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

July 2007

Abstract

This paper estimates quantitatively the degree of hyperbolic discounting in a job search model, using data on unemployment spells and accepted wages from the NLSY. The results point to a substantial degree of hyperbolic discounting for low and medium wage workers. The structural estimates are then used to evaluate alternative policy interventions aimed at reducing unemployment. The estimated effects of a given policy can vary by up to 40%, depending on the assumed type of time discounting. Some interventions may raise the long-run utility of hyperbolic workers, and at the same time reduce unemployment duration and lower government expenditures.

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

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

Pagehead Title: Job Search and Hyperbolic Discounting

**Forthcoming, The Economic Journal.* I thank three anonymous referees, the editor Marianne Bertrand, 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, Univeristy of Toulouse, UC Berkeley, 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.

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, 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 counselors[2]

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

[2] Job 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)

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). 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 unemployment 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]

[3] For 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 *et al.*, 2003, 2005) generates an observational non-equivalence between the hyperbolic and the exponential

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 *et al.*, 1998, 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 (2005, henceforth DV-P). The discount function has the conventional $\beta - \delta$ (or quasi-hyperbolic) form, where β is the degree of short run impatience, and δ is the traditional exponential discounting parameter. For a given set of parameter values, I solve the agent's dynamic optimization problem, calculate the likelihood of observing the actual realisations of unemployment durations and accepted wages in the National Longitudinal Survey of Youth (NLSY), and then iterate over the set of unknown parameters until convergence. The model is estimated separately for workers with low, medium and high wages prior to the unemployment spell, and takes into account observed heterogenity in cognitive test scores, marital status and race, as well as unobserved heterogeneity.

A key issue in the estimation of structural search models is the identification of the parameters of the model. The sources of identification are discussed at length in Section 2.3. As in

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.

[4] Exceptions include Angeletos *et al.* (2001), Fang and Silverman (2004), and Laibson *et al.* (2005).

many structural models, part of the identification does depend on the specific structure of the model. However, given the specific modeling restrictions, it is possible to identify the discounting parameters from specific features of the *data*. In particular, the key to identifying the short-run discounting parameter β is the magnitude of unemployment duration relative to the distribution of accepted wages. The intuition for this result lies in the fact that the short-run discounting parameter (approximately) affects only search effort, while it has no effect on reservation wages: in setting the reservation wage, the agent compares between different streams of payoffs to be received in the future, making the degree of short-run discounting (nearly) irrelevant for this decision.[5] This means that the distribution of accepted wages roughly conveys information only about parameters other than β. Therefore, for a given distribution of accepted wages, a relatively long average duration of unemployment must imply that workers exert little search effort, meaning that β must be relatively low. Conversely, if unemployment duration is relatively low, this must imply a high value of β. To assess how important the specific modeling restrictions are, I perform a number tests to assess the sensitivity of the results to different functional form assumptions.

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, while they are somewhat more sensitive to the assumed functional form of the offered wage distribution. In all specifications, the point estimate for the hyperbolic

[5] I say "approximately" because this statement is precise only for naïve hyperbolic workers (see DV-P, p. 536). For sophisticated hyperbolic workers (as in the empirical analysis in this paper), the reservation wage is affected by the future offer arrival rate, which depends on search intensity, and this clearly depends on β. However, this effect is only of second order importance.

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. I find that the estimated effects of a given policy can vary by up to 40%, depending on whether one assumes hyperbolic or exponential preferences. More dramatic are the welfare implications of some interventions: imposing benefit sanctions on workers who do not search enough can actually raise the long-run utility of hyperbolic workers, while at the same time reducing unemployment duration and government expenditures.

1. The Model

1.1. *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 $t = s+1, s+2, ...$, 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 $t = 0, 1, 2, ..., T$. From today's perspective, the plan from period s onwards yields utility

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

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

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

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 may 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 (1). 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:

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

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 exert higher effort in the present, knowing that in the future she will search less than what is optimal from today's perspective.[6]

1.2. *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, where the time unit can be thought of as a week.[7] Consider an infinitely lived

[6] A *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.

[7] This is mostly a matter of convenience, since the data available is weekly. However, there are many features of job search that do occur at a weekly frequency (for example, classified announcements in newspapers, or meeting your relatives on the weekend).

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.

Search effort s_t is parameterized as the probability of obtaining a job offer, hence $s_t \in [0,1]$. In every period the agent incurs a cost of search $c(s_t)$, an increasing and strictly convex function of s_t. 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 realisation 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. 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.

In every period of unemployment, the worker receives instantaneous utility b_t, which can be either positive (utility of leisure) or negative (stigma, low self-esteem, etc.). This utility varies over time, as it includes Unemployment Insurance (UI) benefits, which run out after a limited period. 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 to the presence of job seekers with hyperbolic preferences.

[8] The assumption of a fixed separation probability q is somewhat unrealistic. An attempt to estimate a model with the separation probability being a declining step function did not yield significantly different results.

1.3. *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

$$\max_{s_t \in [0,1]} b_t - c(s_t) + \beta\delta \left\{ s_t E_F \left[\max \left\{ V^E(w), V_{t+1}^U \right\} \right] + (1 - s_t) V_{t+1}^U \right\}, \tag{2}$$

where the expectation is taken with respect to the distribution of wage offers F. In words, 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. 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 $V^E(w)$[9] — 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

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

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 assume that once UI benefits are exhausted, (i.e., starting

[9] The continuation payoff from employment has no time subscript because it does not depend on the time one becomes employed.

from period $T+1$), the environment becomes stationary: in every period the worker chooses the same intensity of search effort and the same acceptance strategy. The equilibrium solution in the stationary model is characterized by a reservation wage policy.[10] Let w^*_{T+1} and s_{T+1} represent, respectively, the reservation wage and the optimal level of search effort in the stationary equilibrium. The reservation wage is:

$$w^*_{T+1} = [1 - \delta(1-q)]\, V^U_{T+1} - \delta q V^U_{new}.$$

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

$$\begin{aligned} (1-\delta)\, V^U_{T+1} &= b_{T+1} - c\left(s_{T+1}\right) + \frac{\delta s_{T+1}}{1-\delta(1-q)} Q\left(w^*_{T+1}\right); \qquad (3) \\ c'\left(s_{T+1}\right) &= \frac{\beta\delta}{1-\delta(1-q)} Q\left(w^*_{T+1}\right), \qquad (4) \end{aligned}$$

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 backward induction. The dynamic equations characterizing the solution in periods $t = 1, ..., T$ are :

$$\begin{aligned} w^*_t &= [1-\delta(1-q)]\, V^U_{t+1} - \delta q V^U_{new}; \qquad (5) \\ V^U_t - \delta V^U_{t+1} &= b_t - c\left(s_t\right) + \frac{\delta s_t}{1-\delta(1-q)} Q\left(w^*_t\right); \qquad (6) \\ c'\left(s_t\right) &= \frac{\beta\delta}{1-\delta(1-q)} Q\left(w^*_t\right). \qquad (7) \end{aligned}$$

To obtain a solution to this model, we need to specify V^U_{new}, the value of becoming unemployed after having been laid off. Given the features of the UI system, V^U_{new} should depend on the accepted wage and on the duration of the employment spell. However, this would introduce

[10] DV-P prove that the equilibrium in this game exists and is unique (Theorem 1, p. 535).

considerable difficulties in calculating the solution. Instead, I make the simplifying assumption that $V^U_{new} = V^U_{T+1}$. 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 role of the impatience parameters in the job search process. We showed that β operates mostly on search effort, while δ operates mainly on reservation wages. 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 immediate costs of searching and benefits that will start to materialise in the future, once an offer is accepted. This time span is relatively short: in the US 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

[11] I also experimented with two alternative assumptions: a) $q = 0$, i.e., the worker is employed forever. This yielded results qualitatively similar to the ones reported later, with the a lower value of the likelihood at the optimum; and b) $V^U_{new} = V^U_1$: 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^U_1), and the solution involves finding a fixed point to the system

$$V^U_1 = h\left(V^U_1\right)$$

where the function h is defined recursively by equations (3)-(7) with $V^U_{new} = V^U_1$. 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 $V^U_{new} = V^U_{T+1}$.

mainly affected by the degree of long-run discounting.[12] The distinct role played by β and δ in the job search process is important for identification issues. A model in which agents have a high degree of present bias will be distinguishable from one in which agents are exponential but highly impatient, because they will have different implications for the joint distribution of accepted wages and unemployment durations. 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.

2. 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). 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.[13] Finally, I incorporate unobserved heterogeneity in the form of a finite mixture

[12]This result does not depend on the timing of events in the job search process. In DV-P (Appendix D), we show that a continuous-time version of our model, in which wages are paid immediately upon receipt of an offer, 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.

[13]AFQT and marital status were the dimensions along which there was the most variation in exit rates and re-employment wages.

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 3.4 I consider alternative forms in which observed heterogeneity may affect the parameters of the model.

2.1. *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 $i = 1, 2, ..., N$. 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 the product of the probability of receiving an offer, $s_t(\theta, X_i)$, and the probability that this offer exceeds the reservation wage, $w_t^*(\theta, X_i)$. 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). 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. Assuming that the measured wage w_i^{obs} is a noisy measure of the true wage $\tilde{w}_i$, the log-likelihood function is:

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

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 (i.e., who are observed in the data when their unemployment spell is still ongoing); 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 $w_{T_i}^*$.

The dependency on observed individual characteristics is made explicit by decomposing θ into (θ_1, θ_2), and letting the exit rate be dependent on observed characteristics through a linear index function, $\theta_2' X_i$, so that $h_t(\theta, X_i) = \tilde{h}_t(\theta_1; \theta_2' X_i)$. Finally, one can add unobserved heterogeneity to the model by specifying a mixture distribution with discrete finite support for the parameter vector θ:

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

This general formulation allows potentially for unobserved heterogeneity in any one of the individual parameters. In practice, however, unobserved heterogeneity is allowed only for a subset of the parameters. In particular, the parameters of the discount function, β and δ are not allowed to vary by type: we will later argue that one of the keys for identification is that unobserved heterogeneity be present in some, but not all, the model parameters. The log-likelihood function for the model with unobserved heterogeneity becomes

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

2.2. *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.

The wage distribution. I assume 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 the log-normal case, the observed wage w_i^{obs} is equal to the actual wage times a multiplicative error term with a log-normal distribution; in the normal case, the observed wage is the sum of the actual wage and a normally distributed error term. These ditributional assumptions imply that it easy to derive the conditional density of observed wages, conditional on the actual wage being above the reservation wage, using properties of the bivariate normal distribution.[14] The parameters of the actual wage distribution, μ_i and σ_i are assumed to 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). The parameters μ_i, σ_i, and σ_u, the standard deviation of the error term, all need to be estimated from the data.

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

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

The parameter η represents the degree of convexity of the cost function. 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. It is difficult to assess exactly what should be an appropriate magnitude for the cost of search parameters. In the Appendix I show how one can use data on time spent on different methods of job search and on their effectiveness to estimate η. The estimated value of η is approximately 0.4. [15]

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

[14] The details of the calculation are stratightforward and can be found in Paserman (2004).

[15] One can 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 (in 1983 dollars), 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 $16,600$.

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 high cost of search workers, while type 2 workers are low cost workers. No restrictions are placed on $\Delta\mu$.[16]

2.3. *Identification*

Using the functional form specification described above, there are, within each income subgroup, 27 unknown parameters that need to be estimated; the short-run discounting parameter β and the long-run *yearly* discount factor δ;[17] the value of time when unemployed, b_0; the parameter vectors of the wage distribution $\boldsymbol{\mu}$ and $\boldsymbol{\sigma}$ (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 the Appendix. 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.

[16] It 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).

[17] The 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.

A key issue in the estimation of structural search models is the identification of the parameters of the model. The identification problem can best be understood if we consider first as a benchmark a population of homogeneous agents in a stationary search environment. In this case, all agents will have the same expected duration of unemployment and the same distribution of accepted wages, regardless of when they find a job. There are essentially two first moments available for identification of the model parameters: the mean duration of unemployment spells, and the mean of accepted wages. In this scenario, it is clear that at most two parameters can be identified without having to rely on the second moments of the distribution of duration and wages. Moreover, if the two parameters have roughly the same effect on reservation wages and on the search effort, identification may also be problematic.[18] Therefore, in a rich model such as the one presented here, identification depends on the availability of a sufficient number of moments, and on whether the different parameters play sufficiently distinct roles in the search process.

There are four key features of our model that allow identification of the key parameters of interest, and of the short-run discounting parameter in particular: 1) the special role played by the short-run discounting parameter in the search process; 2) exclusion restrictions generated by observed heterogeneity in some parameters of the model, but not in others; 3) the non-stationarity of the environment, which implies that there are additional moments available for identification; 4) the specific functional form assumptions and non-linearities present in the model. I will discuss these four features in turn.

The role of β in the search process. The short run discounting parameter β plays a

[18] Roughly speaking, we could think of a linearization of the first order conditions (3-4) that generates a system of two equations in two unknowns. For identification purposes, the coefficient matrix must not be singular.

special role in the search process, in that it affects mainly the search effort decision, a decision involving a trade-off between immediate costs and benefits that materialise 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). As a result, data on accepted wages could in theory be used to estimate the other model parameters, and, once these were known, the value of β could be backed out from data on unemployment duration. This discussion illustrates that the estimated value of β depends on the relative magnitude of unemployment duration and accepted wages. For a given distribution of accepted wages, a relatively long average duration of unemployment must imply that workers exert little search effort, meaning that β must be relatively low. Conversely, if unemployment duration is relatively low, this must imply a high value of β.

Heterogeneity and exclusion restrictions. In a stationary environment with one type of agent, there are two moments available for identification (the mean duration of unemployment and the mean of accepted wages). Hence, from the data it should be possible to identify no more than two parameters – say, the short run and the long run discount factors. A model with three unknown parameters (e.g., the two discounting parameters *and* the cost of search parameter k) would be underidentified. Suppose instead that there are two types of agents, whose types are observable, and who have different costs of search, but the same discounting parameters. The two types will have different durations of unemployment spells, and different re-employment wages. This means that there are now *four* moments available for the identification of the four unknown structural parameters (the two discounting parameters and the cost of search for the two different types). If the two types differed also in their discounting parameters, the model would contain 6 unknown parameters and would again be underidentified. In other words, observed heterogeneity in some of the parameters of the model, but not in others, generates exclusion restrictions that aid in identification

Non-stationarity. The environment in my model is non-stationary, as I assume that agents receive unemployment benefits, which are exhausted after a limited amount of time. To see how this can help in identification, consider for example the short run discounting parameter β and the value of being unemployed b_0. An increase in β and a decrease in b_0 both lead to an increase in search effort. *Prima facie*, it would seem difficult to identify the two parameters separately in a stationary environment. In a non-stationary environment, by contrast, the value of time when unemployed is $b_0 + b_{UI}$ when one receives benefits, and b_0 when benefits have run out. The parameter b_0 must tie together between search effort in the first and in the later part of the unemployment spell. Therefore it is the relative probabilities of exiting unemployment that allow separate identification of b_0 and β in the non-stationary environment. Unobserved heterogeneity also introduces nonstationarity in the observed time path of optimal reservation wages and search effort, as the composition of the sample varies over time. The nonstationarity assumption generates an additional set of moments that can be used for identification, since now one has to find a set of parameters that match the entire time path of wages and exit rates from unemployment.

Functional form assumptions. Identification hinges on the specific functional form assumptions and the nonlinearities present in the model. As illustrated clearly by Flinn and Heckman (1982), a key element is the functional form of the wage offer distribution. Since the distribution of accepted wages is a truncated distribution, it is impossible to distinguish between a model with a lot of mass in the wage below the reservation wage, from a model with a low arrival rate of job offers but little mass below the reservation wage. Therefore, it is necessary to make a specific functional form assumption for the wage offer distribution. In line with previous research on structural estimation, I experiment with the log-normal and the normal specifications for the distribution of wage offers and assess the sensitivity of the results to the different assumptions.[19]

Finally, it should be said that, while the above discussion hints at what the potential sources of identification are, in a highly nonlinear model such as this, whether the model is identified is ultimately an empirical question: are the confidence bands around the parameters of interest sufficiently tight that we can make meaningful economic statements? It turns out that the answer is yes: the standard error on the short-run discounting parameter is sufficiently small that one can reject the null hypothesis of exponential discounting in nearly all specifications.

3. Results

3.1. *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 labour 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

[19]Both the log-normal and the normal distributions satisfy the *recoverability* condition: it is possible to uniquely determine the untruncated distribution $F(x)$ from knowledge of the truncated distribution with known point of truncation, $F(x|x \geq x^*)$. (Flinn and Heckman, 1982).

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 labour 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.

3.2. *Structural Parameter Estimates*

In Table 2 I present the estimates and standard errors of the structural parameters in the three subsamples.[20] 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 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 (β between 0.40 and 0.48), whereas for high wage workers the degree of short run impatience is relatively small ($\beta = 0.89$). The estimate of β is markedly higher in the normal model (0.81), though still substantially lower than 1. At the bottom of the table I report the results of likelihood ratio tests for the restriction $\beta = 1$. In the low wage

[20] The model was estimated using a pre-conditioned conjugate gradient algorithm with numerical derivatives from the NAG Fortran library. To ensure that the optimization algorithm did not wander off testing nonsensical values, I applied a logistic transformation to all the model parameters, and maximized the likelihood with respect to these parameters, which can take on any value on the real line. Standard errors were obtained using the outer product gradient method. To ensure that the algorithm had in fact converged to a global maximum, the optimization procedure was repeated with different starting values.

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 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 labour market. It could also be the case that higher-income individuals have more access to self-control tools, such as social networks, that may help them overcome their self-control problems.

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 *et al.*, 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. Laibson *et al.*'s (2005) structural estimates of β, based on life-cycle consumption choices, range between 0.51 and 0.82. In a labour market context similar to mine, Fang and Silverman (2004) estimate a model of welfare participation for single mothers, and they estimate β equal to 0.34, with a fairly tight standard error. In their model, time-inconsistent behavior arises because of the tension between the short-run benefits of receiving welfare payments versus the delayed reward from working. As in this paper, identification is feasible because of the differential effects of short-run and long-run discounting on welfare participation, because of functional form restrictions, and because of variation in wages and welfare benefits that generate a sufficiently curved likelihood surface. Altogether, my point estimates are roughly comparable to those found elsewhere.

In all models, the likelihood surface in the long-run yearly discount factor δ is essentially flat, as reflected by the very large standard errors, meaning that the long run discount factor is not identified. In fact, the point estimates indicate that δ would take on values greater than 1 if it were not restricted to the unit interval. It should be stressed that the lack of identification of δ is not due to the presence of the additional parameter β: even when I fix β to be equal to 1, δ is not separately identified from the benefits of being unemployed (b_0) and from the costs of search (k). This is a feature common to many other structural models of job search: previous studies, which were not interested specifically in the discount function, simply fix δ at a prespecified value.[21]

We find that the value of time when unemployed b_0 is substantially negative in most specifications, consistent with previous structural estimates of the job search model, and also with the literature on subjective measures of happiness, which universally finds that unemployment *per se* bears negative utility (Oswald, 1997; Winkelmann and Winkelmann, 1998). The disutility of unemployment 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 ob-

[21] It should be noted that Fang and Silverman were able to obtain a fairly precise estimate of the long run discount factor δ, even when they set $\beta = 1$. This underlines that the welfare participation models analyzed by Fang and Silverman are different from the job search models in some ways that we probably do not fully understand.

served wages in all population subgroups, indicating that a sizeable fraction of wages are rejected. In all specifications, the costs of search is highest for unmarried workers with low AFQT and lowest for married workers with high AFQT. It is conceivable that low-skill and unmarried workers face a higher cost of obtaining a job offer, the latter perhaps because they have a smaller network of acquaintances. The pattern of the parameters mirrors 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.

3.3. *Predicted Values*

Figures 1a and 1b show the actual and predicted survivor functions, hazard functions, and the average re-employment wage as a function of duration, for the two specifications in the medium wage sample.[22] The model does a good job at fitting the survivor function, even though it cannot capture the high week to week variation in exit rates and expected re-employment wages. In both specifications, the exit rate rises as week 26 (the time of benefit exhaustion) approaches, though the normal model seems to be able to fit the spike in exit rates at the time of benefit exhaustion more accurately.

In Table 3, I present the predicted values for a number of variables at different stages of the unemployment spell. In all specifications, the predicted survivor function matches quite well the empirical survivor function. The model also does well in matching the average re-employment wage, while it slightly overestimates the average duration of spells (perhaps because of right-censoring of extremely long spells in the actual data). Underlying these averages, there is a rich

[22] In the low and high wage samples we found similar patterns. The figures are available upon request.

dynamic behavior. Offer probabilities decrease monotonically in three of the four specifications (the changing sample composition between high and low-cost workers dominates the increase in search effort as UI benefits run out), while offer acceptance probabilities rise monotonically in all models. As a result, exit rates can have either a U-shape or can decrease monotonically in the first 26 weeks.

In the lognormal model, predicted acceptance probabilities are quite low during the first few weeks of the spell, ranging from 12% to 20%. In the low and medium wage samples, the proportion of acceptable offers rises to 34%-50% after one year, and to 77% to 90% 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 acceptance probability ranges between 55% and 73% over the entire spell, matching quite closely direct estimates of acceptance probabilities – on the order of 60% – based on survey evidence (Blau and Robins, 1990).[23]

In comparing between different types of workers, the table suggests 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

[23] Estimates from previous one-sided structural models found that essentially all offers are accepted (Wolpin, 1987, Van den Berg, 1990). On the other hand, some structural equilibrium search models also find low values of the acceptance probability (Eckstein and Wolpin, 1995).

in reservation wages (see Devine and Kiefer, 1991).

3.4. *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 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.[24]

The results are presented in Table 4. I focus on the medium wage sample and the lognormal

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

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.[25]

[25] Additional robustness checks involved the cost of search function and the time profile of unemployment 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.

4. 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.[26] Having fully specified the agent's preferences, 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, while the same policy would lower the welfare of an agent with time-consistent preferences.

In this section, I first use the maximum likelihood estimates of Table 2 to evaluate the effect of a number of different policy interventions unemployment duration, on the re-employment wage, on government expenditures and on the individual's long-run welfare. I then assess whether and how the conclusions would have been altered if I had instead estimated the model assuming that the agent's preferences are exponential.

With hyperbolic preferences, the correct notion of individual utility is difficult to define, because of the potential conflict between an individual's different selves: 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

[26] However, 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.

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.

4.1. *Policies*

Following is the decsription of the seven different policies under examination. These are modeled to resemble as closely as possible interventions that are, or have been, actually implemented. I restrict attention to the medium wage sample.

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%.

Shortening the duration of unemployment benefits. In this intervention, UI benefits are paid for only 21 weeks instead of 26.

Changing the time profile of unemployment benefits. Instead of paying a constant stream of UI benefits throughout the unemployment spell, one can change the time profile of benefits. Here I experiment with increasing the benefit payments by 33% in the first 13 weeks of the unemployment spell, and then cutting the payments in half in the subsequent 13 weeks of the spell, so that the total amount of payments to a worker who remains unemployed until the exhaustion of benefits stays unchanged. It is worth examining whether exponential and hyperbolic workers react differently to the intertemporal tradeoffs embodied in this policy.

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% 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%, and that government expenditures rise by $0.1k$ for every week of insured unemployment.[27]

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 *et al.*, 2005; van den Berg *et al.*, 2004; Abbring *et al.*, 2005). 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 $\underline{s} = 0.05$, the worker is no longer eligible to receive UI benefits.

[27] It may be argued that this intervention is not really feasible given the way I have modelled the search cost as a psychological cost. This interpretation is not strictly necessary, and part of the costs of search can be viewed as direct monetary costs. Moreover, the simulations should be viewed primarily as a modeling device that allows me to quantify approximately the effect of an intervention that is often used in practice, even though it does not fit neatly into the framework of the job search model presented here. The same holds for other policies analyzed below, such as monitoring search intensity or monitoring the job acceptance strategy.

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. 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 where able to hold that job for 13 more weeks. Translating this policy to our model is straightforward.

Finally, all the policies are compared to the hypothetical environment in which perfect commitment is assumed. This allows us to obtain a sense of the limits of these interventions that provide only partial commitment.

4.2. *Results*

I present the results of the policy evaluations in Table 5. The top panel refers to the lognormal model, and the bottom panel to the normal model. The first four columns of the table present the effects of the policies in the hyperbolic model. All the parameter values were set at their maximum likelihood estimates, with the exception of δ, which was set at 0.95 to avoid implausible outcomes resulting from the infinite horizon and the absence of discounting.[28] The last four columns present the effects of the policies using parameter estimates from a restricted model, where β is set equal to 1. 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). All the outcomes are population averages, where the average is taken with respect to the distribution of observed and unobserved types.[29,30]

I analyze first the results for the lognormal model. A 20% 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% because of agents' behavioral response to the policy, with little difference between the

[28] As stated before, the likelihood function is essentially flat in δ, hence fixing δ at 0.95 does not have any meaningful effects on any of the results. Maximum likelihood estimates of models with $\delta = 0.95$ are available upon request.

[29] For details of the calculations, see Paserman (2004).

[30] Since utility here is measured in monetary terms, it is possible to report the change in utility averaged over the different types of workers in the population. In Table 7 I highlight how different workers experience difference changes in utility following a policy intervention.

hyperbolic and the exponential model. The utility drop for the hyperbolic agent is somewhat smaller than that for the exponential agent.

In the hyperbolic model, reducing the duration of UI benefits by 5 weeks has a slightly smaller effect on duration than cutting the level of benefits, while the two policies have nearly identical effects in the exponential model. In both models, though, cutting the duration of benefits achieves a smaller reduction in government expenditures than cutting the level of benefits. Changing the time profile of UI benefits has essentially no effect on duration in the hyperbolic model, and only a small effect in the exponential model.

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, in contrast to the empirical findings on job search assistance programs (Meyer, 1995). 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, if the program shifts the distribution of potential wage offers, the cost-benefit analysis could be much more favorable.

Monitoring search intensity results in large decreases in unemployment duration for both hyperbolic and exponential agents, and a 9%-12% decrease in government expenditures. What is most striking about this policy is the contrast between hyperbolic and exponential agents in terms of welfare: despite the drastic reduction in unemployment duration, the policy actually has a positive effect on the hyperbolic worker's long-run utility, and he would be willing to pay 460 dollars to see it implemented. The exponential worker, on the other hand, would be willing to pay a nearly identical amount to *prevent* implementation of this policy. These results are not unexpected: 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.[31]

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 0.63 weeks in the hyperbolic model, and is essentially unchanged in the exponential model. The expected re-employment wage decreases by at most 7 dollars.

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 more than 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.

[31] The magnitude of the effect is somewhat larger than that found in the literature (Ashenfelter *et al.*, 2005; van den Berg *et al.*, 2004; Abbring *et al.*, 2005) but this can be explained by the fact that we are assuming that the government is capable of monitoring search effort perfectly. I have also experimented with a specification with imperfect monitoring, and the results of the analysis were qualitatively similar.

The final row shows that all the policies fall far short of achieving the results that would have been obtained by a hyperbolic worker with perfect ability to commit. In this sense, the proposed interventions provide a very limited partial commitment device.

The bottom panel of the table 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 and changing the time profile of UI benefits have a similar impact on duration. On the other hand, the impact of a change in the level or duration of 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. Again, the hyperbolic worker would like to see this policy implemented, even though the change in utility is smaller than that found previously. The gap between the effects of the policies and the effect of perfect commitment ability is now substantially smaller. Some of the interventions are able to cut unemployment duration by as much as two thirds of the reduction obtained by prefect commitment.

4.3. *Designing an Optimal Policy*

The results above suggest that with hyperbolic workers it may be possible to design a policy that improves the worker's long-run utility, reduces unemployment duration, and lowers government expenditures. Such a policy takes the form of monitoring search effort, and imposing sanctions on workers who do not meet the effort threshold. In Table 6, I examine how varying the effort threshold affects the outcome variables. It turns out that in the lognormal model, setting the effort threshold at 0.05 achieves the largest reduction in unemployment duration, and maximizes long run utility and social welfare (the change in the worker's utility plus the reduction in government expenditures). In the normal model, utility maximization is achieved at $\underline{s} = 0.04$, while expected duration is minimized at $\underline{s} = 0.10$, and social welfare increases monotonically in the threshold (the reduction in government expenditures always exceeds the drop in workers' long run utility).

In Table 7, I investigate more in depth how the policy affects the behavior of different types of workers, for two values of the effort threshold, 0.05 and 0.10. In the log-normal model, setting the threshold at 0.05 induces all the high cost workers to exert effort exactly at the threshold so as not to lose UI benefits, while the behavior of low cost workers is unaffected. The deviation from the unrestricted behavior raises the long-run utility of all the high cost workers: hence, the policy is Pareto-improving in the interpersonal sense.[32] If the threshold is set at 0.10, all the high cost workers choose instead to opt out of the UI system, and their long-run utility falls relative to the benchmark. In the normal model, all the high-cost workers and one low-cost worker search exactly at the threshold when this is equal to 0.05. Again, the policy is Pareto-improving. On the other hand, when the threshold is 0.10, there is a rich variety of outcomes: one type of worker chooses to opt out of the UI system (unmarried, low AFQT, high cost); the remaining high-cost types choose to search at the threshold and experience a drop in long-run utility; two of the low-cost types choose to search at the threshold, and their long-run utility rises; and the behavior of the remaining workers is unaffected.

This exercise illustrates how a given policy can have a rich set of implications for different types of workers. Importantly, with hyperbolic preferences it is possible to raise the utility of all types

[32] We do not perform here an intrapersonal welfare analysis for the different selves of the hyperbolic agent.

of agents, while at the same time reducing unemployment and lowering government expenditures.

5. Conclusion

This paper estimates the structural parameters of a model of job search with hyperbolic discounting and endogenous search effort, using data on duration of unemployment spells and accepted wages from the NLSY. Identification of the hyperbolic discounting parameters is achieved primarily thanks to variation in the relative magnitude of unemployment duration and accepted wages, even though the results depend also on the specific structure of the model and on the functional form assumption for the distribution of offered wages. This is a somewhat undesirable feature of the model: perhaps more data on self-reported search behavior and expectations about offered wages would enable one to pin down more precisely the parameters of the discount function.

In three different subsamples, the short run discounting parameter is smaller than one, although the point estimate is sensitive to the assumed functional form of the wage offer distribution. In most specifications, likelihood ratio tests reject the exponential model. 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.

Appendix: 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

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

Letting λ be the Lagrange multiplier, ($\lambda \geq 0$), the first order conditions are:

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

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

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

$$P(X_1,...,X_K) = 1 - \prod_{j=1}^{K} [1 - f(X_j)]. \tag{8}$$

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

$$f(X_j) = 1 - (1 + X_j)^{-\alpha_j}, \quad \alpha_j > 0. \tag{9}$$

It is easy to verify that this functional form satisfies all the necessary requirements. Using (8), we can express the first order conditions as

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

and, using (9), we obtain

$$\frac{c_j}{c_k} = \frac{\alpha_j}{\alpha_k} \frac{(1 + X_k)}{(1 + X_j)}, \tag{10}$$

which yields

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

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; $S_i = \left\{ s | \sum_{j=1}^{K} 1(X_j(s) > 0) = i \right\}$ 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

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

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.[33] 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.[34]

[33] This is the same data used in Holzer (1988).

[34] These 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 servces 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.

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

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

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

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

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 $c_1, ..., c_{K-1}$ (c_K is normalised to 1) from the first order conditions (10):

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

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, $\alpha_1, ..., \alpha_5$, and for the cost coefficients $c_1, ..., c_5$.

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.[35] The results of this regression are

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

[35] The 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.

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.

References

Ainslie, G. W. (1992). *Picoeconomics.* Cambridge: Cambridge University Press.

Abbring, J. H., van den Berg, G. J. and van Ours, J. C. (2005). 'The effect of unemployment insurance sanctions on the transition rate from unemployment to employment', Economic Journal, vol. 115(505) (July), pp. 602–30.

Angeletos, G-M., Laibson, D. I., Repetto, A., Tobacman, J. and Weinberg, S. (2001). 'The hyperbolic buffer stock model: calibration, simulation, and empirical evaluation', *Journal of Economic Perspectives*, vol. 15(3) (Summer), pp. 47–68.

Ashenfelter, O., Ashmore, D. and Deschênes, O. (2005). 'Do unemployment insurance recipients actively seek work? Randomized trials in four U.S. states', *Journal of Econometrics*, vol. 125(1-2) (March-April), pp. 53–75.

Barron, J. M. and Mellow, W. (1979). 'Search effort in the labor market', *Journal of Human Resources*, vol. 14(3) (Summer), pp. 389–404.

Benzion, U., Rapoport, A. and Yagil, J. (1989). 'Discount rates inferred from decisions: an experimental study', *Management Science*, vol. 35(3) (March), pp. 270–84.

Blau, D. M. and Robins, P. K. (1990). 'Job search outcomes for the employed and unemployed', *Journal of Political Economy* vol. 98(3) (June), pp. 637–55.

Bolles, R. N. (2000). *What Color Is Your Parachute? 2000. A Practical Manual for Job-Hunters and Career-Changers*. Berkeley: Ten Speed Press.

DellaVigna, S. and Paserman, M. D. (2005). 'Job search and impatience', *Journal of Labor Economics*, vol. 23(3) (July), pp. 527–88.

Devine, T. J. and Kiefer, N. M. (1991). *Empirical Labor Economics: The Search Approach.* New York and Oxford: Oxford University Press, 1991.

Diamond, P. and Kőszegi, B. (2003). 'Quasi-hyperbolic discounting and retirement', *Journal of Public Economics*, vol. 87(9-10) (September), pp. 1839–72.

Eckstein, Z. and Wolpin, K. I. (1995). 'Duration to first job and the return to schooling: estimates from a search-matching model', *Review of Economic Studies*, vol. 62(2) (April), pp. 263–86.

Fang, H. and Silverman, D. (2007). "Time-inconsistency and welfare program participation: evidence from the NLSY." Mimeo., Yale University, March 2007.

Flinn, C. J. and Heckman, J. J. (1982). 'New methods for analyzing structural models of labor force dynamics', *Journal of Econometrics*, vol. 18(1), pp. 115–68.

Gruber, J. and Kőszegi, B. (2001). 'Is addiction 'rational'? Theory and evidence', *Quarterly Journal of Economics*, vol. 116(4) (November), pp. 1261–1304.

Harris, C. and Laibson, D. I. (2001) 'Dynamic choices of hyperbolic consumers', *Econometrica*, vol. 69(4) (July), pp. 935–57.

Holzer, H. J. (1988). 'Search method use by unemployed youth', *Journal of Labor Economics*, vol. 6(1) (January), pp. 1–20.

Kirby, K. N. (1997). 'Bidding on the future: evidence against normative discounting of delayed rewards', *Journal of Experimental Psychology: General,* vol. 126(1) (March), pp. 54–70.

Kirby, K. N. and Herrnstein, R. J. (1995). 'Preference reversals due to myopic discounting of delayed reward', *Psychological Science*, vol. 6(2) (March), pp. 83–89.

Laibson, D. I. (1997). 'Golden eggs and hyperbolic discounting', *Quarterly Journal of Economics*, vol. 112(2) (May), pp. 443–77.

Laibson, D. I., Repetto, A. and Tobacman, J. (1998). 'Self control and saving for retirement', *Brookings Papers on Economic Activity*, vol. 0(1), pp. 91–172.

Laibson, D. I., Repetto, A. and Tobacman, J. (2003). 'A debt puzzle', in (P.Aghion *et al.*, eds.), *Knowledge, Information and Expectations in Modern Economics: In Honor of Edmund S. Phelps*, pp. 228–66, Princeton: Princeton University Press.

Laibson, D. I, Repetto, A. and Tobacman, J. (2005). 'Estimating discount functions with consumption choices over the lifecycle', mimeo., Harvard University.

Lippman, S. A. and McCall, J. J. (1976). 'The economics of job search: a survey', *Economic Inquiry*, vol. 14(1) (September), pp. 155–89.

Meyer, B. D. (1995). 'Lessons from the U.S. unemployment insurance experiments', *Journal of Economic Literature*, vol. 33(1) (March), pp. 91–131.

O'Donoghue, T. D. and Rabin, M. (1999). 'Doing it now or later', *American Economic Review*, vol. 89(1) (March), pp. 103–24.

O'Donoghue, T. D. and Rabin, M. (2001). 'Choice and procrastination', *Quarterly Journal of Economics*, vol. 126(1) (February), pp. 121–60.

Oswald, A. J. (1997). 'Happiness and economic performance', ECONOMIC JOURNAL, vol. 107(445) (November), pp. 1815–31.

Paserman, M. D. (2004). 'Job search and hyperbolic discounting: structural estimation and policy evaluation', CEPR Discussion Paper No. 4396.

Phelps, E. S. and Pollak, R.A. (1968). 'On second-best national saving and game-equilibrium growth', *Review of Economic Studies,* vol. 35(2) (April), pp. 185–99.

Thaler, R. H. (1981). 'Some empirical evidence on dynamic inconsistency', *Economics Letters,* vol. 8, pp. 201–07.

van den Berg, G. J. (1990). 'Nonstationarity in job search theory', *Review of Economic Studies*, vol. 57(2) (April), pp. 255–77.

van den Berg, G. J., van der Klaauw, B. and van Ours, J. C. (2004). 'Punitive sanctions and the transition rate from welfare to work', *Journal of Labor Economics*, vol. 22(1) (January), pp. 211-41.

Winkelmann, L. and Winkelmann, R. (1998). 'Why are the unemployed so unhappy? Evidence from panel data', *Economica*, vol. 65(257) (February), pp. 1–15.

Wolpin, K. I. (1987). 'Estimating a structural search model: the transition from school to work', *Econometrica,* . 55(4) (July), pp. 801–17.

Woodbury, S. A. and Spiegelman, R. G. (1987). 'Bonuses to workers and employers to reduce unemployment: randomized trials in Illinois', *American Economic Review,* vol. 77(4) (September), pp. 513–30.

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**	**0.4833**	**0.8140**	**0.8937**
	(0.1075)	**(0.1971)**	**(0.1672)**	**(0.1441)**
δ	0.9962	1.0000*	1.0000*	0.9989
	(0.1848)	(0.0001)	(0.0019)	(0.1798)
Value of time when unemployed				
b_0	-141.61	-164.31	-7.38	-308.78
	(61.16)	(61.43)	(16.54)	(193.53)
Parameters of the wage offer distribution				
μ_1 (unmarried, low AFQT)	4.1545	5.0230	174.08	5.4296
	(0.1996)	(0.1547)	(17.69)	(0.3408)
μ_2 (unmarried, medium AFQT)	4.3613	5.0993	185.64	5.2297
	(0.1789)	(0.1230)	(11.32)	(0.1331)
μ_3 (unmarried, high AFQT)	4.9503	4.8658	163.47	4.5543
	(0.2341)	(0.1425)	(13.62)	(0.1785)
μ_4 (married, low AFQT)	4.3954	5.1694	178.88	5.6692
	(0.2187)	(0.1192)	(15.75)	(0.1956)
μ_5 (married, medium AFQT)	4.2771	5.2790	204.97	5.4365
	(0.1337)	(0.1150)	(10.98)	(0.1361)
μ_6 (married, high AFQT)	4.1175	5.0735	232.06	5.3669
	(0.2683)	(0.0837)	(14.05)	(0.1235)
σ_1 (unmarried, low AFQT)	0.7196	0.5614	153.85	0.7444
	(0.0723)	(0.0585)	(8.31)	(0.1659)
σ_2 (unmarried, medium AFQT)	0.5760	0.4122	122.06	0.5810
	(0.0516)	(0.0453)	(7.54)	(0.0502)
σ_3 (unmarried, high AFQT)	0.2786	0.4793	173.84	0.7108
	(0.2653)	(0.0397)	(7.76)	(0.0944)
σ_4 (married, low AFQT)	0.5413	0.3525	125.61	0.5658
	(0.0770)	(0.0459)	(16.44)	(0.0767)
σ_5 (married, medium AFQT)	0.5187	0.3068	124.87	0.4649
	(0.0423)	(0.0410)	(5.63)	(0.0424)
σ_6 (married, high AFQT)	0.5715	0.3463	88.58	0.5503
	(0.1039)	(0.0320)	(15.77)	(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	0.2256	45.70	0.0990
	(0.0256)	(0.0197)	(5.85)	(0.0182)
Parameters of the cost of search function				
k_1 (unmarried, low AFQT)	678	2,998	3,349	9,085
	(382)	(1,818)	(2,380)	(8,106)
k_2 (unmarried, medium AFQT)	359	1,174	1,389	1,198
	(217)	(797)	(964)	(879)
k_3 (unmarried, high AFQT)	3,362	325	603	304
	(3442)	(273)	(474)	(194)
k_4 (married, low AFQT)	451	1,002	776	8,499
	(343)	(780)	(603)	(6,681)
k_5 (married, medium AFQT)	83	905	885	1,216
	(31)	(667)	(591)	(873)
k_6 (married, high AFQT)	118	170	916	1,073
	(160)	(98)	(591)	(685)
Parameters of the heterogeneity distribution				
$\Delta\mu$	0.5456	0.0201	-20.68	0.2179
	(0.1262)	(0.1185)	(15.68)	(0.1412)
Δk	12,374	19,114	11,199	56,857
	(4,217)	(7,675)	(5,584)	(28,021)
p	0.2962	0.2276	0.2720	0.2974
	(0.0505)	(0.0330)	(0.0345)	(0.0341)
Number of Observations	450	898	898	449
Log-Likelihood	-4041.5	-8484.8	-8509.12	-4270.2
LR statistic for β=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 †

Low Wage

	Empirical survivor	Predicted survivor	Exit rate	Offer Probability	Acceptance probability	Reemployment 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

	Empirical	Predicted
Average duration	24.81	26.06
Re-employment wage	180.25	178.82

Medium Wage (lognormal model)

	Empirical survivor	Predicted survivor	Exit rate	Offer Probability	Acceptance probability	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

	Empirical	Predicted
Average duration	21.57	22.43
Re-employment wage	265.90	259.35

Medium Wage (normal model)

	Empirical survivor	Predicted survivor	Exit rate	Offer Probability	Acceptance probability	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

	Empirical	Predicted
Average duration	21.57	22.35
Re-employment wage	265.90	268.34

High Wage

	Empirical survivor	Predicted survivor	Exit rate	Offer Probability	Acceptance probability	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

	Empirical	Predicted
Average duration	16.47	17.84
Re-employment wage	462.34	455.01

† Source: Author's calculations.

Table 4: 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 β=1	1.45	6.04	0.97	13.4

Table 5: 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	22.21	255.42	2009	-	23.24	257.28	2009	-
Effect of: Change in UI benefits	-0.96	-0.88	-474	-238	-1.19	-1.38	-497	-386
Change in UI Duration	-0.77	-0.61	-236	-71	-1.18	-1.23	-269	-188
Change in UI time profile	-0.04	0.02	251	235	-0.33	-0.39	-227	228
Job search assistance	-1.01	2.71	10,673	465	-1.45	2.85	26,108	403
Monitoring search effort	-4.84	0.31	-200	460	-6.34	-0.09	-250	-511
Monitoring Acceptance	-0.63	-4.71	-422	-439	-0.04	-7.45	-340	-399
Employment Bonus	-2.10	-2.38	578	1085	-2.63	-3.70	548	766
Benchmark with ability to commit	-12.55	13.69	-788	21,065	-	-	-	-

[†] **Notes:** Entries in the Table represent changes 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 β=1. Source: Author's calculations.

Table 5: Policy Evaluations (continued) †

Medium Wage Sample – Normal 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	22.79	267.08	2078	-	22.96	267.03	2082	-
Effect of:								
Change in UI benefits	-2.41	-0.56	-626	-330	-2.30	-0.66	-619	-403
Change in UI Duration	-1.70	-0.42	-294	-138	-1.83	-0.55	-313	-197
Change in UI time profile	0.03	-0.18	325	244	-0.30	-0.26	277	229
Job search assistance	-1.21	0.69	8,758	200	-1.36	0.82	14,165	228
Monitoring search effort	-4.37	0.78	-196	52	-4.72	0.54	-218	-148
Monitoring Acceptance	-5.28	-25.60	-355	-2,278	-5.48	-24.03	-384	-2,316
Employment Bonus	-4.12	-1.48	457	901	-4.23	-1.71	446	753
Benchmark with ability to commit	-6.36	1.18	-361	1,302	-	-	-	-

† **Notes:** Entries in the Table represent changes 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 β=1. Source: Author's calculations.

Table 6: Designing an Optimal Policy for Monitoring Search Effort

1. Lognormal Model

	Expected Duration	Expected wage
Benchmark:	22.21	255.42

	Difference relative to benchmark in:				
Benefit sanctions if effort below:	**Expected Duration**	**Expected wage**	**Gov't expenditure**	**Long-run utility**	**Social welfare**
0.01	0.00	0.00	0	0	0
0.02	-0.53	0.01	-25	64	89
0.03	-2.27	0.08	-93	259	352
0.04	-3.69	0.19	-150	387	537
0.05	*-4.84*	*0.31*	*-200*	*460*	*660*
0.06	-4.26	0.19	-347	232	578
0.07	-1.10	-0.02	-637	-450	187
0.08	-1.10	-0.02	-637	-450	187
0.09	-1.10	-0.02	-637	-450	187
0.10	-1.11	-0.02	-638	-448	190
0.11	-1.20	0.08	-645	-427	218
0.12	-1.30	0.24	-652	-400	252
0.13	-1.40	0.43	-658	-372	285
0.14	-1.49	0.63	-664	-348	316
0.15	-1.57	0.82	-670	-325	344

2. Normal Model

	Expected Duration	Expected wage
Benchmark:	22.79	267.08

	Difference relative to benchmark in:				
Benefit sanctions if effort below:	**Expected Duration**	**Expected wage**	**Gov't expenditure**	**Long-run utility**	**Social welfare**
0.01	0.00	0.00	0	0	0
0.02	-0.55	0.12	-28	20	48
0.03	-1.80	0.36	-85	51	136
0.04	*-3.12*	*0.59*	*-143*	*63*	*206*
0.05	-4.32	0.77	-196	51	247
0.06	-5.44	0.90	-252	23	275
0.07	-6.47	0.96	-307	-18	289
0.08	-7.39	0.97	-359	-69	290
0.09	-7.01	0.79	-468	-182	286
0.10	-7.62	0.73	-509	-232	277
0.11	-5.01	0.13	-745	-453	292
0.12	-5.36	0.14	-779	-467	312
0.13	-4.57	-0.10	-890	-551	339
0.14	-4.37	-0.19	-954	-588	366
0.15	-4.68	-0.14	-990	-582	408

Table 7: Optimal Monitoring of Search Effort and Worker Heterogeneity

	Unmarried, Low AFQT		Unmarried, Medium AFQT		Unmarried, High AFQT		Married, Low AFQT		Married, Medium AFQT		Married, High AFQT	
	High cost	Low cost	High cost	Low cost	High cost	Low cost	High cost	Low cost	High cost	Low cost	High cost	Low cost
1. Lognormal Model												
Benchmark												
Search effort in week 1	0.0146	0.0977	0.0153	0.1966	0.0138	0.5290	0.0158	0.2224	0.0161	0.2390	0.0139	0.8196
***s* = 0.05**												
Search effort in week1	0.0500	0.0977	0.0500	0.1966	0.0500	0.5290	0.0500	0.2224	0.0500	0.2390	0.0500	0.8196
Change in expected Duration	-23.13	0.00	-21.17	0.00	-23.58	0.00	-20.22	0.00	-18.94	0.00	-21.62	0.00
Change in utility relative to benchmark	2197	0	2034	0	1882	0	2034	0	2035	0	1884	0
***s* = 0.1**												
Search effort in week1	0.0211	0.1000	0.0231	0.1966	0.0217	0.5290	0.0239	0.2224	0.0253	0.2390	0.0237	0.8196
Change in expected Duration	-4.27	-0.07	-4.66	0.00	-5.01	0.00	-4.62	0.00	-5.02	0.00	-5.70	0.00
Change in utility relative to benchmark	-1790	15	-1902	0	-1936	0	-1898	0	-2112	0	-2286	0
2. Normal Model												
Benchmark												
Search effort in week 1	0.0101	0.0473	0.0119	0.0968	0.0094	0.1365	0.0111	0.1373	0.0127	0.1380	0.0166	0.1807
***s* = 0.05**												
Search effort in week1	0.0500	0.0500	0.0500	0.0968	0.0500	0.1365	0.0500	0.1373	0.0500	0.1380	0.0500	0.1807
Change in expected Duration	-22.29	-0.06	-15.53	0.00	-19.93	0.00	-15.95	0.00	-12.73	0.00	-9.17	0.00
Change in utility relative to benchmark	94	3	180	0	195	0	178	0	234	0	225	0
***s* = 0.1**												
Search effort in week1	0.0273	0.1000	0.1000	0.1000	0.1000	0.1365	0.1000	0.1373	0.1000	0.1380	0.1000	0.1807
Change in expected duration	-9.63	-5.52	-29.88	-0.06	-37.02	0.00	-30.47	0.00	-25.66	0.00	-20.10	0.00
Change in utility relative to benchmark	-2236	235	-811	2	-899	0	-799	0	-607	0	-409	0

Appendix Table 1: Methods of Search †

	Summary Statistics				Maximum likelihood estimates	
	(1)	(2)	(3)	(4)	(5)	(6)
Method of Search	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.

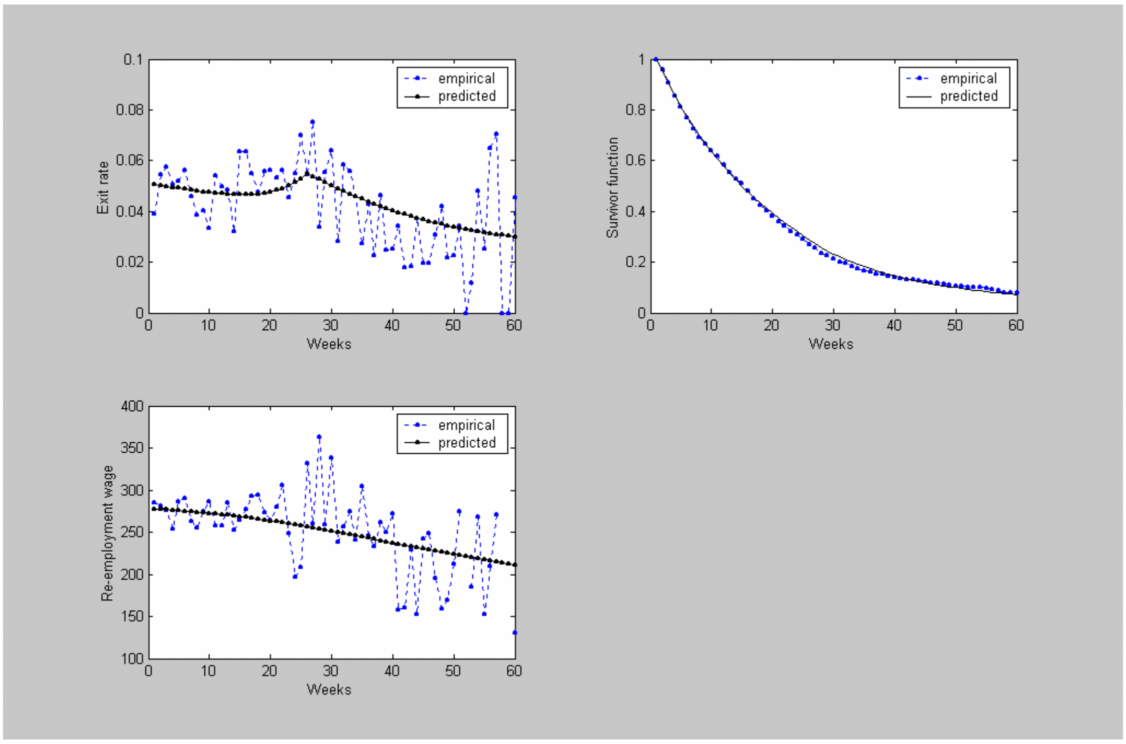

Fig. 1a: Job Search Dynamics, Medium Wage Sample
Lognormal Model

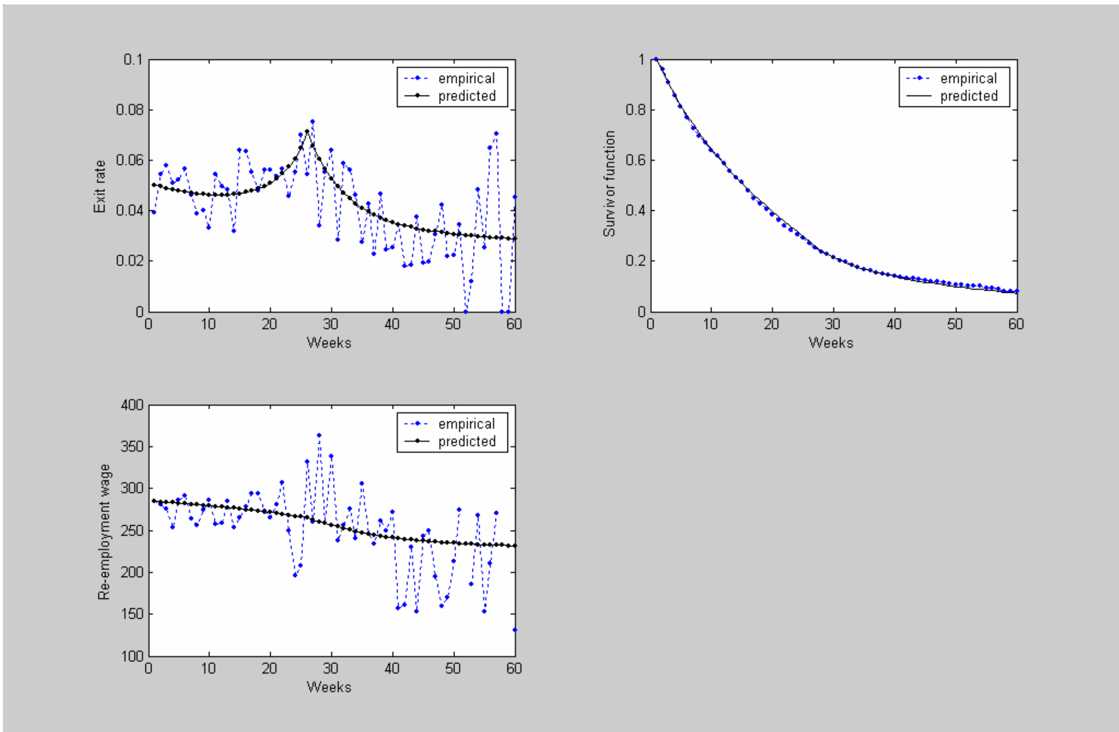

Fig. 1b: Job Search Dynamics, Medium Wage Sample
Normal Model