. cost
Segment 1				
h	6.11	7.51	6.11	6.23
p_0	91.7	91.7	76.3	0.0
p_1	4.8	4.8	23.7	100.0
p_2	3.3	3.3	0.0	0.0
...	...	...	...	...
p_{30}	0.2	0.0	0.0	0.0
N_F	0.510	0.559	0.782	0.957
Output	8.28	9.42	9.96	11.45
Gain		14%	20%	38%
Segment 2				
h	6.11	7.51	6.11	4.66
p_0	91.4	91.4	71.0	0.0
p_1	5.5	5.5	29.0	100.0
p_2	3.1	3.1	0.0	0.0
...	...	...	...	...
p_{30}	0.1	0.0	0.0	0.0
N_F	0.639	0.640	0.858	0.964
Output	11.77	12.82	14.52	16.48
Gain		9%	23%	40%
Segment 3				
h	6.11	7.56	6.11	2.30
p_0	88.5	88.6	70.6	0.0
p_1	4.5	4.5	29.4	100.0
p_2	6.9	6.9	0.0	0.0
...	...	...	...	...
p_{30}	0.1	0.0	0.0	0.0
N_F	0.757	0.767	0.886	0.972
Output	13.56	14.44	15.83	17.31
Gain		7%	17%	28%
Segment 4				
h	6.15	8.29	6.15	0.34
p_0	87.0	87.1	69.0	0.0
p_1	4.9	4.9	31.0	100.0
p_2	8.0	8.0	0.0	0.0
...	...	...	...	...
p_{30}	0.1	0.0	0.0	0.0
N_F	0.809	0.812	0.904	0.975
Output	16.78	17.57	19.07	20.58
Gain		5%	14%	23%
Segment 5				
h	10.61	7.51	6.52	-4.35
p_0	82.6	83.1	52.5	0.0
p_1	6.8	6.8	47.5	100.0
p_2	10.1	10.1	0.0	0.0
...	...	...	...	...
p_{30}	0.4	0.0	0.0	0.0
N_F	0.890	0.901	0.962	0.985
Output	25.81	28.39	29.93	30.94
Gain		10%	16%	20%

The fractions p_a are percentages. Omitted values are equal to (or rounded down to) zero.

Table 10: Bounds on the planner's solution for a general class of search cost functions

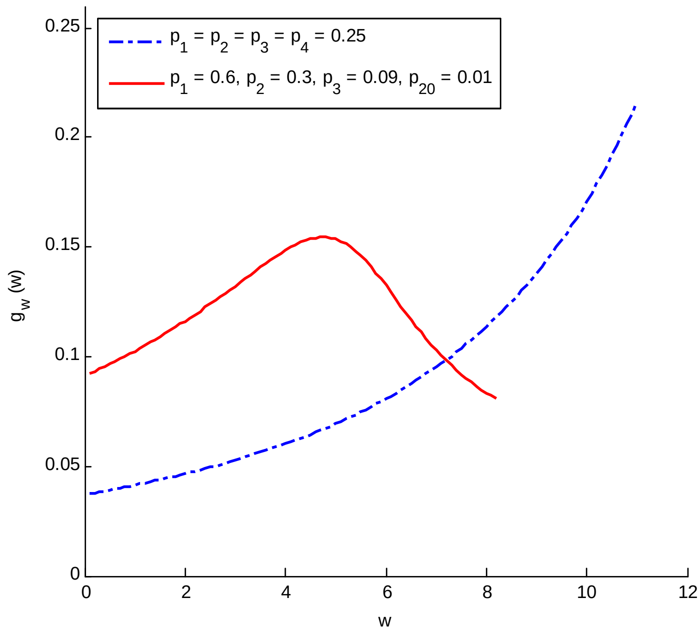

Figure 1: $g_w(w)$ for different values of the fractions $\{p_0, p_1, ..., p_S\}$

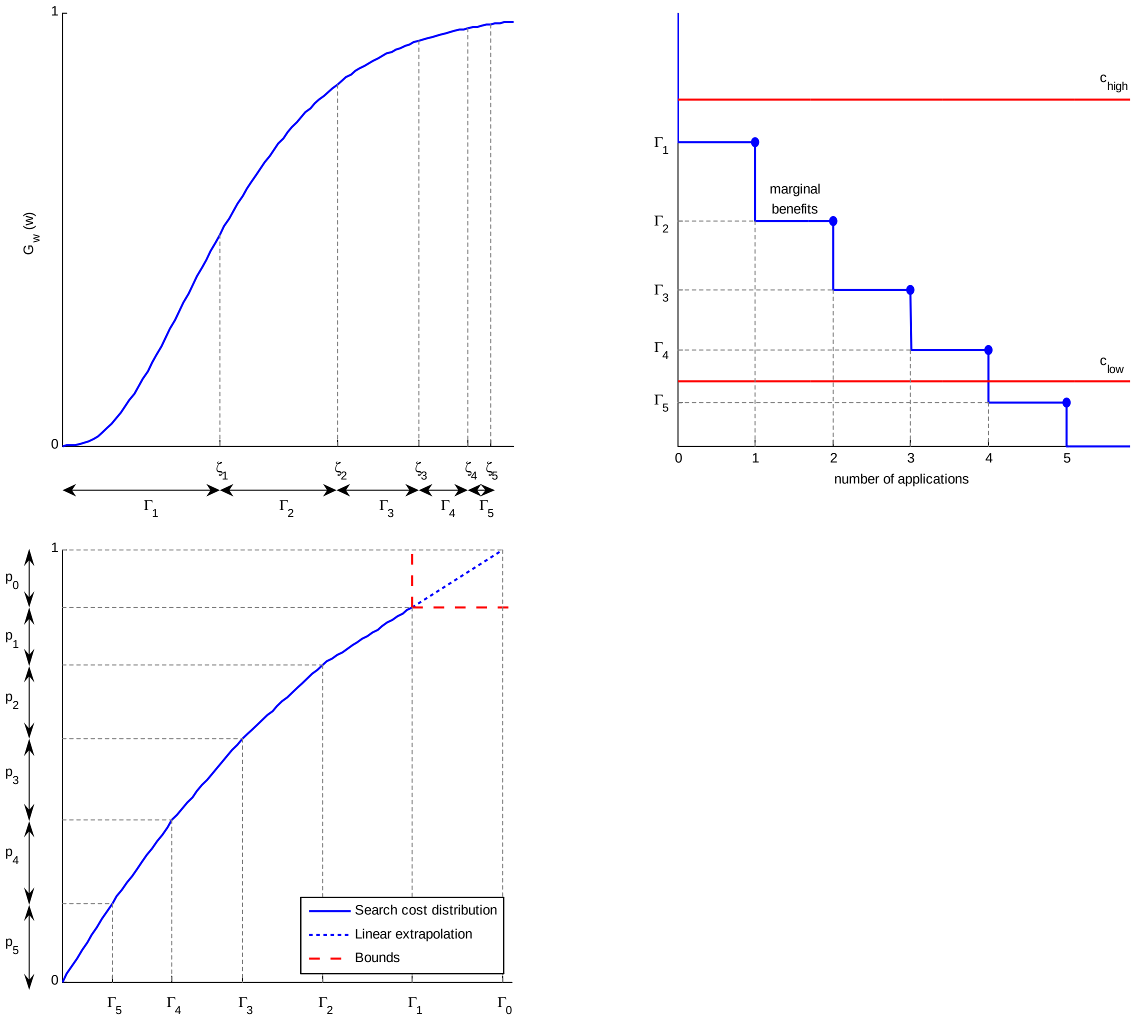

Figure 2: Relation between the wage distribution and the search cost distribution. For the ease of graphical exposition, the figure shows a special case in which $b = 0$, $w_R = 0$, $r + \delta = 1$ and $S = 5$. The figure is purely illustrative and no inferences about the actual distributions of wages and search costs should be made from it.

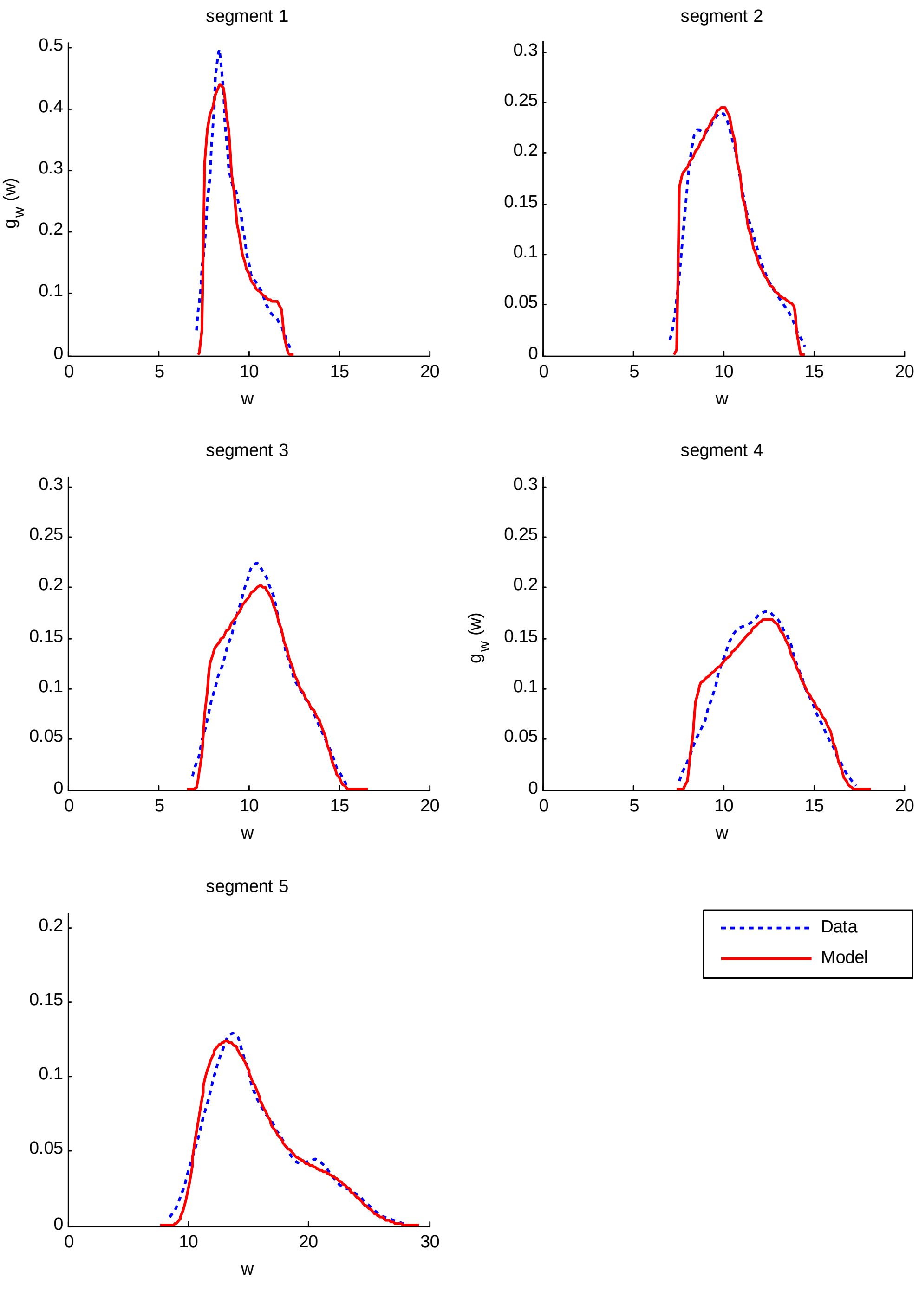

Figure 3: Estimated wage densities

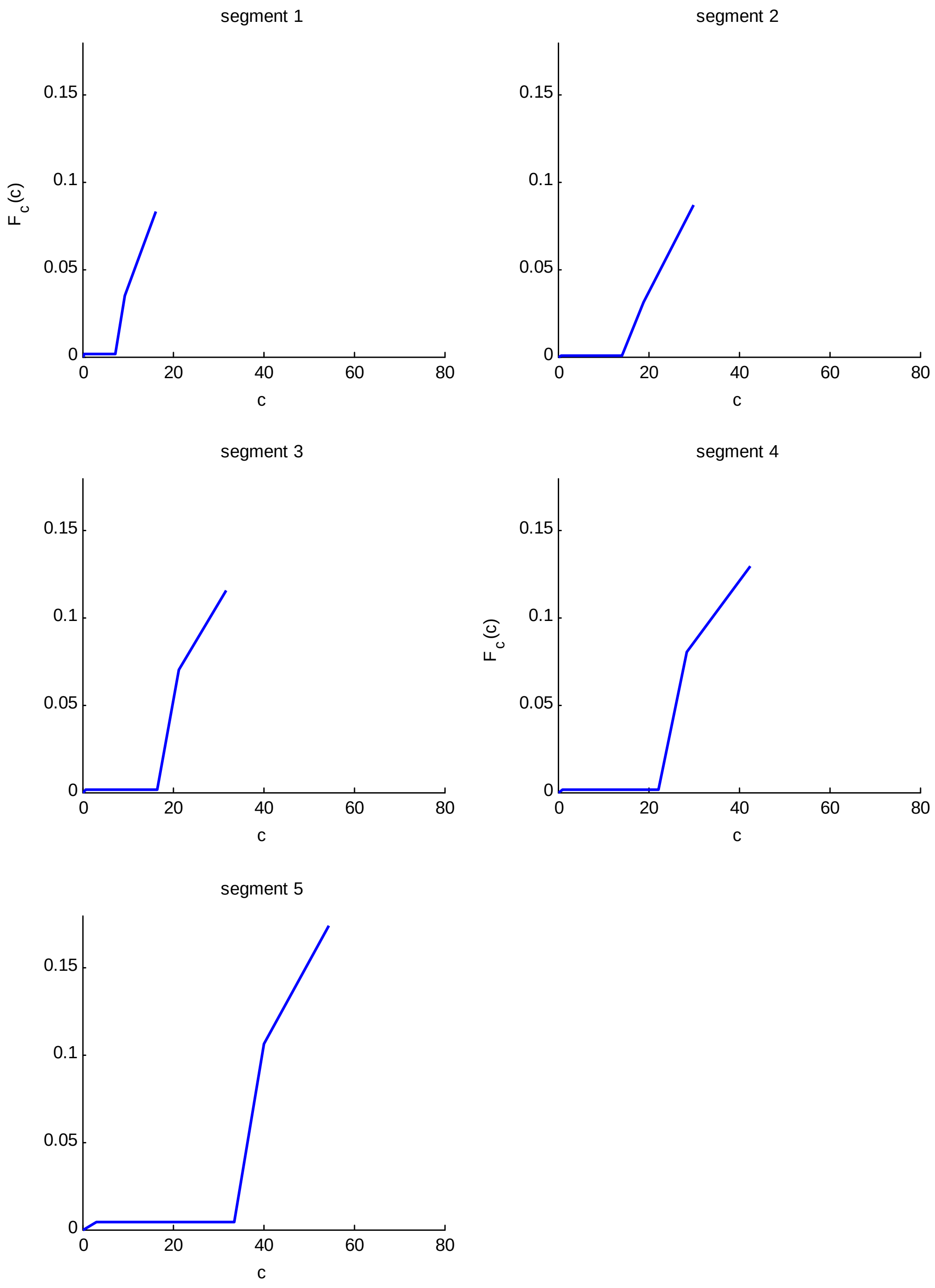

Figure 4: Estimated search cost distributions

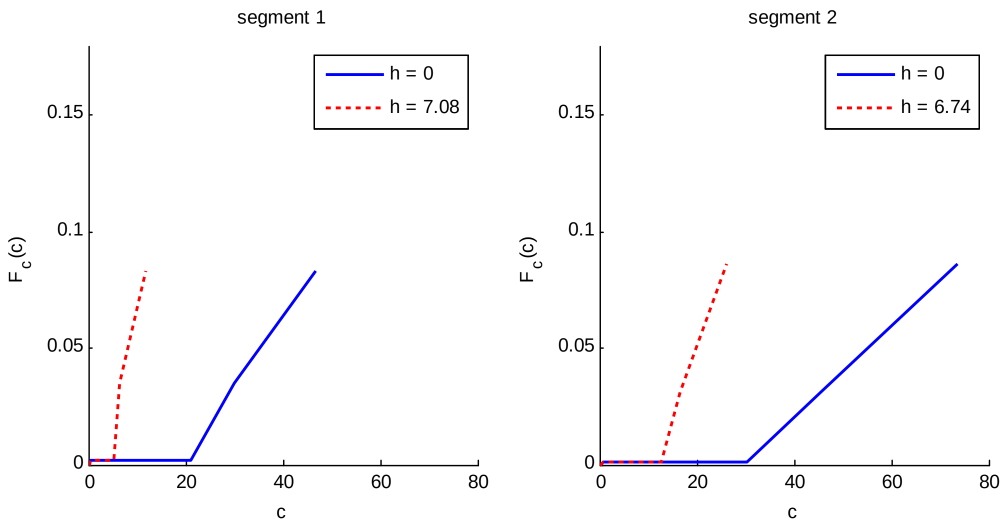

Figure 5: Estimated search cost distributions for segment 1 and 2 using bounds on the value of household production

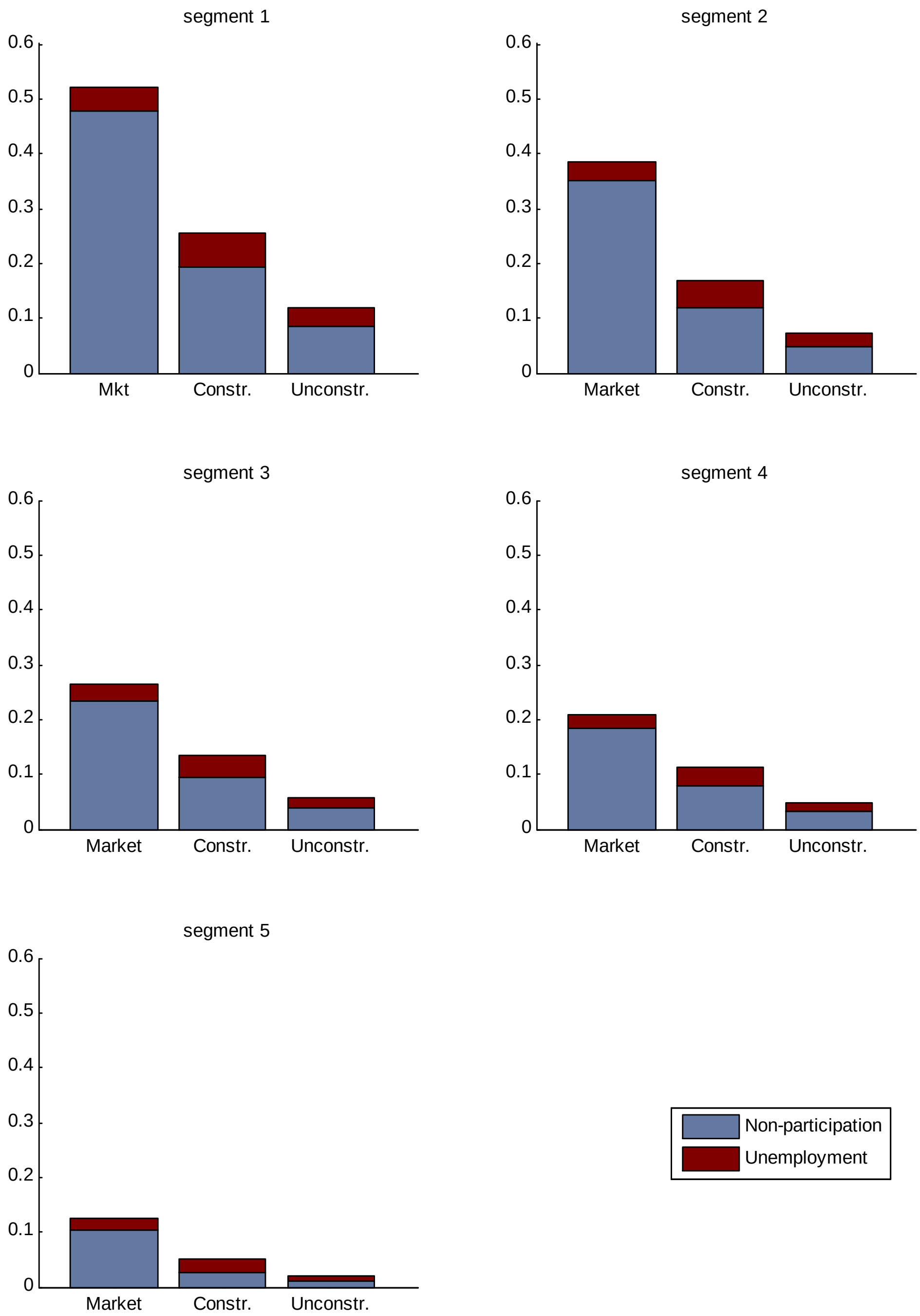

Figure 6: Non-participation and unemployment in the market equilibrium and the planner's solutions