

Job Search and Hyperbolic Discounting: Structural Estimation and Policy Evaluation*

M. Daniele Paserman
Department of Economics, Hebrew University
Jerusalem, 91905, Israel
e-mail: dpaserma@cc.huji.ac.il

June 2004

Abstract

This paper estimates the structural parameters of a job search model with hyperbolic discounting and endogenous search effort. It estimates quantitatively the degree of hyperbolic discounting, and assesses its implications for the impact of various policy interventions aimed at reducing unemployment. The model is estimated using data on unemployment spells and accepted wages from the National Longitudinal Survey of Youth (NLSY). The likelihood function explicitly incorporates all the restrictions implied by the optimal dynamic programming solution to the model. Both observed and unobserved heterogeneity are accounted for. The results point to a substantial degree of hyperbolic discounting, especially for low and medium wage workers. The structural estimates are also used to evaluate alternative policy interventions for the unemployed. Estimates based on a model with exponential discounting may lead to biased inference on the economic impact of policies.

JEL: C6, C73, D90, J64, J68.

Keywords: Job Search, Hyperbolic Discounting, Structural Estimation, Policy Evaluation.

*I thank Gary Chamberlain, Stefano DellaVigna, Eric Gould, David Laibson, Lawrence Katz, Robert Sauer and participants in seminars at the Bank of Israel, Bar Ilan University, Ben Gurion University, Ente Luigi Einaudi, Harvard University, the Hebrew University, University of Toulouse, the Econometrics in Tel Aviv 2001 conference, and the 2002 Econometric Society Winter Meetings for many helpful comments. Michael Hvoshniansky provided helpful research assistance. Financial support from the Maurice Falk Institute for Economic Research in Israel and from the Israeli Science Foundation are gratefully acknowledged. All errors are my own.

1. Introduction

Job search is an unpleasant activity with immediate costs and delayed benefits. The tension between long-run goals and short-run impulses may lead unemployed workers to postpone repeatedly tasks necessary to find a job. In standard economic models, agents are assumed to be time-consistent, so that a contrast between short-run and long-run preferences never arises. However, a growing literature has challenged the conventional view, and allows agents to be time inconsistent by modeling their discount function as *hyperbolic* (as opposed to the standard assumption of *exponential* discounting).¹ Agents with hyperbolic discount functions exhibit a high degree of discounting in the short run, but a relatively low degree of discounting in the long run. Therefore, hyperbolic agents are likely to delay tasks with immediate costs and delayed benefits, whereas they would choose to perform the same task if both costs and benefits were to occur in the future.

There are several reasons for introducing hyperbolic preferences in a job search context. First, hyperbolic preferences may affect the way we think of job search problems, and of policies addressing unemployment. Agents with hyperbolic and exponential preferences will give different weights to the various components of costs and benefits involved in the search process: hyperbolic agents will be particularly sensitive to the immediate and direct costs of search (writing the résumé, contacting employers, making unpleasant phone calls to distant relatives), while impatient exponential agents are more likely to be affected by long-run costs and benefits, such as those associated with waiting longer to obtain a better job. As a result, hyperbolic workers devote little effort to search activities, and possibly less than they wish. This prediction matches the anecdotal evidence of job counselors² to devote more time to search, as well as the puzzling quantitative evidence that unemployed individuals report searching on average only seven hours per week (Barron and Mellow, 1979). In addition, because of the different weights assigned to the different components of the job search process, policies that are targeted at one particular dimension of the job search process may be more effective for one type of worker than for the other.

Secondly, the welfare implications of unemployment policies may differ substantially depending on the type of preferences. In a model with conventional preferences, an agent may experience a long unemployment spell because of bad luck. With hyperbolic preferences, a long unem-

¹See Laibson (1997); O'Donoghue and Rabin (1999); Harris and Laibson (2001).

²Job hunting books routinely warn against searching too little: “If two weeks have gone by and you haven’t even started doing the inventory described in this chapter [...], don’t procrastinate any longer! Choose a helper for your job-hunt.” (Bolles, 2000. *What Color is Your Parachute?*, p. 87)

ployment spell could be due to “bad” choices, with the agent agreeing that his own choices are undesirable from a long-run perspective. Therefore, an intervention that brings the agent to choose actions more in line with his long-run preferences may actually be welfare improving, despite its imposition of restrictions, potentially even highly unpleasant ones, on the unemployed.

Thirdly, job search models may be a particularly fertile ground for testing the hyperbolic model versus the exponential model. This is so because the job search process involves two types of decisions involving intertemporal tradeoffs: the decision on how much effort to devote to search, and the decision on whether to accept a job offer or to continue searching. The timing of costs and benefits differs between these two tradeoffs, implying that different forms of discounting affect the two components of the job search process differently. Consequently, in job search models it may be possible to overcome the observational equivalence problem between hyperbolic and exponential models that plagues tests of hyperbolic discounting in other contexts.³

Despite the recent upsurge of interest in time-inconsistent preferences, direct estimates of the parameters of the discount function are relatively rare.⁴ Knowledge of these parameters has important implications not only for the specific environment studied in this paper (job search and unemployment spells), but also for the multitude of other applications in which hyperbolic preferences have been used, from retirement savings (Laibson, Repetto and Tobacman, 1998 and 2003; Diamond and Kőszegi, 2003) to the consumption of addictive goods (Gruber and Kőszegi, 2001).

In this paper, I provide one of the first structural estimates of the degree of hyperbolic discounting. I set up a model of job search with endogenous search effort and hyperbolic discounting, similar to the one used in DellaVigna and Paserman (2003, henceforth DV-P). I set up a likelihood function that explicitly incorporates all the restrictions implied by the optimal dynamic programming solution to the model. The model is estimated using data on unemployment spells and accepted wages from the National Longitudinal Survey of Youth (NLSY). I control for observed heterogeneity in the worker’s wages before the unemployment spell, in cognitive test scores, in marital status, and in race. Unobserved heterogeneity enters the model as a finite mixture distribution, whose parameters need to be estimated.

³For example, Laibson, Repetto and Tobacman (1998) find that the life-cycle consumption paths of hyperbolic and exponential agents are remarkably similar. Only the introduction of credit card borrowing in more recent work (Laibson, Repetto and Tobacman, 2003 and 2004) generates an observational non-equivalence between the hyperbolic and the exponential models. As in the present paper, the possibility to distinguish between the two models arises because of the contrast between the time horizon of utility flows deriving from two different decisions, namely the accumulation of liquid versus illiquid wealth.

⁴Exceptions include Angeletos et al. (2001), Fang and Silverman (2002), and Laibson, Repetto and Tobacman (2004).

A key issue that should be addressed is that of identification of the model parameters. It is important to differentiate between two aspects of the identification problem: in the literature on time discounting, the crucial question is whether it is possible to distinguish empirically between a model with hyperbolic agents from a model with exponential, but highly impatient, agents; in the job search literature, on the other hand, the central question is whether it is at all possible to identify the time discounting parameters from data on unemployment durations and accepted wages alone.

With respect to the first problem, I argue that in this setting it is possible to distinguish between hyperbolic and exponential agents, assuming that the discount function is identified. The distinguishing feature of my model is that different forms of impatience have contrasting effects on the outcomes of the job search process. Intuitively, impatience lowers the probability of receiving a job offer, but raises the probability of accepting the offer. In DV-P, we showed that for hyperbolic workers the first effect dominates, so that, *ceteris paribus*, workers with a higher degree of short-run impatience experience longer spells of unemployment. On the other hand, if individuals only differ in their exponential discount rate, the second effect in general dominates: more impatient exponential workers have shorter unemployment spells. The intuition behind the main result in DV-P helps to illustrate how the short-run and the long-run discounting parameters are separately identified with data on the length of unemployment spells and accepted wages. The search process is made up of two separate decisions: the search effort decision involves a trade-off between immediate costs and benefits that will be realized in the near future, once an offer is accepted. In the United States, this time span is typically short (the mean duration of unemployment in the United States is 20 weeks). Over this limited time horizon, short-run impatience matters the most. The reservation wage decision, on the other hand, essentially involves only a comparison between payoffs in the more distant future. To a first approximation, the degree of short-run discounting has no effect on the reservation wage.⁵ Then, assuming that all the other model parameters are known, one can back out the long-term discount factor from the estimated reservation wage. The short term discount factor is then identified from data on exit rates from unemployment.

With respect to the second issue, it is useful to return to Flinn and Heckman's (1982) discussion of the identification problem in optimal job search models. Flinn and Heckman analyze a stationary job search model with exogenous offer arrival rates and show that data on unemployment durations and accepted wages is not sufficient to identify the parameters of the model. The basic problem is that one can observe only the part of the wage distribution above the

⁵The statement is exactly true if agents are naive and there is a one period delay between the time a job offer is accepted and the time of the first wage payment. See DV-P, Proposition 1(c).

reservation wage, making it impossible to distinguish between offer arrival rates and acceptance probabilities. Imposing the optimality condition for the reservation wage does not resolve the problem, despite the fact that the condition depends only on the truncated part of the wage offer distribution. This is because two other parameters appear in the optimality condition: the utility of being unemployed and the offer arrival rate. Even if one sets the utility of unemployment equal to UI benefits, the offer arrival rate and the exponential discount factor are not separately identified. To overcome these problems, I need to modify the Flinn-Heckman assumptions and add some structure to the job search model. This is done in three ways. First, in line with previous research (Wolpin, 1997, van den Berg 1990) I specify a particular functional form for the wage offer distribution. This allows me to parametrically identify the reservation wage and the acceptance probability from data on accepted wages alone. Second, the utility of being unemployed is not constant over time, and depends on the time profile of unemployment insurance benefits, making the model non-stationary. Third, I endogenize the offer arrival rate, and I choose a convenient and natural functional form for the cost of search function. This function is then calibrated using data on actual hours spent searching and on the effectiveness of search. In the empirical section, I explore the sensitivity of the results to the particular functional forms adopted for the wage distribution and for the time profile of the utility of being unemployed.

In my preferred specification, the results of the structural estimation point to a substantial degree of present bias for low and medium wage workers, and only a moderate degree of short run impatience for high wage workers. The results are fairly robust to different assumptions about the heterogeneity distribution or the time-profile of the value of unemployment, while they are somewhat more sensitive to the assumed functional form for the offered wage distribution. In all specifications, the point estimate for the hyperbolic discounting parameter is quite far away from the benchmark value of 1, which is equivalent to exponential discounting. The maximum likelihood estimates are also used to evaluate alternative policy interventions for the unemployed. The policies analyzed are a cut in unemployment benefits, a job search assistance program, monitoring search effort, monitoring the job acceptance strategy, and a re-employment bonus. I find that ignoring hyperbolic preferences may lead one to incorrect inferences on the effects of these interventions.

2. The Model

2.1. Hyperbolic discounting

Over the years, psychologists have collected a substantial body of evidence on individual time preferences (for a review, see Ainslie, 1992). Experiments show that agents are extremely impatient if the rewards are to be obtained in the near future, but relatively patient when choosing between rewards to be accrued in the distant future. This form of discounting implies that agents prefer a larger, later reward over a smaller, earlier one as long as the rewards are sufficiently distant in time; however, as both rewards get closer in time, the agent may choose the smaller, earlier reward. In an experiment with monetary rewards an overwhelming majority of subjects exhibit such reversal of preferences (Kirby and Herrnstein, 1995).

Hyperbolic discount functions, introduced by Phelps and Pollak (1968) and first studied in the context of intertemporal one-person decisions by Laibson (1997), provide a convenient representation of the above findings: for a decision maker at *any* time s , the discount function is equal to 1 for $t = s$ and to $\beta\delta^{t-s}$ for $t = s + 1, s + 2, \dots$, with $\beta < 1$. The implied discount factor from period s to the next period is $\beta\delta$, while the discount factor between any two periods in the future is simply δ . However, when period $s + 1$ comes along, the agent faces the same type of preferences: the discount factor between that period and the next is $\beta\delta$. This matches the main feature of the experimental evidence — high discounting at short delays, low discounting at long delays.

These preferences are dynamically inconsistent. To illustrate this, consider a plan of actions q_t that yield instantaneous utility $u(q_t)$, for $t = 0, 1, 2, \dots, T$. From today's perspective, the plan from period s onwards yields utility

$$V^s(q_s, \dots, q_T) = \beta \sum_{t=s}^T \delta^t u(q_t). \quad (2.1)$$

However, from the perspective of the decision maker at time s , the same plan yields utility

$$\tilde{V}^s(q_s, \dots, q_T) = u(q_s) + \beta \sum_{t=s+1}^T \delta^{t-s} u(q_t). \quad (2.2)$$

There is therefore a conflict between the preferences of a given individual at different points in time. Note that this conflict disappears in the special case of $\beta = 1$. In this case, we are back to the time-consistent *exponential* model with discount function δ^t .

We can interpret β as the parameter of short-run patience and δ as the parameter of long-run patience. The degree of short-run patience β is crucial for this theory; any β smaller than one is

sufficient to generate some degree of procrastination in activities with salient costs and delayed benefits, such as searching for a job (O’Donoghue and Rabin, 1999). The implications of this form of time inconsistency in the context of job search are easy to see. A worker with a very high degree of short run impatience may wish to postpone her job search activities to a later period, when, from today’s perspective, the discount rate is relatively low; however, when the later period comes along, the worker once more faces a high degree of short-run impatience, and will choose to postpone her activity again.

For the purposes of this paper, we will restrict attention to the case of a *sophisticated* hyperbolic worker. A sophisticated agent is aware of her time inconsistency problem: she knows that at time s in the future, she will choose an action q_s^* that maximizes the utility function in (2.2). The optimal choice at s depends on the set of actions chosen from today until $s - 1$. Therefore in the present she chooses q_0 that solves:

$$\max_{q_0} u(q_0) + \beta \sum_{t=1}^T \delta^t u[q_t^*(q_0)].$$

The sophisticated agent knows that if she postpones a task until tomorrow, she may wish to postpone it again. She will therefore try to find ways to overcome her procrastination problems. In the context of job search, the sophisticated worker will, assuming that no commitment devices are available, exert higher effort in the present, knowing that in the future she will search less than what is optimal from today’s perspective.⁶

2.2. Setting

The model is a variant of the prototypical job search model (Lippman and McCall, 1976), augmented to include endogenous search effort and hyperbolic discounting. The model is set in discrete time; it is convenient to think of a week as the time unit. Consider an infinitely lived worker who is unemployed at time t . In each period of unemployment, the worker takes two decisions: first, he chooses the amount of search effort; second, conditional on receiving a job offer, he decides whether to accept it or not.

In every period, the worker receives instantaneous utility b_t . I model b_t as the sum of two components: the first component, b_t^{UI} is the monetary value of Unemployment Insurance (UI) benefits. This may vary during the course of the unemployment spell. The second component,

⁶A *naive* worker believes incorrectly that her future preferences will be exponential, and that his procrastination problems are only temporary. In DV-P, we found that the qualitative and quantitative behavior of sophisticated and naive workers in a job search model is similar.

b_0 , represents the psychic value of being unemployed: it can be either positive (utility of leisure) or negative (stigma, low self-esteem, etc.). I assume that b_0 is invariant over time.

Search effort s_t is parameterized as the probability of obtaining a job offer;⁷ therefore, $s_t \in [0, 1]$. In every period the agent incurs a cost of search $c(s_t)$, an increasing and strictly convex function of s_t . The assumption of increasing marginal cost comes naturally. The marginal cost of effort is low when the agent is searching little: a meeting with a friend or a glance at the newspaper raises the probability of finding a job at minimal additional expense. On the other hand, increasing the probability of receiving an offer up to almost certainty is likely to have prohibitive costs. In order to simplify the characterization of the solution, I also assume no fixed costs of search: a worker may be informed about an outstanding job offer at no cost.

Upon receiving a job offer, the worker must decide whether to accept it or not. The job offer is characterized by a wage w , which is a realization of a random variable W with cumulative distribution function F . If the worker accepts the offer, he becomes employed and receives wage w starting from next period. If no offer is received or the offer is declined, the worker searches again in period $t + 1$. We assume F to be known to the worker, constant over time and independent of search effort. In other words, search effort determines how often the individual samples out of F , not the distribution being sampled.

Finally, we allow for the possibility of layoff. At the end of each period of employment, the worker is laid off with known probability $q \in [0, 1]$,⁸ in which case he becomes unemployed starting from next period. With probability $1 - q$, the worker continues to be employed at wage w .

Note that, while the cost function and the wage distribution function are assumed to be invariant over time, we allow the utility associated with unemployment to be a function of the length of the unemployment spell. This is meant to reflect the common feature of the Unemployment Insurance system, where benefits are paid only for a limited amount of time. The model is therefore non-stationary: the optimal amount of search effort and the acceptance/rejection strategy will depend on the current length of the unemployment spell. Finally, we should note that the model is set in partial equilibrium and abstracts from any potential response of firms

⁷Suppose that the worker chooses search effort z , and that a job offer is generated with probability $s = p(z)$. Under the assumption that $p(\cdot)$ is a strictly increasing function of search, there is a one-to-one correspondence between s and z . Without loss of generality, we can then specify the worker's problem more conveniently in terms of the probability of receiving an offer s .

⁸The assumption of a fixed separation probability q is somewhat unrealistic, given the well documented declining shape of the hazard function for employment spells. However, an attempt to estimate a model with the separation probability being a declining step function did not yield significantly different results.

to the presence of job seekers with hyperbolic preferences.

2.3. Solution

For any period t , we can write down the maximization problem of the unemployed worker for given continuation payoff V_{t+1}^U when unemployed and $V_{t+1}^E(w)$ when employed at wage w . The worker chooses search effort s_t and the wage acceptance policy to solve

$$\max_{s_t \in [0,1]} b_t - c(s_t) + \beta\delta [s_t E_F \{ \max(V_{t+1}^E(w), V_{t+1}^U) \} + (1 - s_t) V_{t+1}^U], \quad (2.3)$$

where the expectation is taken with respect to the distribution of wage offers F . Expression (2.3) is easily interpretable: the worker in period t receives benefits b_t and pays the cost of search $c(s_t)$. The continuation payoffs are discounted by the factor $\beta\delta$, where β is the additional term due to hyperbolic discounting (for the exponential worker, $\beta = 1$). With probability s_t the worker receives a wage offer w that he can then accept — thus obtaining, starting from next period, the continuation payoff from employment $V_{t+1}^E(w)$ — or reject, in which case he gains next period the continuation payoff from unemployment, V_{t+1}^U . With probability $1 - s_t$, the worker does not find a job and therefore receives V_{t+1}^U . The continuation payoff from employment at wage w , from the perspective of the decision maker in period t , is

$$V^E(w) = w + \delta [qV_{new}^U + (1 - q)V^E(w)].^9$$

The worker obtains wage w in period $t + 1$; then, with probability $1 - q$ he maintains his job and continues to receive wage w ; with probability q he is laid off, in which case he enters a new spell of unemployment whose value is V_{new}^U . The exact specification of V_{new}^U is deferred until later. Note that, from the perspective of self t , payoffs from period $t + 1$ onwards are discounted exponentially.

To make the model operational, we now assume that, beginning in period $T + 1$, the environment is stationary. The only environmental variable that we allow to vary over time is b_t^{UI} . Once UI benefits are exhausted, the environment becomes stationary: in every period the worker chooses the same intensity of search effort and the same acceptance strategy.¹⁰

The equilibrium solution in the stationary model is characterized by a reservation wage policy. Let w_{T+1}^* and s_{T+1} represent, respectively, the reservation wage and the optimal level of search

⁹Note that the continuation payoff from employment does not depend on the time one becomes employed. Hence we can omit the time subscript.

¹⁰In DV-P we give a detailed account of the stationary solution: this includes a proof of existence and uniqueness of the stationary Markov perfect equilibrium in the intrapersonal game played by the hyperbolic individual's different selves.

effort in the stationary equilibrium. The reservation wage is:

$$w_{T+1}^* = [1 - \delta(1 - q)]V_{T+1}^U - \delta qV_{new}^U. \quad (2.4)$$

The stationary value of being unemployed, V_{T+1}^U , and the optimal level of search effort, s_{T+1} , can then be found by solving the following system of nonlinear equations:

$$(1 - \delta)V_{T+1}^U = b_{T+1} - c(s_{T+1}) + \frac{\delta s_{T+1}}{1 - \delta(1 - q)}Q(w_{T+1}^*); \quad (2.5)$$

$$c'(s_{T+1}) = \frac{\beta\delta}{1 - \delta(1 - q)}Q(w_{T+1}^*), \quad (2.6)$$

where $Q(x) \equiv \int_x^\infty (u - x) dF(u)$. Having obtained the solution for the stationary problem, it is straightforward to solve the entire model by backwards induction. The dynamic equations characterizing the solution in periods $t = 1, \dots, T$ are :

$$w_t^* = [1 - \delta(1 - q)]V_{t+1}^U - \delta qV_{new}^U; \quad (2.7)$$

$$V_t^U - \delta V_{t+1}^U = b_t - c(s_t) + \frac{\delta s_t}{1 - \delta(1 - q)}Q(w_t^*); \quad (2.8)$$

$$c'(s_t) = \frac{\beta\delta}{1 - \delta(1 - q)}Q(w_t^*). \quad (2.9)$$

To obtain a solution to this model, we need to specify V_{new}^U , the value of becoming unemployed after having been laid off. Given the features of the Unemployment Insurance system, V_{new}^U should depend on the accepted wage and on the duration of the employment spell. However, this would introduce considerable difficulties in calculating the solution. Instead, I make the simplifying assumption that $V_{new}^U = V_{T+1}^U$. This implies that a worker is no longer eligible to take UI benefits after his first spell of unemployment. While this assumption is unrealistic for most workers, it is plausible to assume that any bias introduced will not be large.¹¹

In DV-P, we highlighted the relationship between the outcomes of the job search process and the impatience parameters β and δ . Impatience has two contrasting effects on the job search process. Impatient individuals dislike all sorts of investment activities, and therefore exert little

¹¹I also experimented with the assumption that $V_{new}^U = V_1^U$: in other words, once the worker is laid off he starts a new unemployment spell with all parameters taking on the same values as in the current spell. In this case a state variable is added to the dynamic programming model (V_1^U), and the solution involves finding a fixed point to the system

$$V_1^U = h(V_1^U)$$

where the function h is defined recursively by equations (2.5)-(2.9) with $V_{new}^U = V_1^U$. A fixed point did not always exist for all parameter values; where it was possible to obtain a solution, it did not differ substantially from the one obtained assuming $V_{new}^U = V_{T+1}^U$.

search effort. On the other hand, once they have a job offer in hand, they prefer to accept it right away rather than wait an additional period for an even better offer. In other words, impatience lowers the probability of receiving, but raises the probability of accepting a job offer. Both impatience parameters in fact operate in the same direction on the intensity of search effort and on the reservation wage. However, the magnitude of the effects differs sharply: the short run discounting parameter β operates mainly on the intensity of search effort, while the long-run discounting parameter δ influences primarily the reservation wage. As a result, higher short-run impatience is associated with longer unemployment spells (the search effect dominates). In DV-P we show that this result requires only a very mild assumption on the wage distribution, which is implied by all log-concave distributions and is comfortably satisfied by other distributions commonly used in the search literature.¹² By contrast, higher long-run impatience is in general associated with shorter unemployment spells (the reservation wage effect dominates). This result is always true for values of δ close enough to one. Extensive simulations in DV-P show that for plausible parameterizations of the wage distribution and of the layoff probability the exit rate is decreasing in δ for all values of the yearly discount factor greater than 0.8.¹³ In other words, over the range of values that are considered plausible in the consumption and finance literature, the exit rate is decreasing in δ .

For intuition on this result, consider the two separate decisions making up the search process. The decision on intensity of search effort involves a trade-off between the present costs of searching and benefits that will start to materialize in the near future, once an offer is accepted. This time span is relatively short: in the United States most spells end in less than 30 weeks. Over this limited time horizon, short-run impatience matters the most. On the other hand, the reservation wage decision involves a comparison of long-term consequences, once an offer is received: the worker chooses whether to accept the wage or wait for an even better offer. Since most employment spells last for more than a year, the worker is making a choice for the long run. Therefore, the reservation wage is mainly affected by the degree of long-run discounting.¹⁴

The distinct role played by β and δ in the job search process is important for identification

¹²See DV-P, Proposition 2.

¹³See DV-P, Table 1.

¹⁴The reader may be concerned that the results depend crucially on the timing of events in the job search process. In particular, if a worker could start receiving the wage immediately upon acceptance of a job offer, then the first wage receipt would also be discounted at rate $\beta\delta$, thereby invalidating the argument on the irrelevance of short-run discounting for the reservation wage decision. This is not the case. In DV-P (Appendix D), we show that a continuous-time version of our model, with wages being paid off immediately, yields exactly the same first order conditions as the present model. Hence, the exact timing of wage receipts is immaterial for the result on the different effects of short and long-run discounting on search effort and reservation wages.

issues. Despite the fact that both β and δ represent the degree of preferences for the present versus the future, it is possible to identify them separately. The above discussion suggests that, to a first approximation, data on reservation wages alone may be sufficient to identify δ , and data on search intensity may be sufficient to identify β .

3. Estimation

I now turn to the structural estimation of the job search model, using data on the duration of unemployment spells and on re-employment wages, as in the classic works of Wolpin (1987), and van den Berg (1990). My work is new in that it allows for hyperbolic preferences, and centers its attention on the discounting parameters. I also attempt to estimate the parameters of the cost of search function, a feature absent from models with an exogenous offer arrival rate.¹⁵ As in van den Berg, I introduce non-stationarity by allowing for UI benefits to run out after a limited amount of time. I estimate the model separately for three groups of workers, classified by their wage prior to the unemployment spell. In this way, I am able to control for observed heterogeneity along one very important dimension. In addition, in the basic specification I will allow variation in parameters by AFQT scores and marital status within each wage group.¹⁶ Finally, I incorporate unobserved heterogeneity in the form of a finite mixture distribution for the model parameters. This will help to capture the negative duration dependence in unemployment spells typically observed in this type of data. In section 4.6 I consider alternative forms in which observed heterogeneity may affect the parameters of the model.

3.1. The Likelihood Function

The data consists of observations on the length of unemployment spells in weeks, T_i , on re-employment wages w_i , and on a set of individual characteristics X_i for a sample of individuals $i = 1, 2, \dots, N$. I denote by $h_t(\theta, X_i) \equiv s_t(\theta, X_i) [1 - F_\theta(w_t^*(\theta, X_i))]$ the exit rate from unemployment in week t as a function of the model parameters θ and of individual characteristics X_i . The exit rate in week t is simply the product of the probability of receiving an offer in week t , $s_t(\theta, X_i)$, and the probability that this offer exceeds the reservation wage in week t , $w_t^*(\theta, X_i)$. The values of $s_t(\theta, X_i)$ and $w_t^*(\theta, X_i)$ are obtained from the solution of the dynamic programming problem (2.3). Let C_i be a dummy variable indicating whether the observed unemployment spell

¹⁵Bloemen (2004) estimates a structural search model with endogenous search intensity using direct information on different indicators of the intensity of search.

¹⁶AFQT and marital status were chosen as the explanatory variables since it appeared that the most significant variation in exit rates and re-employment wages occurred along these lines.

is complete, and let E_i be a dummy indicating whether the re-employment wage is observed. Assuming that the measured wage w_i^{obs} is a noisy measure of the true wage \tilde{w}_i , the log-likelihood function, in the absence of unobserved heterogeneity, is:

$$\begin{aligned}
L(\theta) = & \sum_{i=1}^N \sum_{\tau=1}^{T_i-1} \log [1 - h_{\tau}(\theta, X_i)] + \\
& \sum_{i:C_i=1} \log [h_{T_i}(\theta, X_i)] + \\
& \sum_{i:C_i=0} \log [1 - h_{T_i}(\theta, X_i)] + \\
& \sum_{i:C_i=1, E_i=1} \log \frac{P(\tilde{w}_i \geq w_{T_i}^*(\theta, X_i) \mid w_i^{obs}) f(w_i^{obs})}{P[\tilde{w}_i \geq w_{T_i}^*(\theta, X_i)]}.
\end{aligned}$$

The first term represents the probability that individual i does not exit unemployment in any of the periods from $t = 1$ to $T_i - 1$; the second term represents the likelihood contribution of individuals who find employment at period T_i ; the third term is the contribution of individuals who are censored at T_i ; the last term is obtained from the probability density of observing a re-employment wage w_i^{obs} , conditional on the true wage being greater than the reservation wage $w_{T_i}^*$. See Appendix A for a full derivation and for the extension to incorporate unobserved heterogeneity.

3.2. Functional form specification

In order to solve the model and to identify its parameters, one needs to specify functional forms for the cost function and the wage distribution function, and set values for those parameters that are not directly estimated. The choice of functional forms is dictated by empirical plausibility and by computational convenience.

The value of time when unemployed. I model the value of time when unemployed as the sum of the monetary value of UI benefits and a time invariant component b_0 .

$$\begin{aligned}
b_t &= b_t^{UI} + b_0 \\
b_t^{UI} &= \begin{cases} b^{UI} & \text{if } t \leq T \\ 0 & \text{if } t > T \end{cases}
\end{aligned}$$

The second component can be thought of as either positive (utility of leisure) or negative (stigma associated with unemployment). It enters as an unknown parameter in the likelihood function.

Consistent with the UI system in most states, I assume in the benchmark specification that benefits are received for $T = 26$ weeks. The monetary value of unemployment benefits b^{UI} is fixed at the average value of actual benefits observed in the population of interest. In alternative specifications, I also estimate a stationary version of the model, with $T = 0$.

The wage distribution. I estimate the model assuming that the actual wage \tilde{w}_i is drawn from either a log-normal or a normal distribution. This allows me to assess the sensitivity of my estimates to different functional form assumptions about the wage distribution. In what follows, I illustrate the estimation procedure for the log-normal case with a multiplicative error term.¹⁷ I assume that the parameters of the log-normal distribution are μ_i and σ_i . The moments of the log wage distribution are assumed to depend on observed characteristics X_i . The vector X_i consists of six indicator variables indicating the possible combinations of marital status and three AFQT dummies (high, medium, and low). Observed wages w_i^{obs} are equal to actual wages times a multiplicative error term with a log-normal distribution, so that

$$\log w_i^{obs} = \log \tilde{w}_i + u_i,$$

and

$$\begin{pmatrix} \log \tilde{w}_i \\ u_i \end{pmatrix} \sim N \left[\begin{pmatrix} \mu_i \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_i^2 & 0 \\ 0 & \sigma_u^2 \end{pmatrix} \right].$$

This specification implies that log wages are measured with a white noise measurement error: this assumption is plausible if one considers that many of the other characteristics of a job, pecuniary and not, are not observed in our data. The joint normality of actual log wages and the error term implies joint normality between actual log wages and observed log wages. Then, using the properties of the multivariate normal distribution, we can derive

$$\begin{aligned} f(w_i^{obs} | \tilde{w}_i > w_{T_i}^*(\theta)) &= \frac{P(\tilde{w}_i \geq w_{T_i}^*(\theta) | w_i^{obs}) f(w_i^{obs})}{P(\tilde{w}_i \geq w_{T_i}^*(\theta))} \\ &= \frac{\left[1 - \Phi \left(\frac{\log w_{T_i}^* - a_{1i} - a_{2i} \log w_i^{obs}}{v_i} \right) \right] \left(\frac{1}{w_i^{obs} \sqrt{\sigma_i^2 + \sigma_u^2}} \phi \left(\frac{\log(w_i^{obs}) - \mu_i}{\sqrt{\sigma_i^2 + \sigma_u^2}} \right) \right)}{1 - \Phi \left(\frac{\log w_{T_i}^* - \mu_i}{\sigma_i} \right)}, \end{aligned}$$

with $a_{2i} = \frac{\sigma_i^2}{\sigma_i^2 + \sigma_u^2}$, $a_{1i} = \mu_i(1 - a_{2i})$, and $v_i^2 = \sigma_i^2 - a_{2i}^2(\sigma_i^2 + \sigma_u^2)$, and $\Phi(\cdot)$ and $\phi(\cdot)$ represent the standard normal cumulative distribution function and probability density function respectively. The parameters μ_i , σ_i , and σ_u all need to be estimated from the data.

¹⁷Adapting this procedure to the normal distribution with an additive error term is straightforward.

The cost function. A convenient and natural functional form that captures the main features of the cost function is

$$c_i(s) = k_i s^{1+\eta}, \quad \eta > 0.$$

The parameter η represents the degree of convexity of the cost function. It is also equal to the (constant) elasticity of marginal cost of search with respect to effort: the higher the value of η , the higher the proportional increase in the marginal cost of search for a given percentage increase in search effort. The parameter k_i represents the scale of the cost function: it tells us how costly it would be to obtain a job with probability one. I allow this parameter to depend on observed characteristics X_i .

The parameter η is estimated using data from the 1981 wave of the NLSY that provides detailed information on the search activities of employed and unemployed youth.¹⁸ I consider an extension of the model described in the text in which workers have already chosen the optimal probability of receiving an offer s , and must now decide how to optimally allocate their time between alternative search methods. Assume that there are K different methods that can be used to generate offers. Once the worker has chosen s , he must decide how intensively to use each method of search. Assume also that the effort exerted in this period has no bearing on the probability of receiving an offer in any of the latter periods, and that each search method generates offers from the same wage distribution. These assumptions make the model effectively static, and allow us to abstract from the issues of time discounting.

The worker's problem becomes one of optimally choosing intensity of search for each method so as to minimize search costs, subject to the constraint that the probability of receiving an offer be at least s . Formally, the problem is

$$\begin{aligned} & \min_{X_1, \dots, X_K} \sum_{j=1}^K c_j X_j \\ \text{s.t.} \quad & P(X_1, \dots, X_K) \geq s \\ & X_j \geq 0, \quad j = 1, \dots, K, \end{aligned}$$

where X_j is the amount of hours devoted to search using method j , and c_j is the cost of one hour of search using method j . In Appendix B, I show that an appropriate choice of functional form for $P(X_1, X_2, \dots, X_K)$ yields a closed form solution for the optimal hours of search for each method, and these in turn generate a cost of search function of the form

$$\tilde{c}(s) = A(1-s)^{-\varepsilon} - B, \tag{3.1}$$

¹⁸This is the same data used in Holzer (1988).

where A, B , and ε are functions of productivity and the cost of each search method. The productivity of search parameters are estimated by maximum likelihood, using data on whether each search method resulted in a job offer. Then one can use the first order conditions to back out the cost of search parameters c_j , up to a constant of proportionality (details in Appendix B). Given α and c , one can evaluate the cost function $\tilde{c}(s)$ at various levels of search intensity s . These estimated values are then used to estimate the parameters of the constant marginal cost elasticity cost function, $c(s) = ks^{1+\eta}$, used in the dynamic programming model. The estimation is performed by running a weighted least squares regression of $\log \tilde{c}(s)$ on $\log s$, where the weights are proportional to a normal density centered at the estimated mean probability of receiving a job offer.¹⁹ The results of this regression are

$$\begin{aligned} \log \tilde{c}(s) &= 4.412 + 1.408 \log s \\ &\quad (0.003) \quad (0.0011) \end{aligned}$$

The estimated constant marginal cost elasticity η is then equal to the coefficient on $\log s$ minus one. This gives rise to the benchmark value of $\eta = 0.4$.

One can then use this value and information on weekly offer arrival probabilities to gauge the magnitude of k . For example, if one believes that the weekly cost of search is roughly of the same order of magnitude as the weekly wage, say \$250,²⁰ then a weekly offer probability of 0.1 and the estimated value of η imply a value for k of $250/(0.1)^{1.4} \approx 6,300$. If the weekly offer probability is instead 0.05, the implied value for k is approximately 16,600.

Unobserved Heterogeneity. I introduce unobserved heterogeneity by assuming that the population is composed of two unobserved types, in proportions p and $1 - p$. Type 1 workers (in proportion p) have cost of search parameter $k_i + \Delta k$ and face a wage offer distribution with mean $\mu_i + \Delta \mu$. Type 2 workers have cost of search parameter k_i and face a wage offer distribution with mean μ_i . The parameter Δk is restricted to be positive, indicating that type 1 workers can be viewed as low cost of search workers, while type 2 workers are high cost workers. No restrictions are placed on $\Delta \mu$.²¹

¹⁹The regression is weighted because the actual cost function $\tilde{c}(s)$ and the approximate cost function $c(s)$ can differ substantially at extreme values of s . Therefore we give more weight to those values of s that have higher density empirically.

²⁰In 1983 dollars.

²¹It is not necessary to interpret the mixture distribution as actual heterogeneity. One may also view the sample as composed of homogeneous workers, who own a fixed stock of two types of “search capital” at the beginning of their unemployment spell: one type of capital generates job offers at low cost but depreciates rapidly (think of this as search through friends and relatives); the second type of capital generates job offers at high cost but depreciates more slowly (search through formal channels).

To illustrate behavior of the key variables in the solution, I depict in Figure 1 the dynamic evolution of the probability of exiting unemployment $h(t)$ and of the expected re-employment wage $E(w|w > w_t^*)$. The figure shows the dynamics of these variables for low cost workers, high cost workers, and for the aggregate of the two types. For both types the exit rate is monotonically increasing, and the expected re-employment wage is monotonically decreasing. When we look at the aggregate variables, we see that the introduction of heterogeneity generates a non-monotonic shape for the hazard function, with a moderate spike at the time of exhaustion of UI benefits.

3.3. Identification

Using the functional form specification described above, there are 27 unknown parameters that need to be estimated; the two discounting parameters, β and δ ;²² the value of time when unemployed, b_0 ; the parameter vectors of the wage distribution μ and σ (six elements each); the standard deviation of the measurement error in log wages, σ_u ; the cost of search parameter vector \mathbf{k} (six elements); and the parameters of the heterogeneity distribution, Δk , $\Delta\mu$ and p . In addition, one must specify a value for the marginal cost elasticity η and for the probability of layoff q . The marginal cost elasticity η is set at 0.4, following the calculations described in Section 3.2. I estimate q separately for the three wage groups using employment spells from the NLSY from 1985 to 1996. This results in $q = 0.0111$, $q = 0.0105$, and $q = 0.0087$ for the low, medium, and high wage groups respectively.

It is important to differentiate between two aspects of the identification problem: on one hand, a key question in the job search literature concerns the identifiability of the time discounting parameter from data on unemployment durations and accepted wages alone. On the other hand, in the literature on time discounting, the crucial question is whether it is possible to distinguish empirically between a model with hyperbolic agents from a model with exponential, but highly impatient, agents. The first problem is well understood (Flinn and Heckman, 1982). To overcome it, I rely on a tight parameterization of the wage offer distribution, of the cost of search function, and on the time profile of the utility of unemployment, and then assess robustness of the empirical results to different choices of functional form.

In this section I give a heuristic description of the reasons that permit separate identification of the two parameters of the discount function. Consider first an unrestricted model in which workers choose search effort and reservation wages, possibly not based on any optimality principle. The parameters of this model, the reservation wage and the probability of receiving an offer

²²The time unit for the model is a week, but, for convenience of interpretation, I choose to present all the results in terms of the *yearly* discount factor, which, with slight abuse of notation, is denoted by δ .

in any period, are identified using data on reservation wages and duration data alone, thanks to the assumption that the entire wage offer distribution is known. In any period t , the reservation wage w_t^* can be estimated by the minimum observed wage for those who exit unemployment in that period. Given the assumption of stationarity from period T onwards, one can estimate the stationary reservation wage as the minimum of all observed wages among those with unemployment spells longer than T weeks. If the wage distribution is *recoverable*, (Flinn and Heckman), the distribution of observed wages for those unemployed longer than T periods, together with the reservation wage, identifies the parameters of the wage distribution. Having identified the parameters of the wage distribution and the reservation wages, it is easy to calculate the conditional probability of acceptance in each period. This, together with the data on the duration of unemployment spells, helps to identify the level of search effort in every period, s_t : in the model, s_t is simply the probability of receiving a job offer, and is equal to the ratio between the exit rate from unemployment and the probability of acceptance. Having identified s_t and w_t^* , identification of the other model parameters follows in principle by imposing the restrictions implied by the dynamic optimization problem.

In practice, however, it may be difficult to separately identify all the model parameters. I should reiterate, though, that the discounting parameters β and δ can be separately identified when the other model parameters are known. As discussed in Section 2.3, the two parameters play fundamentally distinct roles in the job search process: the short-run discount factor β affects mainly the search effort decision, a decision involving a trade-off between immediate costs and benefits that materialize in the near future; on the other hand, the reservation wage decision involves primarily a trade-off between payoffs to be received at two different points in the more distant future, and therefore is nearly unaffected by β (it is affected by β , but only indirectly through the effect of β on the continuation payoff). This suggests that, to a first approximation, one can use data on reservation wages to identify δ . Then, for a known value of δ , it is straightforward to identify β using duration data.

To better illustrate this point, I generate some simulated unemployment histories,²³ and then study the behavior of the likelihood function assuming that data is available either on both duration and re-employment wages, or on duration alone. Figure 2a shows the log-likelihood and likelihood profiles as functions of β assuming either complete or incomplete data.²⁴ It

²³I generated 5000 simulated unemployment histories according to the model described in Section 2. The true values of β and δ were 0.4833 and 0.95 respectively. All the other parameter values were set at their maximum likelihood estimates for the medium wage sample (Table 3).

²⁴The likelihood and log-likelihood are normalized to facilitate comparison between the complete and incomplete data functions.

is apparent that data on wages does not contribute significantly to the identification of β . On the other hand, the identification problem stands out when we analyze the likelihood profiles as functions of the long-term discount factor δ . When data on wages is not available, the log-likelihood as a function of δ is essentially flat, and the likelihood is substantially spread out; incorporating information on accepted wages allows us to identify δ .

4. Results

4.1. Data

My data contains information on the duration of unemployment spells and re-employment wages for males in the NLSY. I use the Work History files to construct a week by week account of every male worker's labor force status from 1978 to 1996. A worker is defined to be unemployed if he is out of a job but willing to work. I classify as unemployment spells all periods of nonemployment in which at least some active search took place. This measure differs from the conventional definition in that a worker who does not actively search during the entire spell can still be classified as unemployed. The re-employment wage is taken to be the average weekly wage (in 1983 dollars) in the first job after the end of the unemployment spell.

I retain only those spells that were reported in 1985 or later by male respondents with no health problems, who were not part of the military subsample, and were not enrolled in school. This ensures that my sample of spells includes mainly workers with strong attachment to the labor force. In addition, since I am interested in estimating a model with time-varying unemployment benefits, I retain only spells in which Unemployment Insurance was received for at least one week. I discard observations where the logarithm of the re-employment wage is more than three and a half standard deviations below or above the mean, or with missing data on the wage prior to the unemployment spell. This leaves me with 1797 observations, representing 1008 different individuals. For the purposes of this paper, I assume that different spells by the same individual can be treated as independent spells.

The empirical analysis is performed separately for three groups classified by their previous earnings. The low earnings group includes workers in the bottom quartile of the sample distribution of previous weekly wages; the medium earnings group includes individuals in the middle two quartiles; the top earnings group includes individuals in the top quartile. Summary statistics on the duration of unemployment spells, on re-employment wages, and on a set of demographic characteristics are presented in Table 1.

There are marked differences between the three groups in all the variables. Workers with higher previous earnings have shorter unemployment spells and are employed at higher wages at the end of their spell. Moreover, as we move across the earnings distribution, we find workers who are older, more educated, more likely to be married, and with higher levels of cognitive ability as measured by the AFQT score. By contrast, the replacement rate falls from about 70% of the previous wage for low wage workers to about 30% for high wage workers. Overall there is substantial heterogeneity in the sample.

4.2. Implementation

The theoretical model imposes restrictions on the range of possible parameter values. For example the long term discount factor δ must be between 0 than 1. Hyperbolic discounting posits that the short-run discount factor β be smaller than unity, but the model does allow values of β greater than one, provided that β is smaller than $1 + \eta$.²⁵ To ensure that the optimization algorithm does not wander off testing nonsensical values, I apply a logistic transformation to all the model parameters: for any structural parameter θ_j restricted to the interval (c_1, c_2) , I define $\gamma_j = \log [(\theta_j - c_1) / (c_2 - \theta_j)]$. I then reparameterize the likelihood in terms of the γ 's, which are allowed to take on any values on the real line. The maximum likelihood estimate for the original structural parameter θ_j is easily obtained as $\hat{\theta}_j^{ML} = [c_1 + c_2 \exp(\hat{\gamma}_j^{ML})] / [1 + \exp(\hat{\gamma}_j^{ML})]$. Standard errors for γ are obtained using the outer product gradient method, and standard errors for θ are derived from these using the delta method. The results of the structural estimation are given below.²⁶

²⁵Solution of the stationary system in (2.5) and (2.6) yields the following two equations in two unknowns, w^* and s^* :

$$\begin{aligned} \frac{Q(w^*)}{(w^* - b)^{1/(1+\eta)}} &= \left(\frac{k}{\beta}\right)^{1/(1+\eta)} \frac{(1+\eta)(1-\delta)(1-q)}{\delta(1+\eta-\beta)^{1/(1+\eta)}} \\ s^* &= \left[\frac{\beta(w^* - b)}{k(1+\eta-\beta)}\right]^{\frac{1}{1+\eta}}. \end{aligned}$$

The solution is well defined for $\beta < 1 + \eta$.

²⁶The model is estimated using a pre-conditioned conjugate gradient algorithm with numerical derivatives from the NAG Fortran library. Convergence to the maximum and calculation of standard errors took approximately 15 minutes on a dual processor 800 MhZ Pentium computer. To ensure that the algorithm had in fact converged to a global maximum, the optimization procedure was repeated with different starting values.

4.3. Structural parameter estimates

In Table 2 I present parameter estimates and standard errors for the structural parameters in the three subsamples. For the medium wage sample I report the results of estimation assuming both a log-normal and a normal wage distribution. In the other two subsamples the normal model had difficulties to converge, hence only results for the log-normal model are reported.

The first observation to be made is that the estimate of the hyperbolic discounting parameter β in the lognormal model increases as we move from the low to the high end of the wage distribution. For low and medium wage workers we find a high degree of present bias, whereas for high wage workers the degree of short run impatience is relatively small. The estimate of β is markedly higher in the normal model, though still substantially lower than 1. The discrepancy in the estimates of β between the normal and the lognormal models is due to the fact that we can only observe wages above the reservation wage, so that in practice identification of the acceptance probability relies on the functional form assumption. This is a limitation of the current model, which could be overcome by incorporating outside information on the actual or perceived distribution of wage offers.

The ranking of β across the distribution of previous wages is not surprising: we expect workers who are able to delay gratification and have relatively little self-control problems to be rewarded in the labor market. I test the restriction $\beta = 1$ using a likelihood ratio test. The results of these tests are reported at the bottom of Table 2. In the low wage sample and in the lognormal medium wage samples, the exponential model is soundly rejected. The evidence in the normal medium wage and high wage sample is ambiguous, with p-values that slightly exceed conventional significance levels.

The point estimates for β in the low and the lognormal medium wage samples are somewhat lower than those found in the experimental literature, (Thaler, 1981; Benzion, Rapoport and Yagil, 1989; Kirby, 1997), but not totally implausible, especially if we take into account that the experimental subjects are likely not representative of the whole population. Fang and Silverman (2002) estimate a similar model of job search for women receiving welfare, and they estimate β equal to 0.61, with quite a large standard error. Laibson, Repetto and Tobacman's (2004) structural estimates of β , based on life-cycle consumption choices, range between 0.51 and 0.82. Altogether, my point estimates are roughly comparable to those found elsewhere.

The estimate of the degree of long-run discounting δ lies extremely close to the boundary of the parameter space, and its confidence interval is essentially uninformative. Interestingly, we find that the value of time when unemployed b_0 is substantially negative in all three samples based

on the lognormal distribution, and essentially zero in the normal model. Large negative values for b_0 imply that the mere fact of being unemployed bears disutility. This is consistent with the literature on subjective measures of happiness (Oswald, 1997; Winkelmann and Winkelmann, 1998): these studies find that unemployment has a universally negative effect on subjective measures of well-being, which goes well beyond that predicted by the income drop associated with joblessness. Moreover, other structural estimates of the job search model also find a similar result: van den Berg (1990) reports that workers “dislike being unemployed ... for non-pecuniary reasons,”²⁷ and Wolpin(1987) estimates a positive weekly cost of search ranging from 104 to 223 dollars. Note also that the disutility is greater as we move up the wage distribution: the psychic cost of being unemployed is larger for workers who experience unemployment (either personally or through acquaintances) relatively rarely.

The estimated mean of the offered wage distribution lies substantially below the mean of observed wages in all population subgroups, indicating that a sizeable fraction of wages are rejected. Similarly, the estimated standard deviation of the wage distribution is larger than the observed standard deviation in all subsamples. Measurement error accounts for only a small fraction of the total variation in log wages. Altogether, there is substantial heterogeneity in the parameters of the wage distribution across population subgroups, for all three samples.

The cost of search parameter vector \mathbf{k} follows a general pattern in all specifications, with costs of search being the highest for unmarried workers with low AFQT and the lowest for married workers with high AFQT. It is conceivable that low-skill workers do indeed face a higher cost of obtaining a job offer with a given probability. Similarly, the result on marital status can be explained by the fact that married workers have a broader network of acquaintances that may facilitate the job search process. It is worth noting that the pattern of the parameters is the same pattern that appears in the distribution of unemployment duration (which is correlated negatively with cognitive ability and marriage): it appears that the cost of search parameter is governing most of the variation in exit rates across groups.

4.4. Predicted values

We now move to analyzing how the various specifications compare in terms of predicted values for the variables of interest. The results are presented in Table 3. I compare the mean values for duration and re-employment wages in the sample to those predicted by the model.²⁸ Likewise,

²⁷ van den Berg (1990), page 270.

²⁸The predicted values are the expectations of the relevant variables, evaluated analytically at the maximum likelihood estimates, and integrated over the distribution of both observed and unobserved types. See Appendix

I present predicted values for the expected weekly acceptance probability, and the expected weekly cost of search. All the models do a fairly good job of matching the average re-employment wage. The expected unemployment duration is consistently overestimated, but this was to be expected, given that in the calculation of predicted unemployment duration we give positive weight to extremely long spells that would be censored in the actual data.

The predicted values for the unobserved quantities raise some questions. In the lognormal models, the predicted acceptance probability ranges between 14 and 26 percent, while it rises to 59 percent in the normal model. The latter figure matches quite closely acceptance probability estimates of around 60 percent based on survey evidence (Blau and Robins, 1990), but is lower than estimates from previous structural models, which found that essentially all offers are accepted (Wolpin, 1987, Van den Berg, 1990).²⁹ The substantial fraction of rejected offers may be due to the very high degree of long run patience, which implies that workers are quite selective in their job acceptance strategy.

4.5. Dynamics

The expected values in Table 3 hide a rich set of dynamic behavior in the model. Table 4 and Figures 3a-3d show the dynamic evolution of several variables of interest over the course of an unemployment spell. Dynamics in the model (see Figure 1) are generated by the limited availability of unemployment insurance, and by the changing composition of the sample as the unemployment spell progresses: workers with an unobserved low cost of search exit unemployment much more frequently, so that the pool of workers who are still unemployed is increasingly dominated by the high cost individuals. These two forces operate in opposite directions on the exit rate from unemployment. As benefits run out, workers increase the intensity of search and begin to accept a larger fraction of wage offers, so that, for a given worker type, the exit rate increases in the first 26 weeks. On the other hand, the aggregate exit rate decreases as the sample is increasingly dominated by high cost types with low exit rates. The expected re-employment wage, by contrast, decreases monotonically: as the time of benefit exhaustion approaches both types of workers set a lower reservation wage, and, because high cost workers have lower reservation wages, the aggregate expected re-employment wage also declines as more and more low-cost workers exit unemployment.

Figures 3a-3d show that in all samples and specifications the model does a good job at fitting

B for details.

²⁹Some structural equilibrium search models also find low values of the acceptance probability (Eckstein and Wolpin, 1995).

the survivor function, even though it cannot capture the high week to week variation in exit rates and expected re-employment wages. In the low and medium wage samples the exit rate is initially flat and begins to rise as week 26 (the time of exhaustion of UI benefits) approaches. In the high wage sample the exit rate falls monotonically. In the medium wage sample, the normal model seems to be able to fit the spike in exit rates at the time of benefit exhaustion more accurately than the lognormal model.

In comparing between different types of workers,³⁰ Table 5 suggest that most of the difference in exit rates between high and low wage workers can be attributed to differences in (endogenously determined) offer arrival rates. In fact, conditional on receiving offers, low-wage workers accept a higher proportion of wages, but have longer unemployment spells because of low offer probabilities. This finding is in accordance with much of the previous literature, which finds that variation in offer arrival rates plays a much larger role in explaining unemployment duration than variation in reservation wages (see Devine and Kiefer, 1991). On the other hand, in the comparison between medium wage and low wage workers, we see that the former have higher exit rates because of higher acceptance probabilities.

In the lognormal model, predicted acceptance probabilities are quite low during the first few weeks of the spell, ranging from 12 to 20 percent. In the low and medium wage samples, the proportion of acceptable offers rises to 34-50 percent after one year, and to 77 to 90 percent after two years. This simply reflects the fact that after two years the sample is composed almost entirely of high-cost types who accept nearly any offer. In the high wage sample, the acceptance probability remains at very low levels even after two years. In the normal model, the dynamics of acceptance probabilities does not show such sharp differences over time.

The cost of search behaves in a very non-linear fashion: in the low and medium wage samples it increases in the first 26 weeks, and then decreases; in the high wage sample it dips in the first few weeks, then rises up to week 26, and then decreases. This pattern is explained by the fact that as the time of benefit exhaustion approaches, all worker types search more intensively, driving up search costs; but this effect is offset by the increasing proportion of high cost types, who search less intensively and actually end up paying a lower overall cost of search.

4.6. Alternative forms of heterogeneity

A natural question that arises is whether the results are robust to different specifications of the structural model. In particular, the basic specification adopts a very simple structure for

³⁰We focus here on the lognormal model where the comparison is feasible.

the way in which observed and unobserved heterogeneity affect the model parameters. I now consider three alternative forms of heterogeneity. In Alternative 1, I let the distribution of the unobserved component have three mass points rather than two. In Alternative 2, I add race as one of the explanatory variables for the parameters of the cost of search function and of the wage offer distribution, while maintaining a two-point distribution for unobserved heterogeneity. The model estimated in Table 2 is nested in these first two alternative models, so that we can perform likelihood ratio tests to assess whether the alternatives significantly improve the fit of the model.

In Alternatives 3 I follow a substantially different approach. I allow observed covariates to be correlated with type probabilities: I eliminate the direct dependence of μ and k on X_i , and instead specify that the probability that individual i is of type j is equal to

$$P(\text{individual } i \text{ is of type } j) = \frac{e^{\alpha'_j X_i}}{\sum_k e^{\alpha'_k X_i}}.$$

In practice, I let there be three types, and X_i is a five-element vector containing a constant, dummies for medium and high AFQT scores, a dummy for whether the individual is married and a dummy for non-whites.³¹

The results are presented in Table 5. I focus on the medium wage sample and the lognormal model, but I also present the estimation results for alternative 3 for the normal model. Alternatives 1 and 2 significantly improve the fit of the model compared to Table 2, as can be seen from the likelihood ratio test statistics and its associated p-values. However, in terms of the point estimate of β , there is little difference between the two models and the basic specification. While the standard errors around β are lower in the richer specifications, the likelihood ratio test statistic for the null hypothesis of exponential discounting gives somewhat weaker results, and in fact we cannot always reject the null hypothesis of exponential discounting. The third alternative, which adopts a substantially different framework for incorporating observed and unobserved heterogeneity, yields a slightly larger point estimate for the short-run discounting parameter, with a relatively large standard error, in both the log-normal specification and the normal specification. In contrast with the results in Table 2, the likelihood ratio test rejects the null hypothesis of exponential discounting in the normal model, despite the slightly higher point estimate. To summarize, introducing richer structures for the effect of observables and unobservables on the model parameters does not substantially affect our estimate of the degree of present bias, but comes at a cost of a loss in precision.³²

³¹In this model, there are 20 unknown parameters: β , δ , b_0 , μ (scalar), σ (scalar), σ_u , $\mathbf{d}\mu$ (vector of length 2), $\mathbf{d}\mathbf{k}$ (vector of length 2), and $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$, a vector of length 10. The obvious normalization is $\alpha_3 = 0$.

³²Additional robustness checks involved the cost of search function and the time profile of unemployment

5. Policy Evaluation

One of the main advantages of structural estimation is that it allows one to simulate the effects of different policy interventions in a behaviorally consistent manner.³³ In addition, it is possible to evaluate the effects of a given policy on outcomes that are not always observable in other contexts: for example, by fully specifying the agent's preferences and assuming that the agent follows an optimal strategy, we can carry out welfare comparisons between alternative policies. In our setting, this can be particularly important, because a hyperbolic agent's dynamic inconsistency may imply that some policy intervention can raise his or her long-term welfare. Alternatively, one may wish to assess whether the estimated effect of a given policy is sensitive to the assumption made about intertemporal preferences.

For a range of different policies and different outcome variables, I ask two questions:

1. Using the maximum likelihood estimates from Tables 3, and assuming that the worker has hyperbolic preferences, what is the effect of the policy on the outcome?
2. How is the conclusion altered if one assumes instead that agents have exponential preferences?

To answer this second question, I re-estimate the model parameters under the restriction that $\beta = 1$, and then evaluate the effects of the different policies. I call these *calibrated* policy evaluations.³⁴

I evaluate the effect of any policy on four outcome variables: a) the expected duration of unemployment; b) the expected re-employment wage; c) the expected level of government payments; d) the individual's utility. With hyperbolic preferences, the correct notion of individual utility is difficult to define, because of the potential conflict between an individual's different selves:

benefits. When the parameter η in the cost of search function was allowed to vary freely, its maximum likelihood estimate was 1.01, and the estimate of β was 0.27; however, standard errors increased considerably, and the restriction $\eta = 0.4$ could not be rejected using a likelihood ratio test. When the value of time when unemployed was assumed to be independent of unemployment benefits (essentially making the model stationary), the model failed to converge.

³³However, one should keep in mind that this is a partial equilibrium model and all the policy experiments conducted here ignore the demand side of the labor market. Therefore, the policy experiments should be interpreted with caution.

³⁴One could also perform non-calibrated policy evaluations: holding all other parameters fixed, what is the effect of a given policy for different values of β ? I omit these calculations because as β changes the benchmark values for the parameters of interest also vary substantially, making comparisons difficult.

future events are discounted differently at different points in time, and an optimal strategy from today's perspective may no longer be optimal in the future. I follow here O'Donoghue and Rabin (2001), and evaluate the different policies using the perspective of the long-run self. The long-run self's utility is simply the utility derived from following the strategy chosen by the hyperbolic agent, discounted exponentially. Because of dynamic inconsistency, this strategy is not optimal from the long run self's perspective. The long run criterion can be thought of as the utility criterion used by a voter who is not currently unemployed when deciding whether to implement a change to the UI system.

The five different policies are: a) cutting the level of unemployment benefits; b) a job search assistance program; c) monitoring the intensity of search effort; d) monitoring the job search acceptance strategy; e) a re-employment bonus program. I describe the policies below. The policies are modeled to resemble as closely as possible interventions that are, or have been, actually implemented. I restrict attention to the medium wage sample.

5.1. Policies

Benchmark. In the benchmark model, the unemployed worker receives Unemployment Insurance benefits for the first 26 weeks of the unemployment spell. The level of benefits for each population subgroup is taken as the average level of benefits for that group in the medium wage sample.

Cutting the level of unemployment benefits. This is the policy that is most commonly analyzed. I model it straightforwardly by cutting the level of unemployment benefits b_{UI} by 20 percent.

Job search assistance program. Many government UI agencies attempt to improve claimants' re-employment prospects by providing a variety of job search assistance programs (see Meyer, 1995, for a survey). The program may include classroom training, help with writing resumes, facilitating contact between employers and job seekers, or one-on-one counseling sessions. I model these types of programs as a 10 percent reduction in search costs faced by the unemployed for the first 26 weeks of the unemployment spell. I assume that the government bears the burden of this cost reduction dollar for dollar. In terms of the model parameters, this means that the cost level parameter k faced by the worker falls by 10 percent, and that government expenditures rise by $0.1k$ for every week of insured unemployment.

Monitoring search intensity. Some UI systems combine the carrot represented by job search assistance programs, with the stick represented by a tightening of the eligibility requirements

for receipt of benefits (see Meyer, 1995; Ashenfelter, Ashmore, and Deschênes, 1999; van den Berg, van der Klaauw and van Ours, 2004). This typically involves more frequent contacts with an employment agency representative to demonstrate active job seeking. I model this policy by assuming that the government can observe the level of search effort s exerted by the worker: if this level falls below $\underline{s} = 0.1$, the worker is no longer eligible to receive UI benefits.

Monitoring the job acceptance strategy. In many UI systems, in order to be eligible for benefits, claimants must be actively looking for work, and must be available to start a job immediately. However, there might be some period at the beginning of a spell in which the claimant is allowed to restrict availability to jobs in his or her occupation or on the basis of pay.³⁵

In modeling this policy, I assume that the government can observe the wage w a worker is offered, and that the worker will lose eligibility to benefits if he or she rejects an offer below a certain threshold w_{min} . I set w_{min} at 0, meaning that any rejected offer will imply the loss of benefits.

Re-employment bonus. Several US states have experimented with re-employment bonus programs (Meyer, 1995), modeled on the successful Illinois Re-employment Bonus Experiment (Woodbury and Spiegelman, 1987). A typical re-employment bonus program could involve paying a bonus equal to 10 weeks of UI benefits to workers who found a job within 13 weeks, and then were able to hold that job for 13 more weeks. Translating this policy to our model is straightforward.

Figures 4a-4e present a graphical depiction of the effects of each policy.

5.2. Results: calibrated policy evaluation

I present the results of the calibrated policy evaluations in Tables 6a and 6b. Table 6a refers to the lognormal model, and Table 6b to the normal model. The first four columns of the Table present the effects of the policies using parameter estimates from the hyperbolic model. These are simply the parameter estimates from Table 3. The last four columns present the effects of the policies using parameter estimates from a restricted model, where β is set equal to 1. For

³⁵The following excerpt is taken from the United Kingdom’s Jobseekers’ Allowance Charter (www.employmentservice.gov.uk).

“If you are looking for work in your usual occupation, you may be allowed a period of between one and 13 weeks at the beginning of your claim, ... during which you are able to restrict your availability to that occupation and your normal rate of pay. You cannot normally refuse a job on the basis of pay after six months.”

each policy, I report its estimated effect both in levels and in logs. The reported change in utility can be interpreted as the certainty equivalent of each policy: it is the lump sum amount that the worker would be willing to pay immediately in order to see the policy implemented (alternatively, it is the lump sum amount by which one would need to compensate the worker immediately in order to implement the policy).

We analyze first the results for the lognormal model. A 20 percent decrease in the level of UI benefits leads to roughly a one week drop in expected duration in the hyperbolic model, and a 1.14 weeks drop in expected duration in the exponential model. This corresponds to a 4.36 to 5.13 percentage decrease in duration, or an elasticity of duration with respect to benefits between 0.22 and 0.26. This estimate falls squarely in the range of previous findings in the literature (see Devine and Kiefer, 1991). The effect of the change in benefits on the expected re-employment wage is negligible: this appears to be true of all policies, and confirms that search effort is the main channel driving variation in exit rates. Expected government expenditures fall by more than 20 percent because of agents' behavioral response to the policy. The effect of the policy is similar for the hyperbolic and the exponential model in these dimensions. However, the utility drop for the hyperbolic agent is somewhat smaller than that for the exponential agent.

The job search assistance program is interesting in that it has a significantly larger effect on outcomes for the exponential agent than for the hyperbolic. The elasticity of expected duration with respect to search costs is 0.46 for the hyperbolic worker, but 0.65 for the exponential. Government expenditures rise by an order of magnitude: this seems to imply that a job search assistance program would always fail a cost-benefit analysis. This stands in contrast to the empirical findings on job search assistance programs: Meyer (1995) documents that four of the five job search assistance programs he analyzes yield higher benefits than costs to the government.³⁶ These differences can be reconciled if one believes that there are returns to scale in job search costs, so that the equivalent of one dollar in job search assistance to the individual worker actually costs less than one dollar to the government. Moreover, there might be other effects that are not taken into account in my simple formulation. For example, if the program shifts the distribution of potential wage offers, the cost-benefit analysis could be much more favorable.

The results of monitoring search intensity are striking. For the hyperbolic worker, this policy reduces expected unemployment duration by nearly eight weeks, and government expenditures (net of administrative costs) fall by 44 percent. Moreover, these drastic changes have actually a positive effect on the worker's utility!³⁷ In contrast, the effect of this policy for the exponential

³⁶See Meyer (1995), Tables 6A and 6B.

³⁷Here, worker type heterogeneity hides some important features of the policy. In the example, the restriction that search effort must be greater than 0.1 is binding only for the high cost types, and has no effect on the low

worker is much smaller, and he would be willing to pay a substantial amount to avoid its implementation. These results were to be expected: the hyperbolic worker searches less than optimally from the long run self's perspective: he would be willing to pay for a commitment device that forces him to search more intensively in the future. The government's threat to cut benefits if he does not exert enough search effort acts exactly as this desirable commitment device. However, the magnitude of the effect does not seem plausible. In fact, Ashenfelter, Ashmore and Deschênes (1999) find no evidence that stricter enforcement and verification of work search behavior results in decreased UI claims and benefits. Van den Berg, van der Klaauw and van Ours (2004) do find an effect of punitive sanctions on search effort for welfare recipients, but not as large as that found here.

Monitoring the job acceptance strategy has a much smaller effect on all the outcome variables, for both the hyperbolic and the exponential models. Expected duration falls by around half a week in the hyperbolic model, and is essentially unchanged for the exponential. The expected re-employment wage decreases by at most 7 dollars. This finding confirms once more the relatively diminished importance of the reservation wage channel in determining the length of unemployment spells.

Finally, the re-employment bonus has quite a sizeable effect on all outcomes. Expected duration falls by more than two weeks for hyperbolic agents, and by two and a half weeks for exponential ones. These effects are somewhat higher than the effects found in the re-employment bonus experiments surveyed by Meyer. This could be due to the fact that the size of the bonus used here is relatively high compared to the actual bonus paid out in most of the experiments. The elasticity of duration with respect to the bonus is similar to that found in the Illinois claimant experiment (which was by far the most successful of the bonus experiments), and higher than that found in the other experiments.

Table 7b presents a similar analysis using the estimates for the normal model in the medium wage sample. Compared to the previous estimates, we find that job search assistance has a similar impact on duration, while the impact of a change in UI benefits and of the re-employment bonus is substantially larger. The most important difference is in the effect of monitoring the acceptance strategy. Contrary to the lognormal model, this policy does have a large impact on unemployment duration. Finally, monitoring job search has large effects in both the hyperbolic and the exponential model. In this parametrization the policy reduces the long-run utility of

cost types. Moreover the required level of search is low enough so that the high cost types prefer to search exactly 0.1 rather than lose their benefits. This puts them closer to the desired level of effort that would have been chosen by a time-consistent agent, so utility rises. A different constraint could actually lower utility.

hyperbolic workers: the commitment value of the policy is dominated by the extra search costs induced.

In a separate analysis, not reported here, I found that most of the differences between the predicted effects of policies are due to differences in the remaining model parameters rather than in hyperbolic discounting per se. This does not mean that the hyperbolic discounting model should be dismissed because it does not deliver significantly different predictions from an exponential model. Quite the contrary: ignoring hyperbolic discounting leads one to estimate substantially different structural parameters, which in turn could lead one to draw incorrect conclusions on the effect of a given policy.

6. Conclusion

This paper estimates the structural parameters of a model of job search with hyperbolic discounting and endogenous search effort. The model is estimated using data on unemployment spells and accepted wages from the NLSY. The estimation strategy controls for both observed and unobserved heterogeneity. The focus of the paper is the estimation of the time discounting parameters. Assuming that the parameters of the wage distribution, the cost function, and the value of time when unemployed are known, data on the duration of unemployment spells and accepted wages identify the time discounting parameters.

The model is estimated separately for three wage groups. In all three subsamples, and in particular for low and medium wage workers, the short run discounting parameter is smaller than one, although the point estimate is somewhat sensitive to the assumed functional form of the wage offer distribution. In most specifications, likelihood ratio tests reject the exponential model. The model does a relatively good job at fitting observed outcomes such as the survivor function or the hazard rate. The paper also uses the structural parameters to evaluate alternative policy interventions for the unemployed. The impact of different policies varies substantially depending on whether the model is calibrated with hyperbolic or exponential preferences. Since hyperbolic workers exert less search effort than what is optimal from the long-run self's perspective, policies that induce hyperbolic workers to search more intensively may actually improve long-run welfare. The actual extent of the utility gain depends however on the exact form of the policy and on the structure of the model. While the structural model can give us insights on the generic effects of a given policy, more detailed data on search costs and wage distributions is necessary to draw precise conclusions for a specific experiment.

A. Appendix A: The likelihood function

The data consists of observations on the length of unemployment spells in weeks, T_i , on re-employment wages w_i , and on a set of individual characteristics X_i for a sample of individuals $i = 1, 2, \dots, N$. Assume for the moment that there are no censored spells and that the re-employment wage is observed for every individual in the sample. I take the week as the time unit, and I denote by $h_t(\theta, X_i) \equiv s_t(\theta, X_i) [1 - F_\theta(w_t^*(\theta, X_i))]$ the exit rate from unemployment in week t as a function of the model parameters θ and of individual characteristics X_i . The exit rate in week t is simply the product of the probability of receiving an offer in week t , $s_t(\theta, X_i)$, and the probability that this offer exceeds the reservation wage in week t , $w_t^*(\theta, X_i)$. The values of $s_t(\theta, X_i)$ and $w_t^*(\theta, X_i)$ are obtained from the solution of the dynamic programming problem (2.3). The likelihood contribution of individual i can be written (omitting the dependency on X_i for notational convenience) as

$$\begin{aligned} \mathcal{L}_i(\theta) &= \left(\prod_{\tau=1}^{T_i-1} [1 - h_\tau(\theta)] \right) \times [h_{T_i}(\theta)] \times \left(\frac{f_\theta(w_i)}{1 - F_\theta[w_{T_i}^*(\theta)]} \mathbf{1}(w_i \geq w_{T_i}^*) \right) \\ &= \left(\prod_{\tau=1}^{T_i-1} \{1 - s_\tau(\theta) [1 - F_\theta(w_\tau^*(\theta))]\} \right) \times \\ &\quad \times \{s_{T_i}(\theta) [1 - F_\theta(w_{T_i}^*(\theta))]\} \times \\ &\quad \times \left(\frac{f_\theta(w_i)}{1 - F_\theta[w_{T_i}^*(\theta)]} \mathbf{1}(w_i \geq w_{T_i}^*) \right). \end{aligned} \tag{A.1}$$

The first term represents the probability that individual i does not exit unemployment in any of the periods from $t = 1$ to $T_i - 1$; the second term represents the probability that the individual finds an acceptable offer in period T_i ; the last term is the probability density of observing a re-employment wage w_i , conditional on that wage being greater than the reservation wage $w_{T_i}^*$. Note that the likelihood is well-defined only if the actual wage is greater than the reservation wage for the week in which the worker actually finds work. This can result in extreme sensitivity of the estimates to a few outliers. To prevent this possibility, I make the plausible assumption, as in Wolpin (1987), that wages are in fact measured with error: let \tilde{w}_i be the true wage, and w_i^{obs} be the observed wage. Then the third element becomes

$$f(w_i^{obs} | \tilde{w}_i > w_{T_i}^*(\theta)) = \frac{P(\tilde{w}_i \geq w_{T_i}^*(\theta) | w_i^{obs}) f(w_i^{obs})}{P(\tilde{w}_i \geq w_{T_i}^*(\theta))}.$$

With a convenient specification of the joint distribution of w_i^{obs} and \tilde{w}_i , this term can be easily calculated.

Censored observations and missing wage data are easily incorporated in this setting. Let C_i be a dummy variable indicating whether the observed unemployment spell is complete, and let E_i be a dummy indicating whether the re-employment wage is observed.³⁸ The likelihood then becomes:

$$\begin{aligned} \mathcal{L}_i(\theta) = & \left\{ \left(\prod_{\tau=1}^{T_i-1} [1 - h_\tau(\theta)] \right) [h_{T_i}(\theta)] \left(\frac{P(\tilde{w}_i \geq w_{T_i}^*(\theta) \mid w_i^{obs}) f(w_i^{obs})}{P(\tilde{w}_i \geq w_{T_i}^*(\theta))} \right) \right\}^{C_i E_i} \times \\ & \times \left\{ \left(\prod_{\tau=1}^{T_i-1} [1 - h_\tau(\theta)] \right) [h_{T_i}(\theta)] \right\}^{C_i(1-E_i)} \times \left\{ \prod_{\tau=1}^{T_i} [1 - h_\tau(\theta)] \right\}^{1-C_i}. \end{aligned}$$

Summing over individuals, and taking logs, yields the following log-likelihood function:

$$\begin{aligned} L(\theta) = & \sum_{i=1}^N \sum_{\tau=1}^{T_i-1} \log [1 - h_\tau(\theta)] + \\ & \sum_{i:C_i=1} \log [h_{T_i}(\theta)] + \\ & \sum_{i:C_i=0} \log [1 - h_{T_i}(\theta)] + \\ & \sum_{i:C_i=1, E_i=1} \log \frac{P(\tilde{w}_i \geq w_{T_i}^*(\theta) \mid w_i^{obs}) f(w_i^{obs})}{P(\tilde{w}_i \geq w_{T_i}^*(\theta))}. \end{aligned}$$

The dependency on observed individual characteristics is made explicit by letting $h_t(\theta, X_i) = \tilde{h}_t(\theta_1; \theta_2' X_i)$: in practice, to keep the dimension of the problem manageable, the control variables will be a set of indicators, so that the hazard rate can be effectively be written as $\tilde{h}_t(\theta_1, \theta_{2i})$. Finally, one can add unobserved heterogeneity to the model by specifying a mixture distribution with discrete finite support for the parameter vector θ :

$$\theta = \begin{cases} \xi_1 & p_1 \\ \xi_2 & p_2 \\ \vdots & \\ \xi_K & p_K \end{cases}, \quad \sum_{k=1}^K p_k = 1$$

³⁸In my sample there are observations with missing re-employment wage data, even if the unemployment spell is complete.

The log-likelihood function for the model with unobserved heterogeneity becomes

$$\begin{aligned}
L_p(\theta) &= \log \int \exp [L(\theta)] dp(\theta) \\
&= \log \sum_{k=1}^K p_k \exp [L(\xi_k)].
\end{aligned} \tag{A.2}$$

B. Appendix B: Estimating the Marginal Cost of Search Elasticity

Model. Consider an extension of the model described in the text in which workers have already chosen the optimal probability of receiving an offer s , and must now decide how to optimally allocate their time between alternative search methods. Assume that there are K different methods that can be used to generate offers. Once the worker has chosen s , he must decide how intensively to use each method of search. Assume also that the effort exerted in this period has no bearing on the probability of receiving an offer in any of the latter periods, and that each search method generates offers from the same wage distribution. These assumptions make the model effectively static, and allow us to abstract from the issues of time discounting described in the text.

The worker's problem becomes one of optimally choosing intensity of search for each method so as to minimize search costs, subject to the constraint that the probability of receiving an offer be at least s . Formally, the problem is

$$\begin{aligned}
&\min_{X_1, \dots, X_K} \sum_{j=1}^K c_j X_j \\
s.t. \quad &P(X_1, \dots, X_K) \geq s \\
&X_j \geq 0, \quad j = 1, \dots, K
\end{aligned}$$

This problem is very similar to that of a firm minimizing costs subject to a technological constraint, or to the dual problem of a worker who minimizes expenditure subject to a utility constraint. Letting λ be the Lagrange multiplier, ($\lambda \geq 0$), the first order conditions are:

$$\begin{aligned}
-c_j + \lambda \frac{\partial P(X_1, \dots, X_K)}{\partial X_j} &\leq 0, \quad \text{for } j = 1, \dots, K, \text{ with equality when } X_j > 0; \\
P(X_1, \dots, X_K) &\geq s.
\end{aligned}$$

These conditions give rise to the optimal intensities of search $X_1(c_1, \dots, c_K, s), \dots, X_K(c_1, \dots, c_K, s)$, and to the minimized cost function $\tilde{c}(c_1, \dots, c_K, s)$. In what follows I show

how an appropriate choice of functional form can generate a simple tractable form for the cost function.

Analytical Solution. We make two important assumptions about the function $P(\cdot)$.

1. The probability of receiving an offer using method j is independent of intensity of search in method j' , $j' \neq j$.
2. The probability of receiving an offer using method j , $f(X_j)$, is an increasing and concave function of intensity of search in method j . Moreover, the probability of receiving an offer using method j when one does not use the method at all is equal to zero. Therefore, $f'(\cdot) > 0$, $f''(\cdot) < 0$, and $f(0) = 0$.

We can then write

$$P(X_1, \dots, X_K) = 1 - \prod_{j=1}^K [1 - f(X_j)]. \quad (\text{B.1})$$

This expression tells us that the probability of receiving an offer is equal to one minus the probability of receiving no offers at all using any of the K methods.

Now let

$$f(X_j) = 1 - (1 + X_j)^{-\alpha_j}, \quad \alpha_j > 0. \quad (\text{B.2})$$

It is easy to verify that this functional form satisfies all the necessary requirements. Using (B.1), we can express the first order conditions as

$$\frac{c_j}{c_k} = \frac{f'(X_j) [1 - f(X_k)]}{f'(X_k) [1 - f(X_j)]},$$

and, using (B.2), we obtain

$$\frac{c_j}{c_k} = \frac{\alpha_j (1 + X_k)}{\alpha_k (1 + X_j)} \quad (\text{B.3})$$

which yields

$$(1 + X_j) = \frac{\alpha_j c_k}{c_j \alpha_k} (1 + X_k).$$

Now, using the fact that $s = 1 - \prod_{j=1}^K [1 - f(X_j)] = 1 - \prod_{j=1}^K (1 + X_j)^{-\alpha_j}$, we can solve for the individual X_j 's:

$$X_j(s) = \max \left\{ \frac{\alpha_j}{c_j} \prod_{k=1}^K \left(\frac{c_k}{\alpha_k} \right)^{\alpha_k / \sum_i \alpha_i} (1 - s)^{-1 / \sum_i \alpha_i} - 1, \quad 0 \right\}$$

Then, the minimized cost function is

$$\begin{aligned}\tilde{c}(s) &= \sum_{j=1}^K c_j X_j(s) \\ &= \sum_{i=1}^K \left[\left(\sum_{J_i} \alpha_j \right) \left(\prod_{J_i} \left(\frac{c_j}{\alpha_j} \right)^{\frac{\alpha_j}{\sum_{J_i} \alpha_l}} \right) (1-s)^{-\frac{1}{\sum_{J_i} \alpha_j}} - \sum_{J_i} c_j \right] \times 1(s \in S_i)\end{aligned}$$

where $1(\cdot)$ is the standard indicator function; $S_i = \left\{ s \mid \sum_{j=1}^K 1(X_j(s) > 0) = i \right\}$ is the set of all possible s values such that exactly i methods of search are used; and $J_i = \{j : X_j(s) > 0, \text{ for } s \in S_i\}$ is the set of indicators for which search methods are used, when exactly i methods are used. For large s , when all search methods are used, this function takes the form

$$\tilde{c}(s) = A(1-s)^{-\varepsilon} - B. \quad (\text{B.4})$$

with $\varepsilon = 1/\sum_j \alpha_j$. The elasticity of the marginal cost of search is equal to $s(1+\varepsilon)/(1-s)$.

Estimation. We estimate the parameters of the cost function using the 1981 wave of the NLSY, that provides detailed information on the search activities of employed and unemployed youth.³⁹ For those who searched for a job in the four weeks prior to the NLSY interview, we have information on whether search resulted in a job offer for each of eleven different methods.⁴⁰ We also have information on the number of hours searched in the past week. I assume that the number of hours of search was constant in each of the past four weeks, and that the probability of receiving a job offer in a given week is independent of search effort in any other week. Let X_{ij} be the number of hours spent searching by individual i using method j , and let O_{ij} be a dummy variable indicating whether method j resulted in a job offer for individual i in any of the past four weeks. The probability that individual i received a job offer using method j in any of the past four weeks is

$$1 - (1 - f(X_{ij}))^4 = 1 - (1 + X_{ij})^{-4\alpha_j}$$

³⁹This is the same data used in Holzer (1988).

⁴⁰These were: 1) checked with the state employment agency; 2) checked with a private employment agency; 3) asked friends and relatives about jobs; 4) placed or answered newspaper ads; 5) took the civil services test or applied for a government job; 6) contacted any public organization; 7) contacted a school placement office; 8) asked teachers or professors about jobs; 9) checked with a labor union; 10) checked directly with employers; 11) other methods.

My estimation strategy consists of working with the conditional likelihood function:

$$\begin{aligned} \prod_{i=1}^N f(O_{ij}|X_{ij}, \alpha) &= \prod_{i=1}^N \prod_{j=1}^K f(O_{ij}|X_{ij}, \alpha_j) \\ &= \prod_{i=1}^N \prod_{j=1}^K \left[1 - (1 + X_{ij})^{-4\alpha_j}\right]^{O_{ij}} \left[(1 + X_{ij})^{-4\alpha_j}\right]^{1-O_{ij}}. \end{aligned}$$

Note that each α can be estimated consistently by maximizing the likelihood separately for each search method. This simplifies the calculations considerably. Given α , one can back out the cost coefficients c_1, \dots, c_{K-1} (c_K is normalized to 1) from the first order conditions (B.3):

$$\hat{c}_j = \frac{\hat{\alpha}_j (1 + \bar{x}_K)}{\hat{\alpha}_K (1 + \bar{x}_j)}$$

The estimation is carried out on the subsample of youth who were searching while unemployed, and did not list being in school as their main activity during the survey week. I focus on the four most popular search methods (state employment service, friends and relatives, newspaper ads, direct employer contact), and aggregate the other methods into the “other methods” category.

The first three columns of Appendix Table 1 present the percentage of the sample using each method, and, conditional on using the method, the average hours spent searching and the probability of receiving an offer in any of the past four weeks. I also report the implied weekly probability of receiving an offer. The next two columns report the estimates and standard errors for the parameters of the probability function, $\alpha_1, \dots, \alpha_5$, and for the cost coefficients c_1, \dots, c_5 .

Given α and c , one can evaluate the cost function $\tilde{c}(s)$ at various levels of search intensity s . These estimated values are then used to estimate the parameters of the constant marginal cost elasticity cost function, $c(s) = ks^{1+\eta}$, used in the dynamic programming model. The estimation is performed by running a weighted least squares regression of $\log \tilde{c}(s)$ on $\log s$, where the weights are proportional to a normal density centered at the estimated mean probability of receiving a job offer.⁴¹ The results of this regression are

$$\begin{aligned} \log \tilde{c}(s) &= 4.412 + 1.408 \log s \\ &\quad (0.003) \quad (0.0011) \end{aligned}$$

The estimated constant marginal cost elasticity η is then equal to the coefficient on $\log s$ minus one. This gives rise to the benchmark value of $\eta = 0.4$ used in the text.

⁴¹The regression is weighted because the actual cost function $\tilde{c}(s)$ and the approximate cost function $c(s)$ can differ substantially at extreme values of s . Therefore we give more weight to those values of s that have higher density empirically.

C. Appendix C: Calculation of Predicted Values

Let us first define the *aggregate* exit rate and the *aggregate* survivor function. Let $h_{jt} = s_{jt} \left[1 - F(w_{jt}^*)\right]$ be the exit rate for a worker of type j ($j = HI, LO$) in week t of an unemployment spell. The exit rate is the product of the probability of receiving an offer s_{jt} and the probability of accepting the offer $\left[1 - F(w_{jt}^*)\right]$. Let D be the random variable indicating the length of the unemployment spell. The survivor function at week t is the probability that a spell lasts t weeks or longer. The survivor function at week t for type j is

$$P(D \geq t | j) = z_{jt} = \prod_{s=1}^{t-1} (1 - h_{js}),$$

with the convention that $h_{j0} = 0$. Let ϕ_{j1} be the proportion of workers of type j in the sample, and let ϕ_{jt} be the proportion of workers of type j who are still unemployed at time t :

$$\begin{aligned} \phi_{jt} &= P(j | D \geq t) = \frac{P(D \geq t | j) P(j)}{\sum_k P(D \geq t | k) P(k)} \\ &= \frac{\phi_{j1} z_{jt}}{\sum_k \phi_{k1} z_{kt}}. \end{aligned} \tag{C.1}$$

The aggregate exit rate in week t is H_t :

$$\begin{aligned} H_t &= P(D = t | D \geq t) = \sum_j P(D = t | D \geq t, j) P(j | D \geq t) \\ &= \sum_j h_{jt} \phi_{jt}. \end{aligned} \tag{C.2}$$

The aggregate survivor function in week t is Z_t :

$$Z_t = P(D \geq t) = \prod_{s=1}^{t-1} (1 - H_s). \tag{C.3}$$

Expected Duration. Aggregate expected duration is given by:

$$\begin{aligned} E(D) &= \sum_{t=1}^{\infty} t P(D = t) = \sum_{t=1}^{\infty} t P(D = t | D \geq t) P(D \geq t) \\ &= \sum_{t=1}^{\infty} t H_t Z_t. \end{aligned} \tag{42}$$

Expected Re-employment Wage. Let \tilde{w}_{jt} be the expected re-employment wage for workers of type j who exit unemployment in week t . The *aggregate* expected re-employment wage in

week t is given by:

$$\begin{aligned}
\tilde{W}_t &= E\left(\tilde{W} \mid D = t\right) = \sum_j E\left(\tilde{W} \mid D = t, j\right) P(j \mid D = t) \\
&= \sum_j \tilde{w}_{jt} \frac{P(D = t \mid j) P(j)}{\sum_k P(D = t \mid k) P(k)} = \sum_j \tilde{w}_{jt} \frac{z_{jt} h_{jt} \phi_{j1}}{\sum_k z_{kt} h_{kt} \phi_{k1}} \\
&= \sum_j \tilde{w}_{jt} \frac{h_{jt} \phi_{jt} \sum_k \phi_{k1} z_{kt}}{\sum_k h_{kt} \phi_{kt} \sum_k \phi_{k1} z_{kt}} = \frac{\sum_j \tilde{w}_{jt} h_{jt} \phi_{jt}}{H_t}, \tag{C.4}
\end{aligned}$$

where the second to last equality is derived by using (C.1) in both the numerator and the denominator. The ex-ante expected re-employment wage is then

$$E\left(\tilde{W}\right) = \sum_{t=1}^{\infty} \tilde{W}_t H_t Z_t.$$

Average weekly cost of search, long run utility. The predicted values for the expectation of the average weekly cost of search and of long run utility are derived in analogous manner to the expected re-employment wage. Let $\bar{c}_{jt} = t^{-1} \sum_{s=1}^t c(s_{jt})$ be the average weekly cost of search for a worker of type j who exits unemployment in week t . Let u_{jt} be the expected long run utility for a worker of type j who exits unemployment at time t . Then, using a similar line of argument as for the calculation of the expected re-employment wage, we obtain the expected aggregate cost of search

$$E\left(\bar{C}\right) = \sum_{t=1}^{\infty} \bar{C}_t H_t Z_t = \sum_{t=1}^{\infty} \sum_j \frac{\bar{c}_{jt} h_{jt} \phi_{jt}}{H_t},$$

and the expected aggregate utility

$$E\left(U\right) = \sum_{t=1}^{\infty} U_t H_t Z_t = \sum_{t=1}^{\infty} \sum_j \frac{u_{jt} h_{jt} \phi_{jt}}{H_t}.$$

Conditional Probability of Acceptance. The conditional probability of acceptance is calculated as the expected total number of accepted offers divided by the expected total number of offers. Since eventually all workers exit unemployment, the expected total number of accepted offers is equal to 1. The expected total number of offers is the sum, over all periods t , of the probability of receiving an offer conditional on being unemployed at t , times the probability of being unemployed at t . Let $S_t = \sum_j \phi_{jt} s_{jt}$ be the aggregate expected probability of receiving an offer in period t . Then, the conditional probability of acceptance is:

$$\Pr(\text{Acceptance} \mid \text{Offer}) = \frac{\sum_{t=1}^{\infty} H_t Z_t}{\sum_{t=1}^{\infty} S_t Z_t} = \frac{1}{\sum_{t=1}^{\infty} S_t Z_t}.$$

References

- [1] **Ainslie, George W.** *Picoeconomics*. Cambridge: Cambridge University Press, 1992.
- [2] **Angeletos, George-Marios; Laibson, David I.; Repetto, Andrea; Tobacman, Jeremy and Weinberg, Stephen.** “The Hyperbolic Buffer Stock Model: Calibration, Simulation, and Empirical Evaluation.” *Journal of Economic Perspectives*, Summer 2001, 15(3), pp. 47–68.
- [3] **Ashenfelter, Orley; Ashmore, David and Deschênes, Olivier.** “Do Unemployment Insurance Recipients Actively Seek Work? Randomized Trials in Four U.S. States.” National Bureau of Economic Research (Cambridge, MA) Working Paper No. W6982, February 1999.
- [4] **Barron, John M. and Mellow, Wesley.** “Search Effort in the Labor Market.” *Journal of Human Resources*, Summer 1979, 14(3), pp. 389–404.
- [5] **Benzion, Uri; Rapoport, Amnon and Yagil, Joseph.** “Discount Rates Inferred from Decisions: an Experimental Study.” *Management Science*, March 1989, 35(3), pp. 270–84.
- [6] **Blau, David M. and Robins, Philip K.** “Job Search Outcomes for the Employed and Unemployed.” *Journal of Political Economy*, June 1990, 98(3), pp. 637–55.
- [7] **Bloemen, Hans G.** “Job Search, Search Intensity and Labour Market Transitions: An Empirical Analysis.” *Journal of Human Resources*, forthcoming, 2004.
- [8] **Bolles, Richard Nelson.** *What Color Is Your Parachute? 2000. A Practical Manual for Job-Hunters and Career-Changers*. Berkeley: Ten Speed Press, 2000.
- [9] **DellaVigna, Stefano and Paserman, M. Daniele.** “Job Search and Impatience.” Mimeo., Hebrew University of Jerusalem, December 2003.
- [10] **Devine, Theresa J. and Kiefer, Nicholas M.** *Empirical Labor Economics: The Search Approach*. New York and Oxford: Oxford University Press, 1991.
- [11] **Diamond, Peter and Kőszegi, Botond.** “Quasi-Hyperbolic Discounting and Retirement.” *Journal of Public Economics*, 2003, forthcoming.
- [12] **Eckstein, Zvi and Wolpin, Kenneth I.** “Duration to First Job and the Return to Schooling: Estimates from a Search-Matching Model.” *Review of Economic Studies*, April 1995, 62(2), pp. 263–286.

- [13] **Fang, Hanming and Silverman, Dan.** “Time-inconsistency and Welfare Program Participation: Evidence from the NLSY.” Mimeo., Yale University, April 2002.
- [14] **Flinn, Christopher J. and Heckman, James J.** “New Methods for Analyzing Structural Models of Labor Force Dynamics.” *Journal of Econometrics*, January 1982, *18*(1), 115-168.
- [15] **Gruber, Jonathan and Köszegi, Botond.** “Is Addiction ‘Rational’? Theory and Evidence.” *Quarterly Journal of Economics*, November 2001, *116*(4), pp. 1261–1304.
- [16] **Harris, Christopher, and Laibson, David I.** “Dynamic Choices of Hyperbolic Consumers.” *Econometrica*, July 2001, *69*(4), pp. 935-957.
- [17] **Holzer, Harry J.** “Search Method Use by Unemployed Youth.” *Journal of Labor Economics*, January 1988, *6*(1), pp. 1–20.
- [18] **Kirby, Kris N.** “Bidding on the Future: Evidence Against Normative Discounting of Delayed Rewards.” *Journal of Experimental Psychology: General*, March 1997, *126*(1), pp. 54–70.
- [19] **Kirby, Kris N. and Herrnstein, R.J.** “Preference Reversals Due to Myopic Discounting of Delayed Reward.” *Psychological Science*, March 1995, *6*(2), pp. 83–89.
- [20] **Laibson, David I.** “Golden Eggs and Hyperbolic Discounting.” *Quarterly Journal of Economics*, May 1997, *112*(2), pp. 443-77.
- [21] **Laibson, David I.; Repetto, Andrea and Tobacman, Jeremy.** “Self Control and Saving for Retirement.” *Brookings Papers on Economic Activity*, 1998, *0*(1), pp. 91–172.
- [22] **Laibson, David I.; Repetto, Andrea and Tobacman, Jeremy.** “A Debt Puzzle,” in *Knowledge, Information and Expectations in Modern Economics: In Honor of Edmund S. Phelps*, Philippe Aghion, Roman Frydman, Joseph Stiglitz and Michael Woodford eds., Princeton: Princeton University Press, 2003.
- [23] **Laibson, David I; Repetto, Andrea and Tobacman, Jeremy.** “Estimating Discount Functions from Lifecycle Consumption Choices.” Mimeo., Harvard University, January 2004.
- [24] **Lippman, Steven A. and McCall, John J.** “The Economics of Job Search: A Survey.” *Economic Inquiry*, September 1976, *14*(1), pp. 155–189.

- [25] **Meyer, Bruce D.** “Lessons From the U.S. Unemployment Insurance Experiments.” *Journal of Economic Literature*, March 1995, *33*(1), pp. 91–131.
- [26] **O’Donoghue, Ted D. and Rabin, Matthew.** “Doing It Now or Later.” *American Economic Review*, March 1999, *89*(1), pp. 103–124.
- [27] **O’Donoghue, Ted D. and Rabin, Matthew.** “Choice and Procrastination.” *Quarterly Journal of Economics*, February 2001, *126*(1), pp. 121–160.
- [28] **Oswald, Andrew J.** “Happiness and Economic Performance.” *Economic Journal*, November 1997, *107*(445), pp. 1815–1831.
- [29] **Phelps, Edmund S. and Pollak R.A.** “On Second-Best National Saving and Game-Equilibrium Growth.” *Review of Economic Studies*, April 1968, *35*(2), pp. 85–199.
- [30] **Thaler, Richard H.** “Some Empirical Evidence on Dynamic Inconsistency.” *Economics Letters*, 1981, *8*, pp. 201–207.
- [31] **van den Berg, Gerard J.** “Nonstationarity in Job Search Theory.” *Review of Economic Studies*, April 1990, *57*(2), pp. 255–277.
- [32] **van den Berg, Gerard J.; van der Klaauw, Bas and van Ours, Jan C.** “Punitive Sanctions and the Transition Rate from Welfare to Work.” *Journal of Labor Economics*, 2004, forthcoming.
- [33] **Winkelmann, L. and Winkelmann, R.** “Why are the Unemployed so Unhappy? Evidence from Panel Data.” *Economica*, February 1998, *65*(257), pp. 1-15.
- [34] **Wolpin, Kenneth I.** “Estimating a Structural Search Model: The Transition from School to Work.” *Econometrica*, July 1987, *55*(4), pp. 801–817.
- [35] **Woodbury, Stephen A. and Spiegelman, Robert G.** “Bonuses to Workers and Employers to Reduce Unemployment: Randomized Trials in Illinois.” *American Economic Review*, September 1987, *77*(4), pp. 513–530.

Table 1: Summary Statistics †

	Low wage		Medium Wage		High Wage	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean.	Std. Dev.
Duration	24.81	27.04	21.57	24.60	16.47	18.67
Re-employment wage	180.25	89.09	265.90	104.17	462.34	201.55
Log (Re-employment wage)	5.10	0.43	5.51	0.41	6.04	0.46
Previous Wage	150.10	37.31	277.27	50.05	532.48	204.98
UI benefits	95.39	42.53	138.52	94.71	167.81	47.70
Replacement rate	0.72	0.65	0.51	0.37	0.34	0.12
Age	28.18	4.13	29.07	3.89	30.49	3.90
Education	11.67	1.83	11.90	1.70	12.61	1.83
Married	0.33	0.47	0.44	0.50	0.57	0.50
AFQT	36.13	25.92	47.22	26.82	55.72	25.10
Number of observations	450		898		449	

† **Notes:** The sample includes all spells of unemployment for males not enrolled in school and not in the military, reported after 1985, in which Unemployment Insurance benefits were received for at least one week.
Source: Author's calculations from the NLSY.

Table 2: Estimated Model Parameters [†]

		Low Wage Sample	Medium Wage Sample		High Wage Sample
		Lognormal	Lognormal	Normal	Lognormal
Discounting Parameters					
	β	0.4021 (0.1075)	0.4833 (0.1971)	0.8140 (0.1672)	0.8937 (0.1441)
	δ	0.9962 (0.1848)	1.0000* (0.0001)	1.0000* (0.0019)	0.9989 (0.1798)
Value of time when unemployed					
	b_0	-141.61 (61.16)	-164.31 (61.43)	-7.38 (16.54)	-308.78 (193.53)
Parameters of the wage offer distribution					
	μ_1 (unmarried, low AFQT)	4.1545 (0.1996)	5.0230 (0.1547)	174.08 (17.69)	5.4296 (0.3408)
	μ_2 (unmarried, medium AFQT)	4.3613 (0.1789)	5.0993 (0.1230)	185.64 (11.32)	5.2297 (0.1331)
	μ_3 (unmarried, high AFQT)	4.9503 (0.2341)	4.8658 (0.1425)	163.47 (13.62)	4.5543 (0.1785)
	μ_4 (married, low AFQT)	4.3954 (0.2187)	5.1694 (0.1192)	178.88 (15.75)	5.6692 (0.1956)
	μ_5 (married, medium AFQT)	4.2771 (0.1337)	5.2790 (0.1150)	204.97 (10.98)	5.4365 (0.1361)
	μ_6 (married, high AFQT)	4.1175 (0.2683)	5.0735 (0.0837)	232.06 (14.05)	5.3669 (0.1235)
	σ_1 (unmarried, low AFQT)	0.7196 (0.0723)	0.5614 (0.0585)	153.85 (8.31)	0.7444 (0.1659)
	σ_2 (unmarried, medium AFQT)	0.5760 (0.0516)	0.4122 (0.0453)	122.06 (7.54)	0.5810 (0.0502)
	σ_3 (unmarried, high AFQT)	0.2786 (0.2653)	0.4793 (0.0397)	173.84 (7.76)	0.7108 (0.0944)
	σ_4 (married, low AFQT)	0.5413 (0.0770)	0.3525 (0.0459)	125.61 (16.44)	0.5658 (0.0767)
	σ_5 (married, medium AFQT)	0.5187 (0.0423)	0.3068 (0.0410)	124.87 (5.63)	0.4649 (0.0424)
	σ_6 (married, high AFQT)	0.5715 (0.1039)	0.3463 (0.0320)	88.58 (15.77)	0.5503 (0.0436)

[†] **Notes:** Standard errors in parentheses. Standard errors were obtained using the outer product gradient method.

* : Indicates that the estimate is at the border of the parameter space: standard errors should be viewed with caution.

Source: Author's calculations from the NLSY.

Table 2: Estimated Model Parameters – Continued [†]

	Low Wage Sample	Medium Wage Sample		High Wage Sample
	Lognormal	Lognormal	Normal	Lognormal
Standard Deviation of the Measurement Error in Log Wages				
σ_u	0.1330 (0.0256)	0.2256 (0.0197)	45.70 (5.85)	0.0990 (0.0182)
Parameters of the cost of search function				
k_1 (unmarried, low AFQT)	678 (382)	2,998 (1,818)	3,349 (2,380)	9,085 (8,106)
k_2 (unmarried, medium AFQT)	359 (217)	1,174 (797)	1,389 (964)	1,198 (879)
k_3 (unmarried, high AFQT)	3,362 (3442)	325 (273)	603 (474)	304 (194)
k_4 (married, low AFQT)	451 (343)	1,002 (780)	776 (603)	8,499 (6,681)
k_5 (married, medium AFQT)	83 (31)	905 (667)	885 (591)	1,216 (873)
k_6 (married, high AFQT)	118 (160)	170 (98)	916 (591)	1,073 (685)
Parameters of the heterogeneity distribution				
$\Delta\mu$	0.5456 (0.1262)	0.0201 (0.1185)	-20.68 (15.68)	0.2179 (0.1412)
Δk	12,374 (4,217)	19,114 (7,675)	11,199 (5,584)	56,857 (28,021)
p	0.2962 (0.0505)	0.2276 (0.0330)	0.2720 (0.0345)	0.2974 (0.0341)
<hr/>				
Number of Observations	450	898	898	449
Log-Likelihood	-4041.5	-8484.8	-8509.12	-4270.2
LR statistic for $\beta=1$	17.2268	13.2375	1.8977	3.0826
p-value ¹	0.00	0.00	0.17	0.08

[†] **Notes:** Standard errors in parentheses. Standard errors were obtained using the outer product gradient method.

* : Indicates that the estimate is at the border of the parameter space: standard errors should be viewed with caution.

Source: Author's calculations from the NLSY.

¹ Based on a $\chi^2(1)$ distribution.

Table 3: Predicted Outcomes – Averages [†]

LOW WAGE

	Empirical	Predicted
Duration	24.81	26.06
Re-employment wage	180.25	178.82
Log (re-employment wage)	5.10	5.18
Conditional acceptance probability	-	0.20
Cost of search ¹	-	70.73

MEDIUM WAGE

(lognormal model)

	Empirical	Predicted
Duration	21.57	22.43
Re-employment wage	265.90	259.35
Log (re-employment wage)	5.51	5.55
Conditional acceptance probability	-	0.26
Cost of search ¹	-	121.89

MEDIUM WAGE (normal model)

	Empirical	Predicted
Duration	21.57	22.35
Re-employment wage	265.90	268.34
Log (re-employment wage)	5.51	5.59
Conditional acceptance probability	-	0.59
Cost of search ¹	-	69.88

HIGH WAGE

	Empirical	Predicted
Duration	16.47	17.84
Re-employment wage	462.34	455.01
Log (re-employment wage)	6.04	6.11
Conditional acceptance probability	-	0.14
Cost of search ¹	-	760.92

[†] **Notes:** Entries in the Table represent the ex-ante expected values of the relevant variables. Source: Author's calculations from the NLSY.

¹ The cost of search is the average weekly cost of search over the unemployment spell.

Table 4: Predicted Outcomes: Dynamics †**Low Wage**

	Empirical survivor function	Predicted survivor	Exit rate	Offer Probability	Conditional acceptance probability	Expected re-employment wage	Cost of search
Week 1	1.00	1.00	0.040	0.263	0.152	186.37	63.36
Week 13	0.60	0.62	0.039	0.226	0.173	183.44	66.75
Week 26	0.32	0.36	0.047	0.212	0.224	175.74	88.48
Week 52	0.14	0.12	0.034	0.099	0.344	169.59	80.41
Week 104	0.03	0.03	0.025	0.033	0.770	161.32	74.48

Medium Wage (lognormal model)

	Empirical survivor function	Predicted survivor	Exit rate	Offer Probability	Conditional acceptance probability	Expected re-employment wage	Cost of search
Week 1	1.00	1.00	0.051	0.246	0.207	277.67	111.07
Week 13	0.56	0.55	0.047	0.202	0.232	270.18	112.72
Week 26	0.27	0.29	0.055	0.186	0.296	255.87	151.00
Week 52	0.10	0.09	0.033	0.067	0.493	220.87	124.18
Week 104	0.03	0.02	0.025	0.028	0.901	188.41	111.18

Medium Wage (normal model)

	Empirical survivor function	Predicted survivor	Exit rate	Offer Probability	Conditional acceptance probability	Expected re-employment wage	Cost of search
Week 1	1.00	1.00	0.050	0.091	0.552	284.06	47.98
Week 13	0.56	0.56	0.046	0.081	0.568	277.17	62.14
Week 26	0.27	0.28	0.071	0.117	0.611	264.83	160.75
Week 52	0.10	0.09	0.030	0.042	0.715	233.67	128.15
Week 104	0.03	0.02	0.026	0.036	0.727	226.91	117.96

High Wage

	Empirical survivor function	Predicted survivor	Exit rate	Offer Probability	Conditional acceptance probability	Expected re-employment wage	Cost of search
Week 1	1.00	1.00	0.070	0.562	0.125	496.16	732.72
Week 13	0.47	0.45	0.060	0.434	0.137	470.31	681.07
Week 26	0.19	0.21	0.054	0.323	0.168	427.92	700.25
Week 52	0.05	0.06	0.038	0.167	0.225	352.03	543.47
Week 104	0.01	0.01	0.026	0.069	0.377	272.42	367.55

† Source: Author's calculations.

Table 5: Alternative Forms of Heterogeneity

	Alternative 1	Alternative 2	Alternative 3	Alternative 4
	Model: Lognormal	Model: Lognormal	Model: Lognormal	Model: Normal
Observed Heterogeneity Variables	AFQT, marital status	AFQT, marital status, race	AFQT, marital status, race	AFQT, marital status, race
Number of mass points in unobserved heterogeneity distribution	3	2	3	3
Link between unobserved heterogeneity and control variables	Independent	Independent	$P_{ij} = \frac{e^{\alpha_k X_i}}{\sum_k e^{\alpha_k X_i}}$	$P_{ij} = \frac{e^{\alpha_k X_i}}{\sum_k e^{\alpha_k X_i}}$
β	0.455	0.464	0.697	0.831
s.e. of β	(0.142)	(0.051)	(0.190)	(0.227)
Log-Likelihood	-8478.6	-8453.1	-8452.8	-8492.34
LR statistic for the restrictions imposed in Table 3	11.6	63.4	-	-
p-value	0.003	0.000	-	-
LR statistic for $\beta=1$	1.45	6.04	0.97	13.4
p-value	0.23	0.01	0.33	0.0002

Table 6a: Policy Evaluations[†]
Medium Wage Sample – Lognormal Model

	Unrestricted Model – Hyperbolic				Restricted Model – Exponential			
	Expected Duration (weeks)	Expected Wage (\$)	Gov't Expenditure (\$)	Long- Run Utility	Expected Duration (weeks)	Expected Wage (\$)	Gov't Expenditure (\$)	Long- Run Utility
Benchmark	22.43	259.36	2052	-	22.71	261.25	2024	-
Change in UI benefits								
In levels	-0.96	-0.83	-480	-245	-1.14	-1.31	-494	-393
in logs (×100)	-4.36	-0.32	-26.70	-	-5.13	-0.50	-28.01	-
Job search assistance								
in levels	-1.02	2.66	10,630	502	-1.43	2.82	25,790	440
in logs (×100)	-4.65	1.02	182.15	-	-6.52	1.07	262.06	-
Monitoring search effort								
in levels	-7.69	0.93	-360	485	-1.47	-0.20	-728	-696
in logs (×100)	-41.96	-0.20	-44.72	-	-6.69	-0.08	-44.72	-
Monitoring acceptance								
in levels	-0.61	-4.52	-442	-479	0.02	-7.12	-356	-427
in logs (×100)	-2.77	-1.76	-24.28	-	0.07	-2.76	-19.36	-
Employment bonus								
in levels	-2.12	-2.29	556	1092	-2.56	-3.52	543	768
in logs (×100)	-9.93	-0.89	23.97	-	-11.96	-1.36	23.78	-

[†] **Notes:** Entries in the Table represent changes (in levels and in logs) in the outcomes as a result of a given policy. The unrestricted model parameters are the maximum likelihood estimates in the medium wage sample. The restricted model parameters are the maximum likelihood estimates in the medium wage sample subject to the restriction that $\beta=1$. Source: Author's calculations.

Table 6b: Policy Evaluations [†]
Medium Wage Sample – Normal Model

	<u>Unrestricted Model –</u>				<u>Restricted Model – Exponential</u>			
	<u>Hyperbolic</u>							
	Expected Duration (weeks)	Expected Wage (\$)	Gov't Expenditure (\$)	Long- Run Utility	Expected Duration (weeks)	Expected Wage (\$)	Gov't Expenditure (\$)	Long- Run Utility
Benchmark	22.35	268.34	2052	-	22.31	268.46	2058	-
Change in UI benefits								
In levels	-2.36	-0.53	-619	-349	-2.24	-0.62	-610	-420
in logs (×100)	-11.14	-0.19	-35.70	-	-10.57	-0.23	-35.19	-
Job search assistance								
in levels	-1.20	0.66	8,682	203	-1.34	0.79	13,995	235
in logs (×100)	-5.53	0.25	165.12	-	-6.21	0.29	205.41	-
Monitoring search effort								
in levels	-8.50	0.80	-437	-212	-5.14	-0.26	-705	-707
in logs (×100)	-47.83	0.29	-23.81	-	-26.18	-0.10	-41.95	-
Monitoring acceptance								
in levels	-5.81	-3.04	-348	-501	-6.13	-3.65	-392	-790
in logs (×100)	-30.12	-1.14	-18.47	-	-32.10	-1.37	-21.16	-
Employment bonus								
in levels	-4.02	-1.39	470	892	-4.12	-1.62	462	743
in logs (×100)	-19.85	-0.52	20.56	-	-20.43	-0.61	20.26	-

[†] **Notes:** Entries in the Table represent changes (in levels and in logs) in the outcomes as a result of a given policy. The unrestricted model parameters are the maximum likelihood estimates in the medium wage sample. The restricted model parameters are the maximum likelihood estimates in the medium wage sample subject to the restriction that $\beta=1$. Source: Author's calculations.

Appendix Table 1: Methods of Search[†]

Method of Search	Summary Statistics				Maximum likelihood estimates	
	(1)	(2)	(3)	(4)	(5)	(6)
	Percentage using method	Average hours conditional on using method	Monthly probability of receiving offer	Implied weekly probability of receiving offer	Estimate of α	Estimate of c
State Employment Service	51.66	3.54	13.37	3.52	0.0309 (0.0062)	0.7183 (0.1864)
Friends and Relatives	70.63	3.61	15.57	4.14	0.0387 (0.0045)	0.8874 (0.1796)
Newspaper ads	65.54	4.01	9.28	2.41	0.0200 (0.0038)	0.4209 (0.1067)
Direct employer contact	62.72	4.77	17.46	4.68	0.0358 (0.0046)	0.6566 (0.1368)
Other methods	56.32	3.33	17.77	4.77	0.0409 (0.0068)	1* (-)
All Methods	100.00	7.17	32.93	9.50	-	-

[†] **Notes:** Summary statistics based on author's calculations using the 1981 Job Search questionnaire in the NLSY. Maximum likelihood estimates of α and c based on the model described in the Appendix.

* The cost coefficients are identified only up to a scale factor. The cost coefficient for "Other Methods" is therefore fixed at 1.

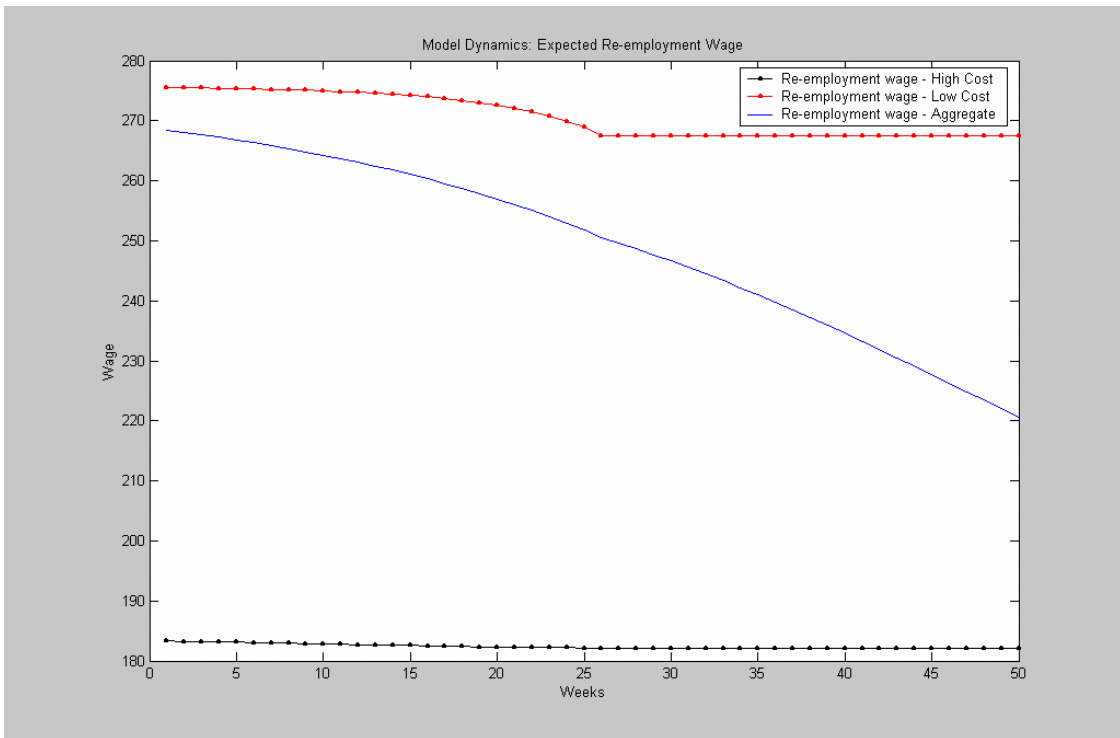
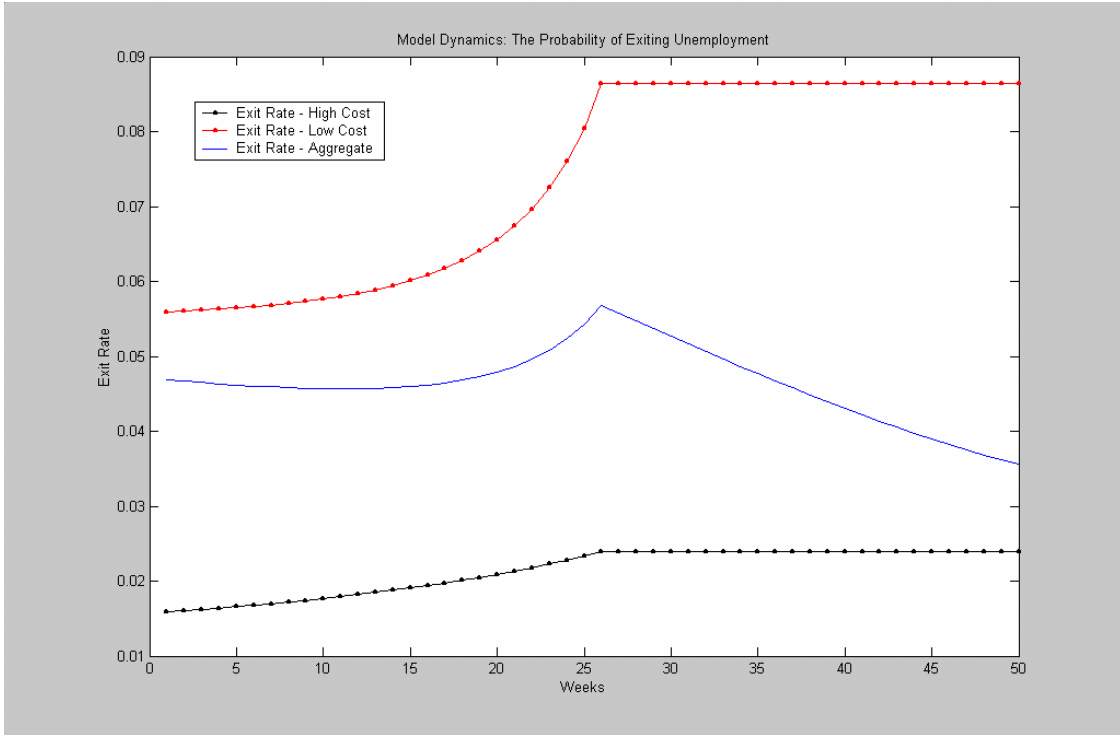


Figure 1: Model Dynamics

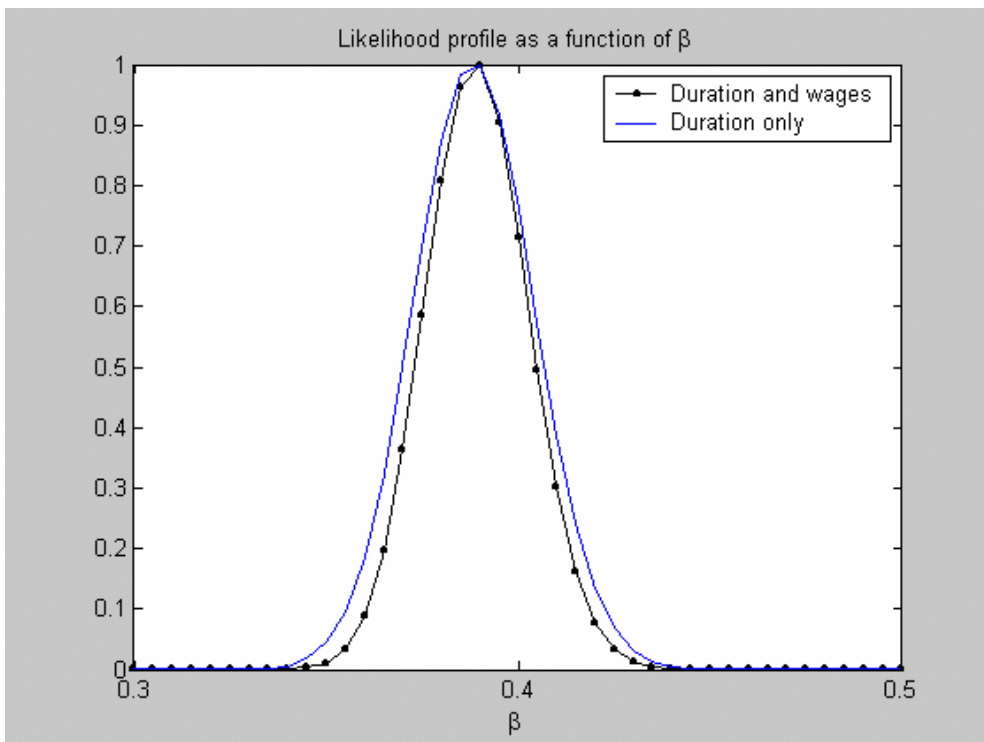
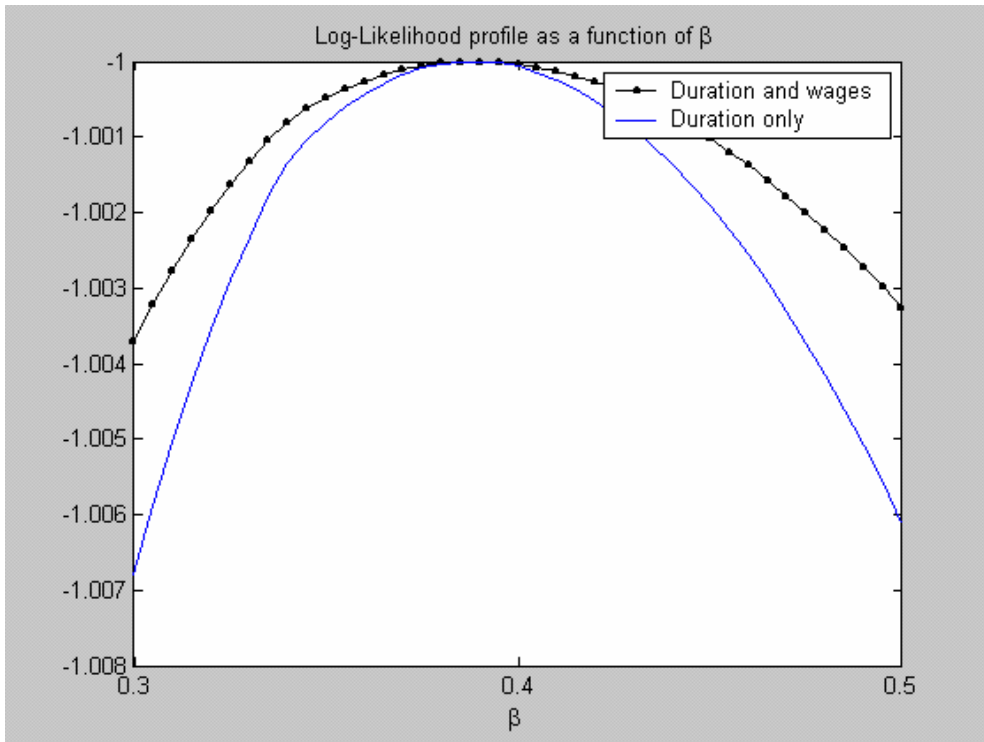


Figure 2a: Likelihood Profiles as a Function of β

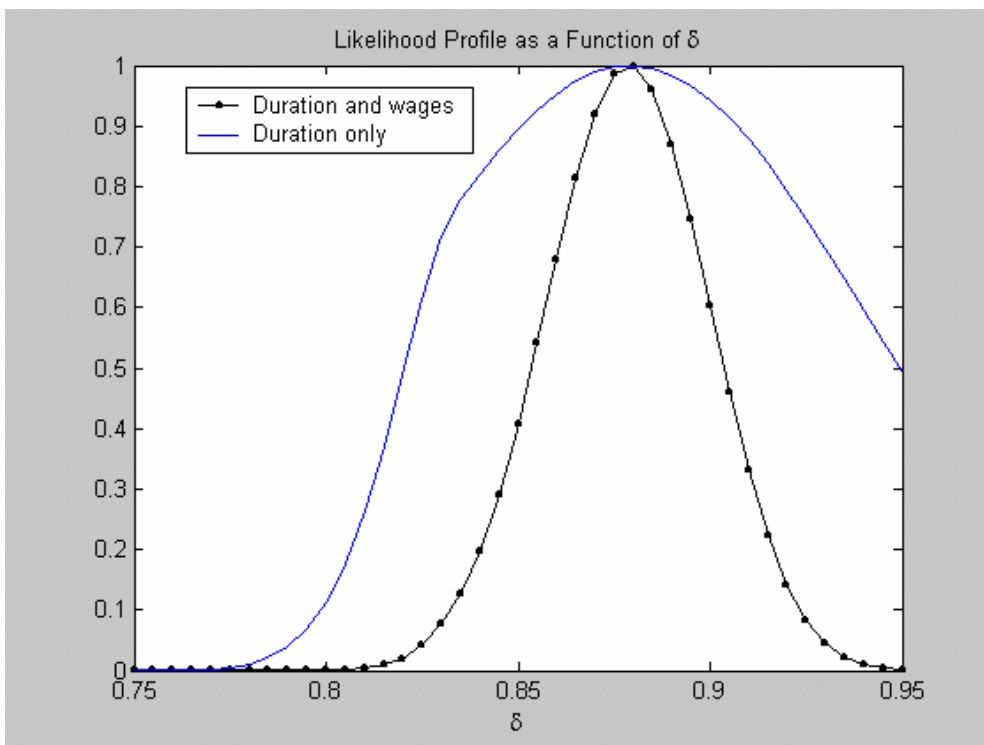
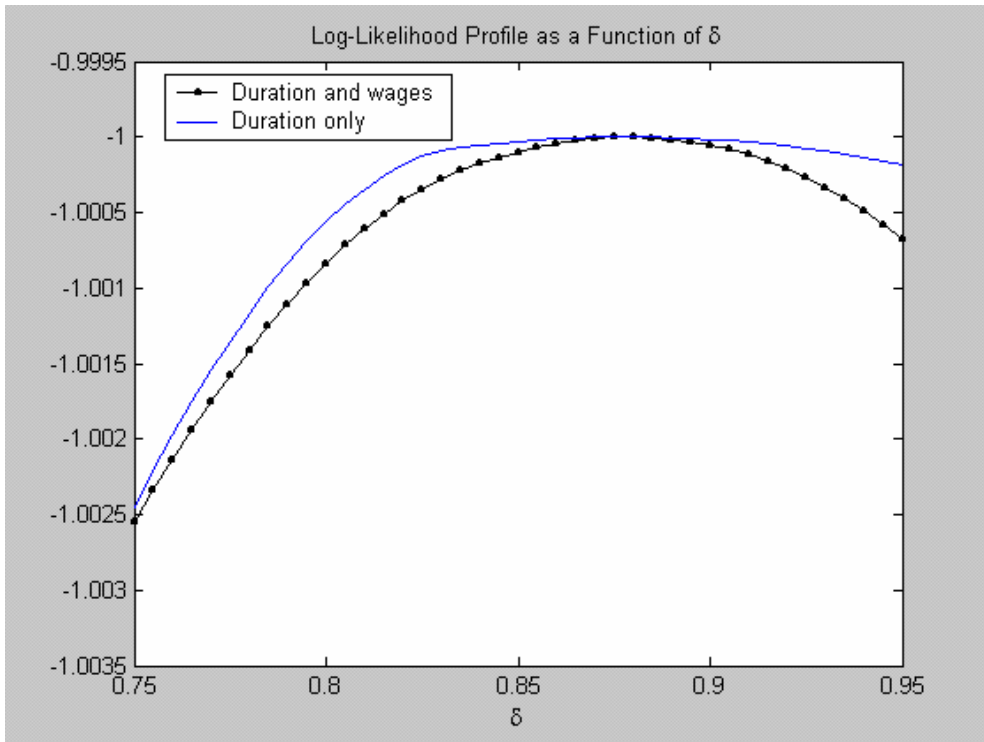


Figure 2b: Likelihood Profiles as a Function of δ

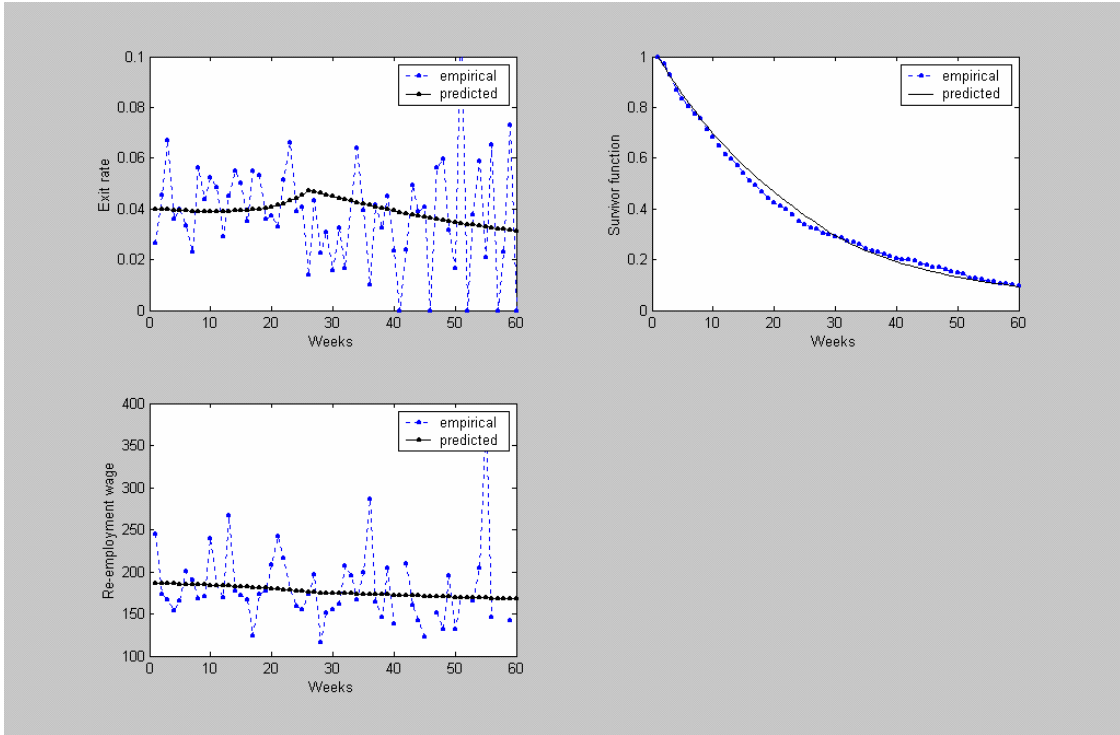
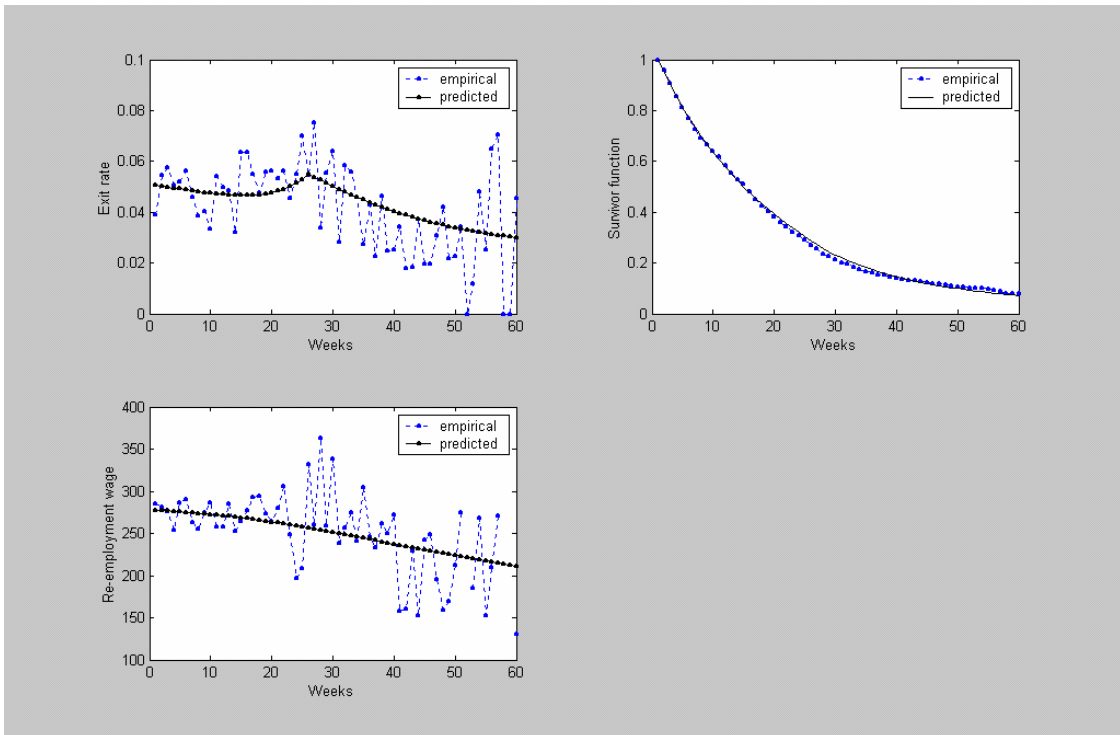


Figure 3a: Job Search Dynamics, Low Wage Sample



**Figure 3b: Job Search Dynamics, Medium Wage Sample
Lognormal Model**

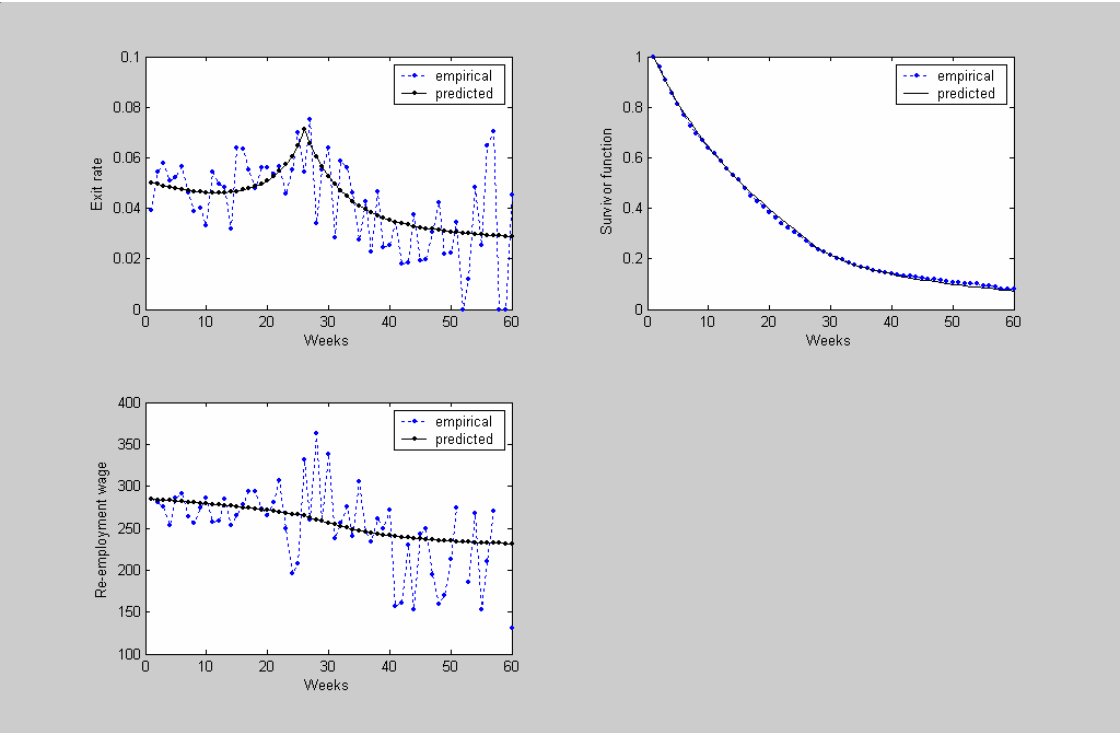


Figure 3c: Job Search Dynamics, Medium Wage Sample Normal Model

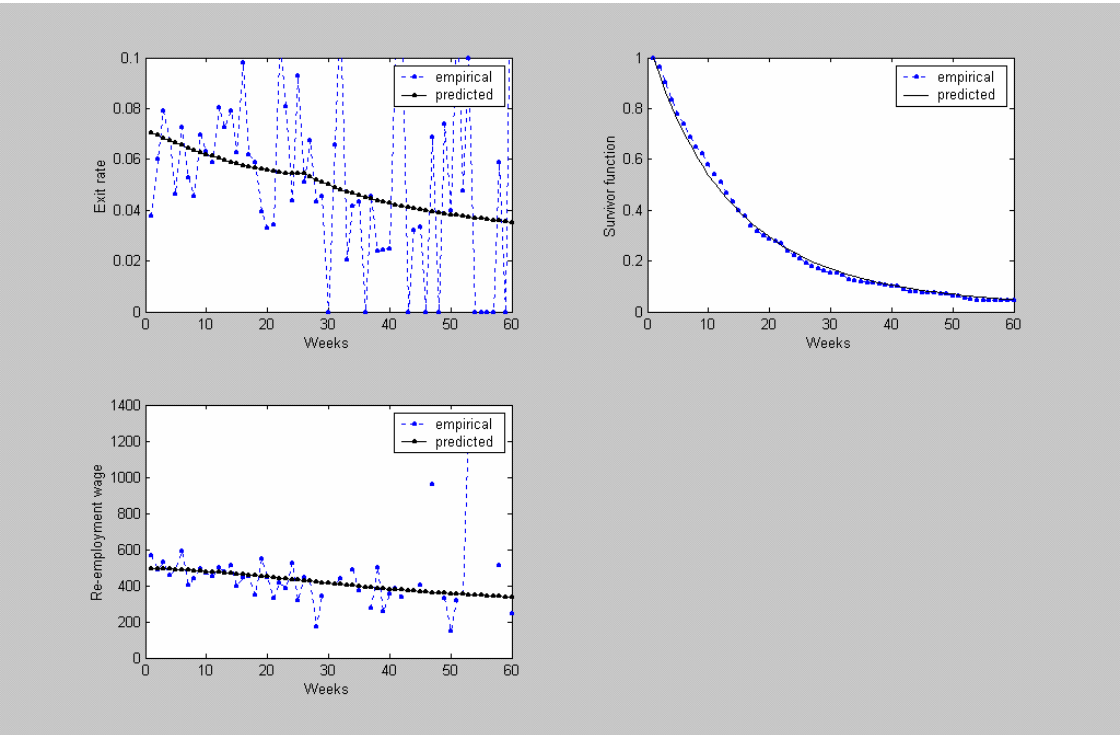


Figure 3d: Job Search Dynamics, High Wage Sample

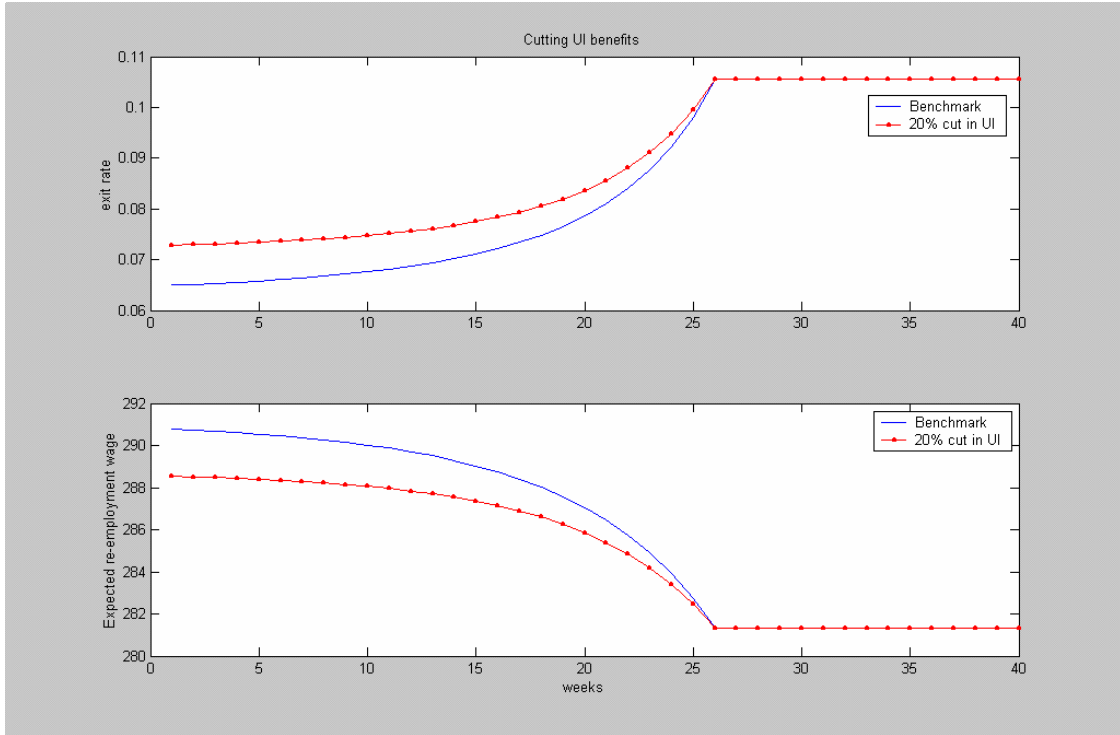


Figure 4a: Cutting UI Benefits

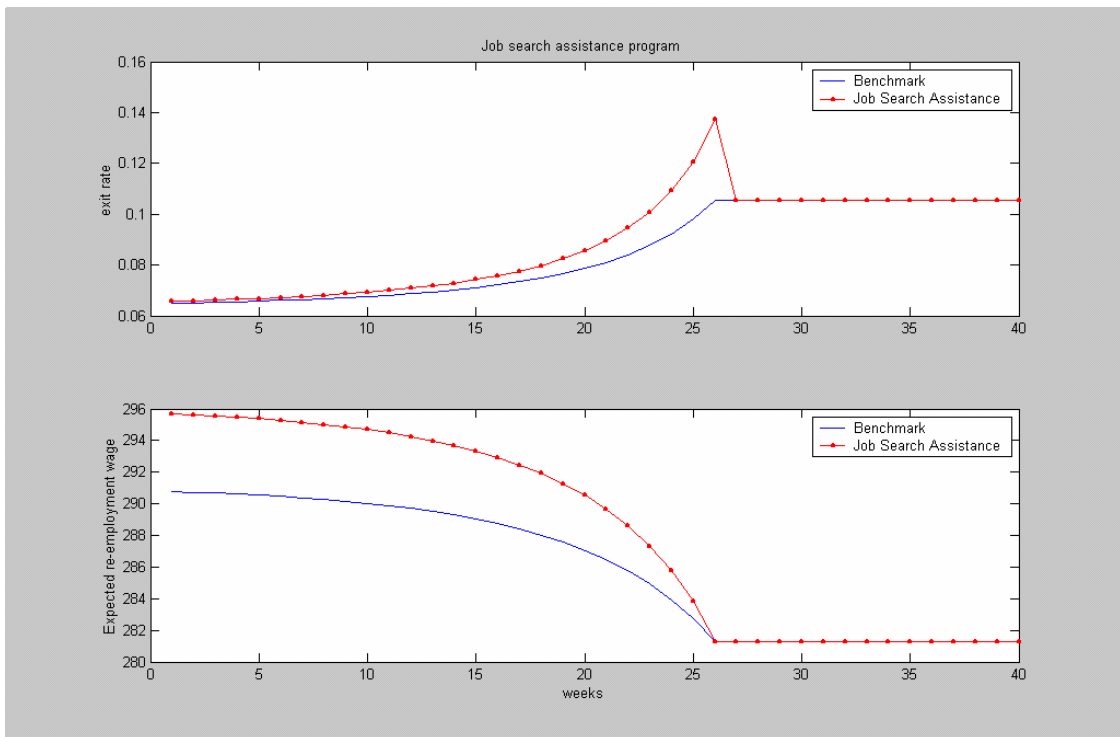


Figure 4b: Job Search Assistance Program

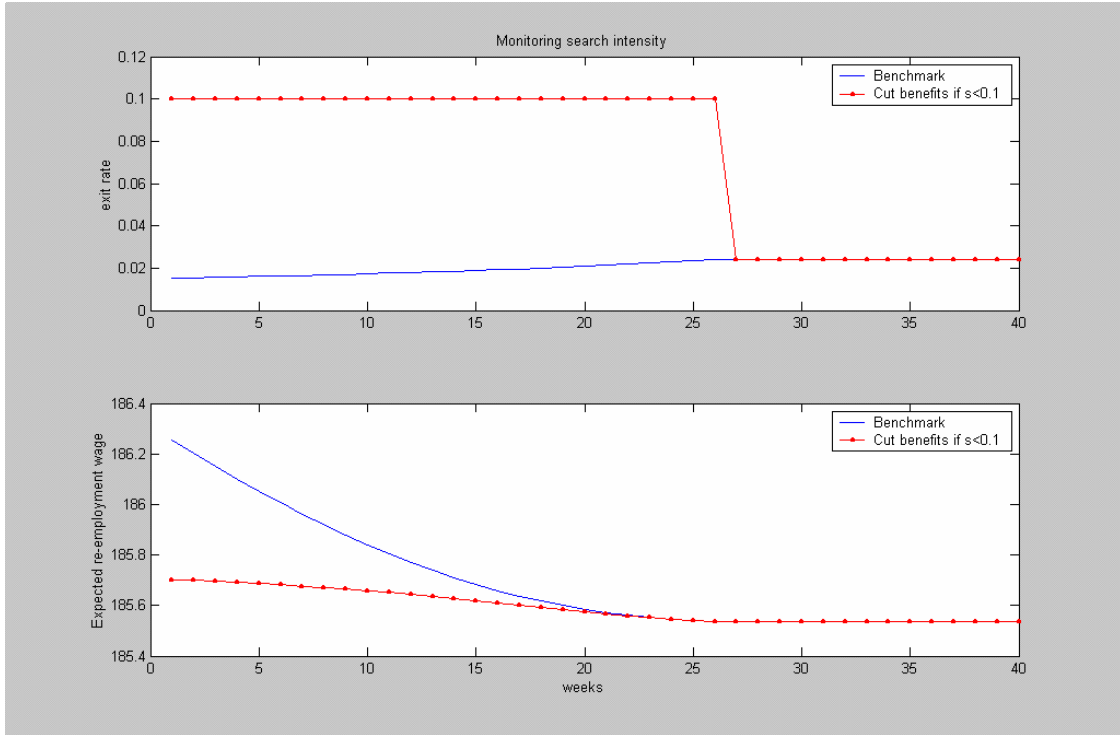


Figure 4c: Monitoring Search Intensity

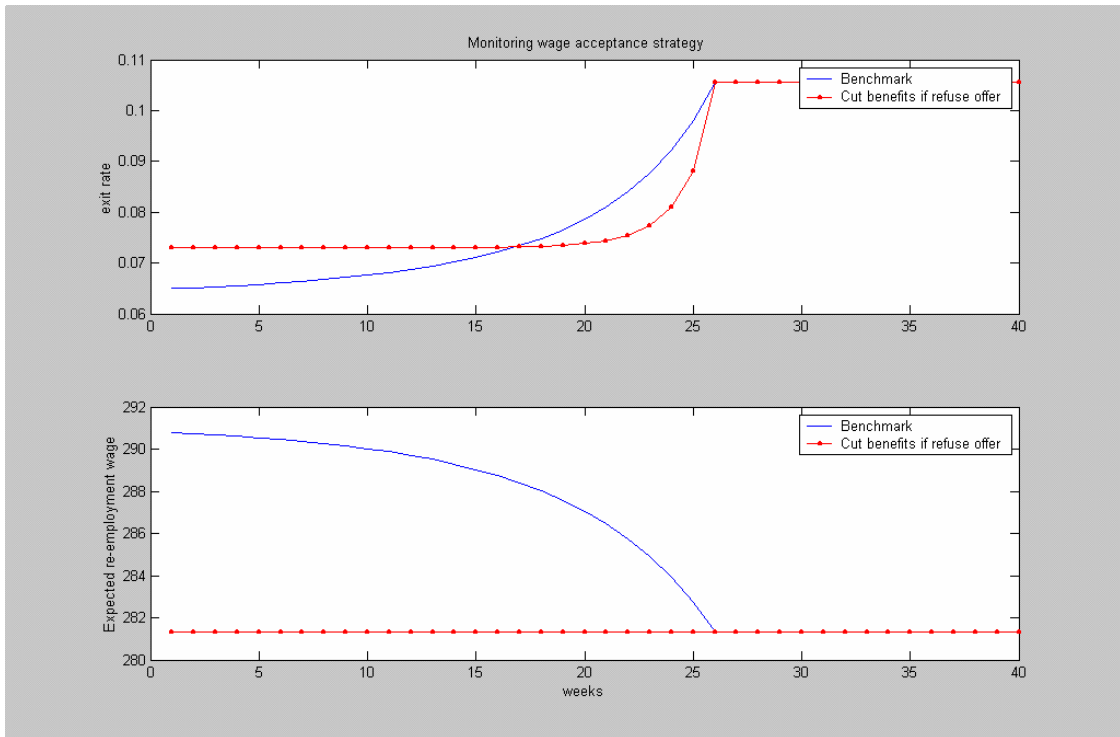


Figure 4d: Monitoring the Wage Acceptance Strategy

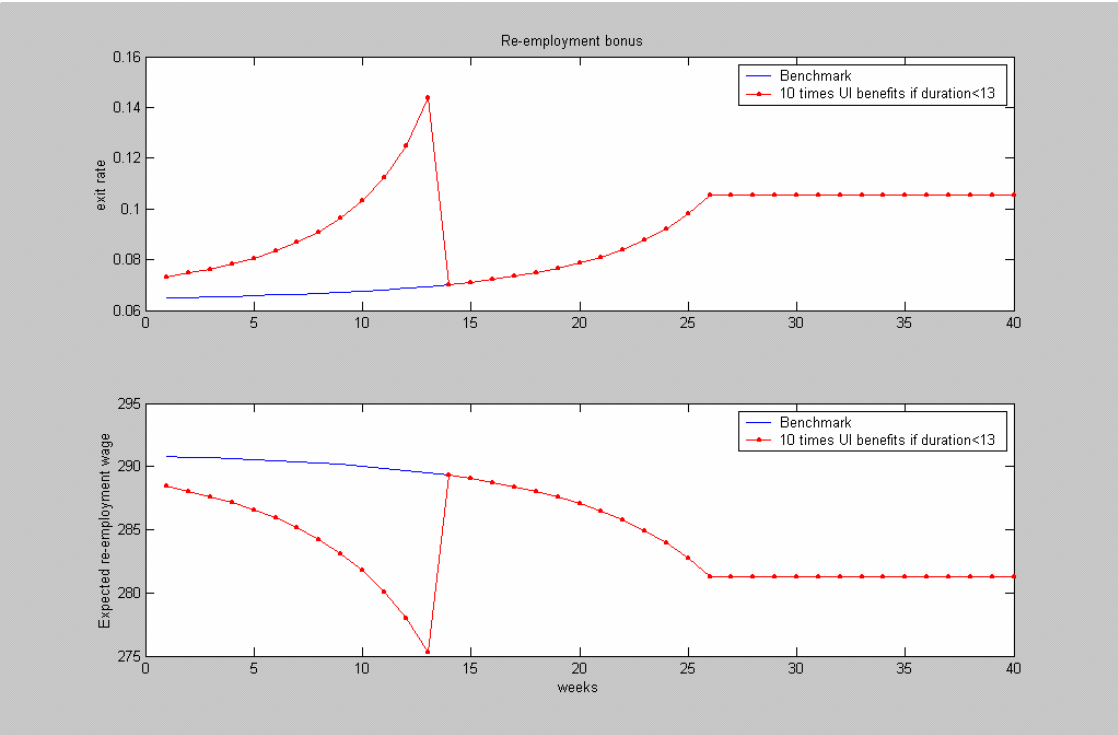


Figure 4e: Re-employment Bonus