

Board of Governors of the Federal Reserve System

International Finance Discussion Papers

Number 710

September 2001

TESTING OPTIMALITY IN JOB SEARCH MODELS

John W. Schindler

NOTE: International Finance Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. References in publications to International Finance Discussion Papers (other than an acknowledgment that the writer has had access to unpublished material) should be cleared with the author or authors. Recent IFDPs are available on the Web at www.federalreserve.gov/pubs/ifdp/.

TESTING OPTIMALITY IN JOB SEARCH MODELS

John W. Schindler^{*}

Abstract: This paper uses Bayesian techniques to compare three definitions of optimality for the basic job search model: the standard income-maximizing definition, an approximation to the standard definition, and a simple alternative. The important role of prior choice in these comparisons is illustrated. Using natural conjugate priors to represent hypothetical samples of data, we find that the simple alternative is preferred to the standard definition of optimality. However, using priors constructed from findings in the literature, we are able to find some evidence in favor of the standard definition of optimality.

Keywords: Bayesian analysis, job search, Bayes factors

^{*} Schindler is an economist in the International Finance Division of the Federal Reserve Board and can be reached by email at john.w.schindler@frb.gov. Helpful comments were provided by Roberto Mariano, Frank Schorfheide, Lawrence Klein, and seminar participants at the Federal Reserve Board. The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any other person associated with the Federal Reserve System.

1 Introduction

Job search models have always assumed that people behave optimally. This assumption leads to the existence of a reservation wage such that any wage offer above it is accepted and any offer below it is rejected. Efforts to determine that people had a reservation wage have been inconclusive (see Lancaster and Chesher (1983)). Despite the lack of evidence in its favor, the optimally determined reservation wage has survived due to its simplicity, its appeal to basic intuition, and the form it gives to the choice of the unemployed person. As job search models have progressed, therefore, the question has turned away from whether people have reservation wages and focused more on determining the definition of optimality that people searching for a job use when choosing among various offers.

Unfortunately, testing reservation wages or definitions of optimality to determine which one is the most accurate representation of reality is difficult, if not impossible, using classical techniques. However, using Bayesian methods and current simulation techniques, it is now possible to determine which, among a competing set of definitions of optimality, is preferred. Recent work done by Lancaster (1997), Kiefer and Steel (1998), and Schindler (2000) showed that the application of Bayesian techniques to job search models was feasible under a broad range of circumstances, and each paper presented the necessary algorithms for sampling from the posterior distribution of a job search model. Following up on that work are papers by Koop (1999), Koop and Poirier (1999), and Schindler (2001), in which Bayesian techniques are used to compare different job search models.

This paper demonstrates how Bayesian techniques can be used to compare different definitions of optimality and illustrates how the choice of prior distribution can be used to aid in the analysis. Specifically, the paper uses the simplest one-sided job search model to examine two

definitions of optimality -- the traditional income-maximizing definition and an approximation to the traditional definition-- and a simple alternative.¹ The most important part of the paper, however, is the specification of the prior distributions. Three types of priors are used. The first type, natural conjugate priors, is used to represent hypothetical samples of data. This allows us to look at the definitions of optimality in different pre-selected job market scenarios, even when we do not have data from a particular scenario. The second type of prior used is a literature-based prior. These are priors constructed from findings in the literature, and hence they represent, in some ways, the best thinking about the various parameters to which priors are assigned. These priors are used to combat the notion that prior choice is completely subjective. Finally, noninformative priors are used in some situations to represent ignorance over the values of the parameters and to “let the data speak for themselves.”

With the definitions of optimality and the priors specified, samples are drawn from the posterior distribution of each combination of optimality definition and prior, and Bayes factors are calculated. The Bayes factors are then interpreted for their evidence for or against the various definitions of optimality in the scenarios represented by the prior distributions.

The results show that the economic conditions in which job search takes place, can affect the definition of optimality that is preferred, and this suggests that people who are looking for a job behave differently in different economic conditions. Interestingly, the traditional income-maximizing definition of optimality is seldom preferred to the other definitions. However, when literature-based priors are used, the evidence against the traditional definition is weaker. The

¹ The results obtained using a simple one-sided model are suggestive of what would be found if the techniques were applied to modern search models. Koop (1998) and Schindler (2001) investigate more complex job search models using Bayesian techniques.

results illustrate the usefulness of Bayesian techniques and the choice of prior distribution for analyzing job search models.

In the next section of the paper, the simple job search model used throughout this paper is described, and the definitions of optimality that will be compared are also outlined. In the third section, the basic Bayesian model and the construction of the various types of prior distributions are discussed. Section four describes the methodology used to compare the definitions of optimality. Section five discusses the results of the comparisons, and the final section concludes.

2 The Basic Job Search Model

The goal of the basic job search model is to model the process that an unemployed worker faces when attempting to get a job. We begin with a few simple assumptions. First, wage offers, w , come from a distribution known to the worker, call it $f(w)$, although the worker is not assumed to know which firms are offering which wages. The model we will use is an infinite horizon model, without layoffs or quits. Hence, once a worker accepts a job, they hold that job forever. On the other hand, if they do not accept a job in a given period, they still have an infinite number of periods in which to receive and accept an offer. So the model we are using is stationary. Each period that the worker is unemployed, we assume that they receive income, b , which can be considered to be unemployment benefits net of search costs.² In addition, we will assume that the worker discounts the future at a rate of $\beta \in [0,1)$. There is, however, no guarantee that a worker will receive a job offer in each period, so let p be the probability that an offer is received in any given period.

2.1 Optimal Reservation Wage

² Unemployment benefits net of search costs, b , is set equal to 1 throughout this paper. Changing this has little effect on the results unless b is set equal to a value close to the minimum accepted wage in the sample.

Using these basic assumptions allows us to derive the optimal search strategy of an unemployed worker who attempts to maximize the expected present value of their future income stream. The result is a reservation wage strategy, whereby the job searcher accepts any wage above the reservation wage, call it \mathbf{x} , and rejects any offer below it. If we make the additional assumption that wages are distributed according to the exponential distribution with parameter, \mathbf{g} , an implicit equation for the optimal reservation wage, \mathbf{x}^{opt} , can be obtained (see Appendix A for a derivation):³

\mathbf{x}^{opt} is such that $h(\mathbf{x}^{opt}; \mathbf{b}, \mathbf{g}, p) = 0$, where

$$h(\mathbf{x}; \mathbf{b}, \mathbf{g}, p) \equiv \frac{\mathbf{g}(\mathbf{x} - \mathbf{b})}{e^{-\mathbf{g}\mathbf{x}}} - \frac{\mathbf{b}}{1 - \mathbf{b}} p.$$

Numerically solving this equation for the reservation wage results in what we will refer to as the optimal reservation wage, and we will refer to a job searcher who has this reservation wage as behaving according to an exactly optimal strategy.

2.2 *Approximately Optimal Reservation Wage*

Some researchers have been critical of the definition of the optimal reservation wage and, in fact, few people believe that people behave according to the exact definition of optimality given above. However, within the context of the simple job search model, few alternatives have been proposed.

There are, of course, many reasons why the exact definition may fail in practice, including the job searcher's ignorance over the values of the specific parameters, its assumption of worker homogeneity, and its assumptions about the function that searchers attempt to maximize.

³ Because the focus of this paper is on illustrating the Bayesian techniques and the use of different types of priors, the use of this simple wage offer distribution is not seen as a problem. The exponential distribution is used for the simplification of calculation, but the techniques illustrated in this paper do not rely upon its use.

Therefore, since we are testing the imposition of optimality in job search models, it makes sense to include some definition of optimality that acknowledges these weaknesses in its definition. One way of doing this is to take the exact reservation wage and add a random error term to it. The error term would represent the unknown misspecification of the reservation wage equation. The new reservation wage that results would represent a wage that is approximately optimal according to the implicit equation, but which tries to account for the fact that the equation is too simple to account for all of the considerations of a worker searching for a job. Such an approximately optimal reservation wage would be represented by:

$$\begin{aligned}\mathbf{x}_i^{aopt} &= \mathbf{x}^{opt} + \mathbf{e}_i, \text{ where} \\ \mathbf{x}^{opt} &\text{ is such that } h(\mathbf{x}^{opt}; \mathbf{b}, \mathbf{g}, p) = 0, \\ h(\mathbf{x}; \mathbf{b}, \mathbf{g}, p) &\equiv \frac{\mathbf{g}(\mathbf{x} - \mathbf{b})}{e^{-\mathbf{g}\mathbf{x}}} - \frac{\mathbf{b}}{1 - \mathbf{b}} p, \text{ and} \\ \mathbf{e}_i &\sim N(0, \mathbf{S}_i), i = 1, 2.\end{aligned}$$

Many different values of \mathbf{S} could be chosen to represent different levels of departure from exact optimality. For the purposes of this paper, two different levels of \mathbf{S} were chosen. For every model (i.e., for each choice of prior distribution), the first value of \mathbf{S} chosen is set equal to one-third the standard deviation of the distribution of the reservation wage when the reservation wage is determined exactly. The second level of \mathbf{S} was then set equal to two-thirds of the standard deviation of the distribution of the reservation wage when the reservation wage is determined exactly.⁴

2.3 Simple Alternative Reservation Wage

⁴ Setting \mathbf{S} equal to values larger than these proved redundant, as the results were not affected.

The simple alternative, against which the exact and approximate definitions of optimality will be compared, does not propose a new definition of optimality and it does not alter the previous definition. Rather, the simple alternative is an estimate of the reservation wage obtained from the data. The estimate used will be the lowest accepted wage offer, w_{min} , and will be denoted \mathbf{x}^{simp} .⁵ If we assume that people are homogeneous and that they have a reservation wage, then this alternative is like the alternative hypothesis in a classical hypothesis test. In the classical sense, if the null hypothesis were that $\mathbf{x} = \mathbf{x}^{opt}$ or $\mathbf{x} = \mathbf{x}^{aopt}$, then this could be represented as the alternative $\mathbf{x} \neq \mathbf{x}^{opt}$ or $\mathbf{x} \neq \mathbf{x}^{aopt}$, respectively.

3 The Bayesian Model

Since the techniques used to compare the different definitions of optimality are Bayesian, we must specify the likelihood and the prior distributions that will be used. Given the likelihood function and the prior, then using Bayes' rule we can determine the posterior, $p(\mathbf{q}|y)$, as follows:

$$p(\mathbf{q}|y) = c \ell(\mathbf{q}) p(\mathbf{q}),$$

$$\text{where } c^{-1} = \int_{\Theta} \ell(\mathbf{q}) p(\mathbf{q}) d(\mathbf{q}),$$

y is the data, \mathbf{q} is the parameter set, $\ell(\mathbf{q})$ is the likelihood, and $p(\mathbf{q})$ is the prior distribution.

3.1 Likelihood Function

The assumptions of the basic job search model outlined earlier lead to the following likelihood function:⁶

$$\ell(t, w) = (pe^{-g\mathbf{x}})^N (1 - pe^{-g\mathbf{x}})^{T-N} \mathbf{g}^N e^{-gN(\bar{w}-\mathbf{x})},$$

where N is the sample size, T is the sum of the unemployment durations for all individuals in the sample, \bar{w} is the average accepted wage, and \mathbf{x} is the reservation wage, which could be \mathbf{x}^{opt} , \mathbf{x}

⁵ For discussion of this estimate, see Christensen and Kiefer (1996).

a^{opt} , or \mathbf{x}^{simp} , depending upon which optimality assumption is used. For this paper, the data consists of the wage that was eventually accepted by the unemployed worker, w , and the duration of the completed spell of unemployment, t . The data used in this paper come from Devine (1988), where $N = 44$, $T = 746$ weeks, $\bar{w} = \$3.67$, and $w_{min} = \$2.18$.⁷

3.2 Prior Distributions⁸

The remaining part of the Bayesian model is a description of the prior distributions used. Much of the contribution of this paper lies in the prior distributions that are constructed and the way in which they are used in the analysis. There will be three types of prior distributions used: noninformative, natural conjugate, and literature-based. Each of these types is used with a specific purpose in mind, as explained below.

3.2.1 Noninformative Priors

Noninformative priors, also called ignorance priors, are used when the researcher does not wish to incorporate any prior knowledge of certain parameters into the model. All possible values of the parameters are considered equally likely. These priors are used to “let the data speak for themselves.” Noninformative priors are used in this paper to represent a baseline model against which other models can be compared. This is particularly important because Bayesian analysis is frequently criticized for the subjectivity of the choice of prior. Using a noninformative prior gives results that are asymptotically equivalent to likelihood results, and hence less open to criticism by researchers who prefer the classical approach.⁹

⁶ For a derivation of the likelihood function, see Appendix B.

⁷ Since this paper is focused on illustrating the uses of the Bayesian techniques and different priors, and because Bayesian econometrics do not suffer the same small-sample problems that classical econometrics do, this data set is useful. Its key advantage is the homogeneity of the small sample. For more details, see Devine (1988).

⁸ See Appendix C for a complete accounting of the form of each type of prior under the various optimality assumptions, as well as listings of the posteriors.

⁹ Noninformative priors are not without controversy. Even amongst Bayesians there is debate about their use. For a discussion of these issues, see Poirier (1995).

3.2.2 Natural Conjugate Priors

The heart of this paper lies in the use of natural conjugate priors. These priors are unique among types of prior distributions because the information they contain can be interpreted as arising from a hypothetical sample of data. This is due to the fact that the functional form of the prior is identical to that of the likelihood function.¹⁰ Thus, natural conjugate priors can be used to conduct experiments along the lines of, e.g., “What if the data indicated that the average unemployment durations was 10 weeks?” or “What if the data indicated that the average accepted wage was only \$2.50?” or some combination of statements about the data.

For the purpose of this paper, the natural conjugate priors will look like:

$$p(\mathbf{b}, \mathbf{g}, p, \mathbf{x}) \propto (pe^{-g\mathbf{x}})^{N_0} (1 - pe^{-g\mathbf{x}})^{T_0 - N_0} \mathbf{g}^{N_0} e^{-gN_0(\bar{w}_0 - \mathbf{x})}, \quad y < \mathbf{x} < w_{\min_0},^{11}$$

where $\mathbf{x} = \mathbf{x}^{opt}$, \mathbf{x}^{aopt} , or \mathbf{x}^{simp} , depending upon which type of optimality is being examined. The parameters of the natural conjugate prior, called hyperparameters, are N_0 , T_0 , \bar{w}_0 , and w_{\min_0} . N_0 has been fixed at 200 and can be interpreted as the sample size of the hypothetical data set. The size was chosen to be significantly larger than the actual data used. That gives the prior distribution most of the weight in the posterior distribution, and hence the ability to influence the results. That means that we can look at different scenarios and be assured that we are seeing the effects of the scenarios and not the data, when we look at the posterior results. In addition to fixing N_0 , w_{\min_0} has been fixed at \$2.18, the same value as the minimum accepted wage, w_{\min} , in the data sample used.¹²

¹⁰ For a discussion of natural conjugate priors and their interpretation, see Poirier (1995).

¹¹ Note that if the optimal or approximately optimal definitions are being examined then \mathbf{x} should not appear as a random variable in the joint distribution because it is a function of the other parameters.

¹² Multiple values could have been chosen for w_{\min_0} , but doing so has little value because choosing $w_{\min_0} > w_{\min}$ has no effect on results. This is because only the minimum of the two values appears in the posterior. Choosing a

The other two hyperparameters, T_0 and \bar{w}_0 , are each allowed to range over five possible values, giving 25 different natural conjugate priors. This allows us to determine if certain job market conditions lead us to prefer one definition of optimality to another. The values of T_0 and \bar{w}_0 chosen for the natural conjugate priors are $T_0 = (200, 2000, 4000, 6000, 8000)$, and $\bar{w}_0 = (2.50, 3.00, 3.50, 4.00, 4.50)$. The values of T_0 correspond to average unemployment durations of 1 week, 10 weeks, 20 weeks, 30 weeks, and 40 weeks, respectively. The data set used indicates an average unemployment duration of about 17 weeks and an average accepted wage of \$3.67. Hence, the range covered by the natural conjugate priors is considerable relative to the data set.¹³

3.2.3 Literature-Based Priors

One of the primary criticisms of Bayesian analysis centers on the subjectivity of the choice of prior distributions. Someone who disagrees with the investigator's choice of prior distributions, for example, may question a particular set of results. Frequently, an investigator using Bayesian techniques is asked to justify their choice of prior distribution as reasonable. Natural conjugate priors are typically good priors for their convenience and tractability, but some critics may not think the hyperparameters of the natural conjugate prior distribution that were chosen were reasonable. In order to attempt to answer this criticism, we have used a wide range of natural conjugate priors and, in addition, we have constructed a set of prior distributions for the parameters, which we refer to as literature-based priors. These priors were constructed by collecting results from papers in the job search literature. In that sense, we have taken advantage of the accumulated knowledge and results to date to construct prior distributions. While priors

value of $w_{\min_0} < w_{\min}$ makes little sense unless you have a strong belief that wage offers are substantially above the reservation wage, thus making the minimum accepted wage offer a poor estimate of the reservation wage.

¹³ Models using other values of T_0 and \bar{w}_0 did not yield results that were significantly different, and hence results from those models are not presented.

are typically taken to represent the prior beliefs of the investigator, the literature-based priors constructed here could be interpreted as reflecting the collective prior beliefs of those who have contributed to the job search literature.

In particular, literature-based priors have been constructed for the discount factor and the probability of receiving an offer. Literature-based priors have not been constructed for the mean of the wage offer arrival distribution or the reservation wage. Doing so would be almost impossible due to the fact that those parameters depend very heavily on the particular data set used. The other parameters are more universal (though admittedly not completely so), in that they tend to have very similar values across a wide variety of data sets and are more independent over time. The values of the discount factor and the offer probability that were collected are presented in Table 1. Notes on the particulars of the data sets from which they were taken are also included with the table.¹⁴ Having used the literature to get reasonable values for these two parameters, it was still necessary to choose a distributional form with which to represent them.

The obvious choice for the prior distribution of the discount factor, \mathbf{b} , is the beta distribution, because of the distribution's support. Thus, we assume $\mathbf{b} \sim \text{Beta}(a, b)$. Two things, then, need to be determined: the parameters of the beta distribution and the quality of the fit of the beta distribution to the prior beliefs collected from the literature. The parameters of the beta distribution are determined using the mean and standard deviation of the selected sample values. These are 0.98 and 0.03 respectively. Using the formulas for the mean and standard deviation of

¹⁴ Attempts were made to take parameter values from papers that used data sets similar to the one used in this paper, although perhaps not with complete success. The criteria for selecting values are, admittedly, subjective.

a beta distribution, it is easy to determine that $a = 12.2$ and $b = 0.2$.¹⁵ The question, then, is how well does $Beta(12.2, 0.2)$ fit the data that were collected. The data is graphed in Figure 1 along with the beta distribution with the parameters just determined. Visual inspection shows that the two distributions have roughly the same shape, and we know that they have the same mean and variance. Hence, this looks like a good fit, at least for the purposes of this paper. So, overall this appears to be an acceptable choice for the prior distribution of the discount factor.

For the probability of receiving an offer, p , we again used a beta distribution to represent the prior beliefs. The mean of the values selected from the job search literature was 0.13, and the standard deviation was 0.22. The values of the parameters of the beta distribution that matched these two moments are $a = 0.2$ and $b = 1.2$. The comparison of the collected values and the appropriate beta distribution appear in Figure 2. Again, the fitted distribution fits the sampled data reasonably well. The literature-based priors constructed for \mathbf{b} and p as just described will be referred to hereafter as the LB-1 priors.

While the literature that was sampled in order to get these prior distributions was quite extensive, some may still object that a particular value that they feel is relevant has been assigned a very low probability. To combat this objection, another set of priors have been constructed that are centered around the same mean as the literature-based priors, but which have a greater variance, and hence give greater probability to other possible values of the discount factor and the offer probability.

¹⁵ The formula for the mean of a beta distribution with parameters a and b , is: $\frac{a}{a+b}$. The formula for the standard

deviation is: $\sqrt{\frac{ab}{(a+b+1)(a+b)^2}}$.

In order to construct these additional literature-based priors, the parameters of the beta distribution were chosen to match the mean of the values selected from the literature, but not the standard deviation. For each parameter, two alternative standard deviation values were used. These standard deviations were chosen to be larger than the standard deviation obtained by sampling from the literature, but not so large as to make the beta distribution bimodal.¹⁶ So, the first alternative was chosen roughly half way between the standard deviation of the values sampled from the literature and the standard deviation that would cause the prior distribution to be bimodal. The second alternative standard deviation was chosen to be as large as possible, with the constraint that the beta distribution remain unimodal for the discount factor prior. For the offer probability, p , the second alternative standard deviation used causes the prior distribution to be slightly bimodal. This was done because a number of researchers have found very high values for the offer probability. This suggests that the collective prior distribution may be bimodal. The literature-based priors constructed with these alternative standard deviations will be referred to hereafter as the LB-2 and LB-3 priors, where LB-3 corresponds to the prior with the largest standard deviation. Table 2 summarizes the literature-based prior distributions.

The literature-based priors will obviously all have the same basic functional form, but will differ in the values of the parameters chosen. For the models in which the optimal or approximately optimal definitions are used, the form of the prior distribution will be:

$$p(\mathbf{b}, \mathbf{g}, p) \propto p^{a-1} (1-p)^{b-1} \mathbf{g}^{c-1} e^{-\mathbf{g}^f} \mathbf{b}^{e-1} (1-\mathbf{b})^{f-1}.$$

For the models in which the simple alternative is used, the form of the prior will be:¹⁷

¹⁶ In order for a beta distribution to be unimodal, one or both of the parameters defining the distribution must be greater than or equal to 1.

¹⁷ Because \mathbf{x} is a parameter in the simple alternative model, it must be assigned a prior distribution. The prior distribution chosen comes from Kiefer and Steel (1998). It has the advantage of being easy to work with and having a shape that is intuitively appealing.

$$p(\mathbf{b}, \mathbf{g}, p, \mathbf{x}) \propto p^{a-1} (1-p)^{b-1} \mathbf{g}^{c-1} e^{-\mathbf{g}^d} \mathbf{b}^{e-1} (1-\mathbf{b})^{f-1} (\bar{w}_0 - \mathbf{x})^{-(N_0+1)}, \mathbf{x} \leq 2.18.$$

The values of a , b , c , d , e , and f for the each of the literature-based priors can be found in Table 3.

4 Methodology

In order to compare the definitions of optimality for the job search model, we will calculate Bayes factors. The Bayes factor is defined as the ratio of the marginalized likelihood function of two models. In order to calculate the Bayes factors, a sample from the posterior distribution is required. A number of different algorithms can be used to simulate from the posterior distributions of job search models, including those used in Lancaster (1997), Kiefer and Steel (1998), Koop (1998), and Schindler (2000). The algorithm used in this paper consists of a Gibbs sampler with a rejection sampler built into the iterations to sample from conditional distributions from which observations are not easily drawn. The algorithm is described in detail in Schindler (2000).

Once a sample has been drawn, calculating estimates of the marginalized likelihood is straightforward, and there are several formulas that can be used. The formula for the marginalized likelihood used in this paper is given by:

$$\bar{p}(y|H) = \left\{ \frac{1}{n} \sum_{i=1}^n p(y|\mathbf{q}, H) \right\}^{-1},$$

where H is the model for which you are calculating the marginalized likelihood and \mathbf{q} are the simulated parameter values from the posterior distribution.¹⁸ The Bayes factor is then calculated as the ratio of the estimates of the marginalized likelihood for the two models.

¹⁸ More about the various ways to calculate marginalized likelihoods and details about this method can be found in Kass and Raftery (1995).

Using Bayes factors to compare models is typically very straightforward. If the Bayes factor is greater than one, that indicates that the model in the numerator is preferred to the one in the denominator for the data sample used.¹⁹ Using Bayes factors to compare different definitions of optimality in job search models, however, is made more difficult by the fact that the likelihood function is not altered by the different definitions of optimality. This does not mean that we cannot use Bayes factors to compare the models. It does, however, mean that understanding why one definition of optimality is preferred to another is not as straightforward. The key is to focus on the non-identified parameters of the model. For our model there is only one non-identified parameter, the discount factor. Note that the discount factor does not appear in the likelihood function, and hence the data does not tell us anything directly about it. However, the reservation wage does appear in the likelihood function, and our beliefs about it are updated based upon information in the data. Since the reservation wage is a function of the discount factor, we can learn about the discount factor indirectly through the posterior distribution of the reservation wage.²⁰ Furthermore, since the likelihood function will be the same regardless of the definition of optimality that we apply, the only difference between the various models will be the way in which the reservation wage is determined. Koop and Poirier (2000) explain how to interpret Bayes factors for comparing models with different definitions of optimality.²¹ They point out that given any \mathbf{g} , p , and \mathbf{x} , you can solve for the value of the discount factor that ensures that the reservation wage was determined according to a given definition of optimality. Hence, every reservation wage can be considered optimal given the right discount factor. The question, then, is

¹⁹ In fact, a Bayes factor greater than one is typically not considered enough evidence in favor of either model. Scales have been suggested for how large a Bayes factor should be before one model is truly to be preferred to another. In this paper, Jeffrey's scale will be made use of, although it should be noted that the results presented in this paper do not depend crucially on this choice of scale. For details, see Jeffrey (1961).

²⁰ This type of indirect learning about non-identified parameters is the subject of Poirier (1998).

whether the choice of discount factor required to make the reservation wage optimal is a reasonable one. Clearly a discount factor greater than one or less than zero would be unreasonable. However, we can say even more because “reasonable”, in the Bayesian sense, is determined in relation to our priors. So, if the discount factors implied by the other parameters lie far out in the tail of the prior distribution, that value would be considered “unreasonable”. The models whose implied discount factors are more unreasonable relative to their priors would most likely not be chosen by Bayes factors. Therefore, it should be kept in mind that when the Bayes factors prefer one model over another, this can be interpreted in terms of the implications of the reservation wage distribution for the discount factors and how those implications compare to the prior beliefs for the discount factors.

5 Comparing Definitions of Optimality

While there are only four variations of optimality examined in this paper (exact, two levels of approximation, and the simple alternative), there are in fact a total of 116 models simulated. This is because for each type of optimality, 29 different prior distributions are used (25 natural conjugate, three literature-based, and one non-informative). From each model, a sample of size 2500 was drawn, which was used to calculate the estimate of the marginalized likelihood. Tables 4-7 provide summaries on the samples drawn from each model. Each table contains entries for only one of the four variations of optimality. Within each table, there are entries for each of the twenty-nine types of priors used. Tables 8-13 contain the Bayes factors that can be used to compare the models. Each of the tables contains the Bayes factors used to compare models that impose a particular type of optimality against models that impose a different type of optimality. In each table, a Bayes factor greater than one indicates evidence in favor of the model listed first

²¹ We alter their discussion slightly since the model they used also included a risk aversion coefficient. See Koop

in the table's title (e.g., A Bayes factor of 59 in the table titled, "Bayes Factors – Simple Alternative versus Exact Optimality" is evidence in favor of the simple alternative model.).

5.1 Comparing Exact Optimality and the Simple Alternative

The Bayes factors for this comparison are in Table 8. Looking first at the natural conjugate priors, the evidence against the exact optimality definition is quite compelling. All told, 22 of the 25 comparisons made using natural conjugate priors indicate that the simple alternative model is preferred. As mentioned above, that means that the posterior distribution of the discount factor is more reasonable under the simple alternative, then under exact optimality, given the prior beliefs about the distribution. The exceptions to this preference for the alternative model are found in the lower left-hand corner of the table. Here there are three models, each with a combination of high wages and short unemployment durations, which show a preference for the exactly optimal models. Hence, in economic scenarios that represent good conditions for job searchers, there is a notable preference for the exactly optimal model.

The preference for the exactly optimal model in the scenarios that represent the best conditions for job searchers suggests that perhaps in those situations people who are searching for jobs are more likely to behave like lifetime income maximizers. One could argue that when conditions are bad, a job searcher is more willing to accept any job that is offered, perhaps at a wage below the one that will maximize expected future income. This could be done in the hopes that conditions would improve later, and a new job could be found. When conditions are good, however, the job searcher can hold out for the income-maximizing offer, knowing that conditions are favorable. In fact, there are many possible explanations, but most of them lie outside the scope of the simple model.

and Poirier (2000) for details.

When we look at the literature-based priors, the results are less overwhelming. The Bayes factor in favor of the simple alternative optimality for the LB-1 prior is just 15.8. While this is still strong evidence, according to the Jeffrey's scale, it indicates that an exactly optimal model that uses reasonably chosen priors may be closer to the truth than one whose priors are chosen by different means. The LB-2 and LB-3 priors, however, quickly revert to decisive evidence in favor of the simple alternative model.

5.2 Comparing Approximate Optimality and the Simple Alternative

When we compare the approximately optimal models to the simple alternative, a similar pattern emerges. Looking at the natural conjugate priors, the two levels of approximate optimality performed similarly to each other, but with some slight differences. We again find that the simple alternative is preferred in the majority of the cases, 23 out of 25 for the first level of approximation and 24 out of 25 for the second level of approximation. As with exact optimality, the exceptions to the preference for the simple alternative are the models found in the lower left-hand corner of the table, where wages are high and unemployment durations are low.

For the literature-based priors, the preference is not as clear. There is some evidence that the simple alternative model is more likely than the first level of approximate optimality for the LB-1 prior, but the Bayes factor is only 2.3. This is not very strong evidence. The second level of approximate optimality, however, is clearly not preferred to the alternative optimality for any of the literature-based priors.

5.3 Comparing Exact Optimality and Approximate Optimality

While the evidence in favor of the simple alternative model is quite strong, another interesting question is which type of optimality, exact or approximate, is preferred. Bayes factors comparing those models are presented in Tables 11 and 12. Beginning with the non-informative

priors, it appears that the approximately optimal models are preferred over the exactly optimal models. The Bayes factor comparing the optimal model to the first level of approximate optimality is 0.36, and the Bayes factor comparing the optimal model to the second level of approximate optimality is $1.2e-6$. The former represents weak evidence in favor of approximate optimality, while the latter represents decisive evidence in favor of approximate optimality. This suggests that a reformulation of the definition of optimality to incorporate job searchers' uncertainty over market conditions and unaccounted factors would be helpful

When we look at the literature-based priors, the evidence in favor of approximate optimality grows. For example, there is very strong evidence that the first approximately optimal model is more likely than the exactly optimal model for the LB-1 and LB-2 priors, although the opposite is true for the LB-3 prior. When the optimal model is compared to the second approximately optimal model, however, the opposite pattern holds -- the exactly optimal model is more likely than the second approximately optimal model for the LB-1 and LB-2 priors, but not for the LB-3 priors.

For the natural conjugate priors, the exactly optimal models are always preferred to the approximately optimal models, although the evidence is weakest for the priors with the highest values of the average accepted wage and unemployment durations. These priors are found in the lower right-hand corner of the tables. This may reflect the fact that when unemployment durations and average accepted wages are highest, we are dealing with the 'choosiest' of job searchers. Those searchers appear willing to remain unemployed for long stretches of time in order to get the wage that they want. In that case then, it seems less likely that they are strictly following the optimal reservation wage profile, because it seems as though they must have turned down wages which were higher than the reservation wage in order for the average

accepted wage to be so high. Thus, while the optimal model is still preferred, perhaps in those situations, job searcher's behavior becomes more approximate in the sense that they are not as concerned with the wages they lose when searching as they are with the wages they earn when working. This would be the profile of a risk-loving, as opposed to risk-neutral, job searcher, and optimality can easily be redefined to incorporate different utility functions.²²

5.4 Comparing Levels of Approximate Optimality

Finally, when comparing the levels of approximate optimality, the evidence is mixed. For the noninformative priors, the higher level of approximation model is clearly preferred. That is also true for the LB-3 prior. That suggests that the greater the dispersion of the prior parameters, the more likely the second level of approximate optimality. The LB-1 and LB-2 prior models, however, clearly point to the first approximately optimal level as being more likely. The natural conjugate prior models point to the first approximate model except for the highest prior values of the average accepted wage and unemployment duration. The Bayes factors get steadily smaller (less in favor of the first level of approximate optimality) as the prior values of the average accepted wage and unemployment duration fall.

6 Conclusions

The comparisons made in this paper led to some interesting results, and those results illustrate the effectiveness of both the Bayesian approach and, more specifically, the use of different types of prior distributions. Both the hypothetical samples that the natural conjugate priors represent and the literature-based priors give researchers new tools to examine job search

²² Risk neutrality has been assumed throughout this paper, and, hence, other risk preferences are outside of the scope of the simple model examined in this paper. However, this assumption can be easily relaxed, as in Koop and Poirier (2000).

models from different perspectives. The results presented in this paper show how important the choice of prior is.

Using the natural conjugate priors to simulate a variety of economic conditions pointed to the fact that the economic conditions are important in determining which model is preferred. The vast majority of the simulations pointed to the fact that the traditional income-maximizing definition of optimality for the basic job search model is rejected in favor of the simple alternative. The exceptions to this, were the cases where the economic conditions were best from the unemployed person's perspective.

Using priors that are constructed based upon findings and assumptions from the job search literature, the results were somewhat different. The standard definition of optimality performed better when the priors were chosen in this way. That suggests that the traditional definition of optimality is better than the natural conjugate prior results would lead one to believe. However, one could argue that the logic here is circular in the sense that the results used to construct the literature-based priors were from papers that assumed the standard definition of optimality. Hence, those priors would be biased towards the optimal models. That notwithstanding, the use of the literature-based priors is still illustrative of the way in which a Bayesian model could be constructed that answers some people's objections to the choice of prior.

In addition to using different types of priors to illustrate some of the power of the Bayesian approach, this paper examined what the possible shortcomings of the traditional definition of optimality might be with respect to the simple job search model. The approximation to the traditional definition of optimality was compared to the simple alternative to see how it performed. The results were mixed but did indicate that the approximation to optimality was preferred in some cases. That suggests that while people may be behaving optimally, the

assumptions of the traditional definition may be too rigid. In particular, the approximation used suggests that assuming unemployed people know all of the characteristics of the job market in which they are searching is unrealistic. However, these results are not definitive in and of themselves. Rather, the results suggest that there is room for improvement to the traditional definition.

Finally, the results in this paper should by no means be considered the final word on this issue. The primary purpose of this paper was to illustrate how Bayesian techniques and the choice of priors could be taken advantage of in the job search literature. Further work is being done to look at models that are more interesting than the simple job search model examined here. Koop (2000) and Schindler (2001) are both attempts to look at more complex models.

References

- Christensen, B. and Kiefer, N. (1991), "The Exact Likelihood Function for the Empirical Job Search Model," *Econometric Theory*, 7, 464-486.
- Devine, T. (1998a), "Interpreting Reemployment Patterns in the Search Framework," Ph.D. Dissertation, Cornell University.
- Devine, T. (1998b), "Arrivals versus Acceptance: The Source of Variation in Reemployment Rates Across Demographic Groups," Working Paper, Pennsylvania State University.
- Devine, T. and Kiefer, N. (1991), *Empirical Labor Economics: The Search Approach*, New York: Oxford University Press.
- Eckstein, Z. and Wolpin, K. (1990), "Estimating an Equilibrium Search Model Using Panel Data on Individuals," *Econometrica*, 58, 783-808.
- Flinn, C. and Heckman, J. (1982), "New Methods for Analyzing Structural Models of Labor Force Dynamics," *Journal of Econometrics*, 18, 115-168.
- Holzer, H. (1986), "Are Unemployed Young Blacks Income Maximizers?" *Southern Economic Journal*, 52, 777-784.
- Jeffreys, H. (1961), *Theory of Probability* (3rd ed.), Oxford: Oxford University Press.
- Kass, R. and Raftery, A. (1995), "Bayes Factors," *Journal of the American Statistical Association*, 90, 773-794.
- Khandker, R. (1988), "Offer Heterogeneity in a Two State Model of Sequential Search," *The Review of Economics and Statistics*, 70, 259-265.
- Kiefer, N. and Neumann, G. (1979), "An Empirical Job-Search Model, with a Test of the Constant Reservation-Wage Hypothesis," *Journal of Political Economy*, 87, 89-107.
- Kiefer, N. and Neumann, G. (1989), *Search Models and Applied Labor Economics*, New York: Cambridge University Press.
- Kiefer, N. and Steel, M. (1998), "Bayesian Analysis of the Prototypal Search Model," *Journal of Business & Economic Statistics*, 16, 178-186.
- Koop, G. (1998), "Bayesian Inference in Models Based on Equilibrium Search Theory," manuscript.
- Koop, G. and Poirier, D. (1999), "Testing for Optimality in Stationary Job Search Models Based on Reservation Wages," manuscript.

- Koop, G. and Potter, S. (1999), "Bayes Factors and nonlinearity: Evidence from economic time series," *Journal of Econometrics*, 88, 251-281.
- Lancaster, T. and Chesher, A. (1983), "An Econometric Analysis of Reservation Wages," *Econometrica*, 51 (6), 1661-1776.
- Lancaster, T. (1997), "Exact Structural Inference in Optimal Job Search Models," *Journal of Business & Economic Statistics*, 15, 165-179.
- Mortensen, D. and Pissarides, C. (1994), "Job Creation and Job Destruction in the Theory of Unemployment," *Review of Economic Studies*, 61, 397-415.
- Mortensen, D. (1986), "Job Search and Labor Market Analyses," in *Handbook of Labor Economics*, Vol II, eds. O. Ashenfelter and R. Layard, Elsevier Science Publishers, Amsterdam, 849-977.
- Narendranathan, W., Nickell, S., and Stern, J. (1985), "Unemployment Benefits Revisited," *Economic Journal*, 95, 307-329.
- Poirier, D. (1995), *Intermediate Statistics and Econometrics*, Cambridge: The MIT Press.
- Poirier, D. (1998), "Revising Beliefs in Non-identified Models," *Econometric Theory*, 14, 483-509.
- Ridder, G. and Gorter, K. (1986), "Unemployment Benefits and Search Behavior: An Empirical Investigation," Manuscript, Cornell University.
- Schindler, J. (2000), "A General Bayesian Approach to Job Search Models: Methodology and Applications," Dissertation, University of Pennsylvania.
- Schindler, J. (2001), "Bayesian Testing of Quits, Layoffs, and On-the-Job Search in Job Search Models," manuscript.
- van den Berg, G. (1990a), "Nonstationarity in Job Search Theory," *Review of Economic Studies*, 57, 255-277.
- van den Berg, G. (1990b), "Search Behavior, Transitions to Non-Participation and the Duration of Unemployment," *Economic Journal*, 100, 842-865.
- Wolpin, K. (1995), "School and Home Parameters," Lecture Notes, University of Pennsylvania, October 1995.

Wolpin, K. (1992), "The Determinants of Black-White Differences in Early Employment Careers: Search, Layoffs, Quits, and Endogenous Wage Growth," *Journal of Political Economy*, 100, 535-560.

Wolpin, K. (1987), "Estimating a Structural Search Model: The Transition from School to Work," *Econometrica*, 55, 801-817.

Appendix A

Derivation of the Optimal Reservation Wage

The optimal strategy of the worker is determined by looking at the decision that the worker faces each period. In any given period, $V(w)$ gives the value to the unemployed worker, where w is the current wage offer. Note that $V(w)$ can be zero because you are not guaranteed to receive an offer in every period. A person following an optimal strategy will have a value in each period equal to:

$$V(w) = \max\{V^e(w), V^u\} = \max\left\{\sum_{t=0}^{\infty} b^t w, b + bEV\right\} = \max\left\{\frac{w}{1-b}, y + bEV\right\}.$$

$V^e(w)$ is the value of being employed at wage w , and is monotonically increasing in w , as can clearly be seen from the rightmost part of the equality. V^u is the value of being unemployed and it, on the other hand, is not a function of w . EV is the expected value of searching for a job in a given period. Therefore the optimal strategy for a person to follow is a reservation wage policy. Define \mathbf{x} to be that wage such that $V^e(w) = V^u$, and call it the reservation wage. Hence the reservation wage is the wage that makes the person indifferent between employment at that wage and unemployment. Then, the optimal strategy will be to accept any wage offer that is greater than \mathbf{x} , and reject all other offers. This reservation wage is defined implicitly by the equation:

$$\mathbf{x}(b, f(w), p) = (1-b)(b + bEV).$$

Through some manipulation and the assumption that wages are distributed according to the exponential distribution with parameter, \mathbf{g} , we obtain the following expression:

$$h(\mathbf{x}; \mathbf{b}, \mathbf{g}, p) \equiv \frac{\mathbf{g}(\mathbf{x} - b)}{e^{-\mathbf{g}\mathbf{x}}} = \frac{\mathbf{b}}{1 - \mathbf{b}} p.$$

Solving for the value of \mathbf{x} , such that $h(\mathbf{x}; \mathbf{b}, \mathbf{g}, p) = 0$, gives us the optimal reservation wage, \mathbf{x}^{opt} .

Appendix B

Derivation of the Likelihood Function

In our simple model, where offers are received based on a Bernoulli process, durations have a geometric distribution with parameter:

$$I = p(1 - F(\mathbf{x})),$$

where F is the CDF of $f(w)$, and $F(\mathbf{x})$ is the CDF evaluated at \mathbf{x} .

Therefore, the distribution of durations is given by:

$$p(t) = I(1 - I)^{t-1}.$$

Next, we need the distribution of wages given that the spell of unemployment is ended, or in other words, the distribution of accepted wages. This is denoted by $p(w|t)$, which is the same as the distribution of wages, w , such that w is greater than \mathbf{x} . This looks like:

$$p(w|t) = p(w|w \geq \mathbf{x}) = \frac{f(w)}{1 - F(\mathbf{x})}.$$

If $f(w)$ is exponential with parameter \mathbf{g} , then the distribution of accepted wages becomes.

$$\begin{aligned} p(w|t) &= \frac{\mathbf{g}e^{-\mathbf{g}w}}{e^{-\mathbf{g}\mathbf{x}}} \\ &= \mathbf{g}e^{-\mathbf{g}(w-\mathbf{x})} \end{aligned}$$

where \mathbf{g} is the parameter of the exponential distribution.

Using $p(t)$ and $p(w|t)$, we get the following likelihood function:

$$\begin{aligned} \ell(t, w) &= \prod_{i=1}^N p(t_i) p(w_i|t_i) \\ &= (pe^{-\mathbf{g}\mathbf{x}})^N (1 - pe^{-\mathbf{g}\mathbf{x}})^{T-N} \mathbf{g}^N e^{-\mathbf{g}N(\bar{w}-\mathbf{x})}, \end{aligned}$$

where $T = \sum_{i=1}^N t_i$ (or the sum of all individual unemployment durations), N is the sample size, and \bar{w} is the average of all accepted wages.

Appendix C Models

Having outlined the different priors (noninformative, natural conjugate, smart, and semi-smart) that will be used and having explained the different types of optimality (exactly optimal, approximately optimal, and the simple alternative) that will be used, each model can now be written out in terms of its uniquely determined posterior distribution.

1. Noninformative Priors

For the simple alternative model, the noninformative prior will take the form:

$$p(\mathbf{b}, \mathbf{g}, p, \mathbf{x}) = f(\mathbf{b})f(\mathbf{g})f(p)f(\mathbf{x}) \propto \frac{1}{\mathbf{g}^{\mathbf{x}}}.$$

This gives the posterior distribution the following form:

$$p(\mathbf{b}, \mathbf{g}, p, \mathbf{x}) \propto (pe^{-\mathbf{g}\mathbf{x}})^N (1 - pe^{-\mathbf{g}\mathbf{x}})^{T-N} \mathbf{g}^{N-1} e^{-\mathbf{g}^N(\bar{w}-\mathbf{x})} \mathbf{x}^{-1}, \quad b < \mathbf{x} = \mathbf{x}^{simp} < w_{\min}.$$

For the optimal and approximately optimal models, the priors have the following form:

$$p(\mathbf{b}, \mathbf{g}, p) = f(\mathbf{b})f(\mathbf{g})f(p) \propto \frac{1}{\mathbf{g}}.$$

This gives the posterior distribution the following form:

$$p(\mathbf{b}, \mathbf{g}, p|t, w) \propto (pe^{-\mathbf{g}\mathbf{x}})^N (1 - pe^{-\mathbf{g}\mathbf{x}})^{T-N} \mathbf{g}^{N-1} e^{-\mathbf{g}^N(\bar{w}-\mathbf{x})}, \quad b < \mathbf{x} = \mathbf{x}^{opt} \text{ or } \mathbf{x}^{aopt} < w_{\min}.$$

2. Natural Conjugate Priors

The natural conjugate prior will take the form:

$$p(\mathbf{b}, \mathbf{g}, p, \mathbf{x}) \propto (pe^{-\mathbf{g}\mathbf{x}})^{N_0} (1 - pe^{-\mathbf{g}\mathbf{x}})^{T_0-N_0} \mathbf{g}^{N_0} e^{-\mathbf{g}^{N_0}(\bar{w}_0-\mathbf{x})}, \quad b < \mathbf{x} < w_{\min},$$

where $\mathbf{x} = \mathbf{x}^{opt}$ or \mathbf{x}^{aopt} or \mathbf{x}^{simp} , depending upon what definition of optimality is used.²³

This gives the posterior distribution the following form:

²³ If the optimal or approximately optimal definitions are used, then \mathbf{x} should not appear as a parameter in the prior distribution (i.e., it should be $p(\mathbf{b}, \mathbf{g}, p)$ rather than $p(\mathbf{b}, \mathbf{g}, p, \mathbf{x})$.)

$$p(\mathbf{b}, \mathbf{g}, p, \mathbf{x} | t, w) \propto (pe^{-g\mathbf{x}})^{N+N_0} (1 - pe^{-g\mathbf{x}})^{T+T_0-(N+N_0)} \mathbf{g}^{N+N_0} e^{-\mathbf{g}(\overline{w} + N_0 \overline{w}_0 - \mathbf{x}(N+N_0))},$$

$$b < \mathbf{x} < \min(w_{\min}, w_{\min_0}),$$

where $\mathbf{x} = \mathbf{x}^{opt}$ or \mathbf{x}^{aopt} or \mathbf{x}^{simp} , depending upon what definition of optimality is used.²⁴

3. Literature-Based Priors²⁵

For the simple alternative model, the literature-based prior will take the following form:

$$p(\mathbf{b}, \mathbf{g}, p, \mathbf{x}) \propto p^{a-1} (1-p)^{b-1} \mathbf{g}^{c-1} e^{-\frac{g}{d}} \mathbf{b}^{e-1} (1-\mathbf{b})^{f-1} (\overline{w}_0 - \mathbf{x})^{-N_0+1}.$$

This gives the posterior distribution the following form:

$$p(\mathbf{b}, \mathbf{g}, p, \mathbf{x} | t, w) \propto p^{a-1} (1-p)^{b-1} (pe^{-g\mathbf{x}})^N (1 - pe^{-g\mathbf{x}})^{T-N}$$

$$\times \mathbf{g}^{N+c-1} e^{-\mathbf{g}N(\overline{w}-\mathbf{x})-\frac{g}{d}} \mathbf{b}^{e-1} (1-\mathbf{b})^{f-1} (\overline{w}_0 - \mathbf{x})^{-N_0+1}, b < \mathbf{x} = \mathbf{x}^{simp} < w_{\min}.$$

For the optimal and approximately optimal models, the priors have the following form:

$$p(\mathbf{b}, \mathbf{g}, p) \propto p^{a-1} (1-p)^{b-1} \mathbf{g}^{c-1} e^{-\frac{g}{d}} \mathbf{b}^{e-1} (1-\mathbf{b})^{f-1}.$$

This gives the posterior distribution the following form:

$$p(\mathbf{b}, \mathbf{g}, p | t, w) \propto p^{a-1} (1-p)^{b-1} (pe^{-g\mathbf{x}})^N (1 - pe^{-g\mathbf{x}})^{T-N}$$

$$\times \mathbf{g}^{N+c-1} e^{-\mathbf{g}N(\overline{w}-\mathbf{x})-\frac{g}{d}} \mathbf{b}^{e-1} (1-\mathbf{b})^{f-1}, b < \mathbf{x} < \min(w_{\min}, w_{\min_0}),$$

where $\mathbf{x} = \mathbf{x}^{opt}$ or \mathbf{x}^{aopt} , depending upon what definition of optimality is used.

²⁴ If the optimal or approximately optimal definitions are used, then \mathbf{x} should not appear as a parameter in the posterior distribution (i.e., it should be $p(\mathbf{b}, \mathbf{g}, p | t, w)$ rather than $p(\mathbf{b}, \mathbf{g}, p, \mathbf{x} | t, w)$.)

²⁵ The parameters of the literature-based priors, a, b, c, d, e , and f , can be found in Table 3.

Figure 1
Collected Discount Factors versus Beta(12.2,0.2)

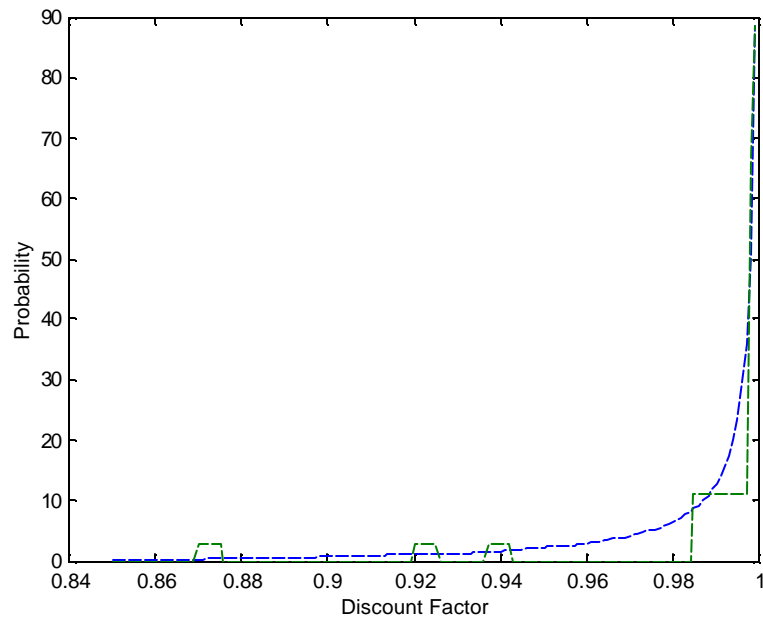


Figure 2
Collected Offer Probability Values versus Beta(0.2,1.2)

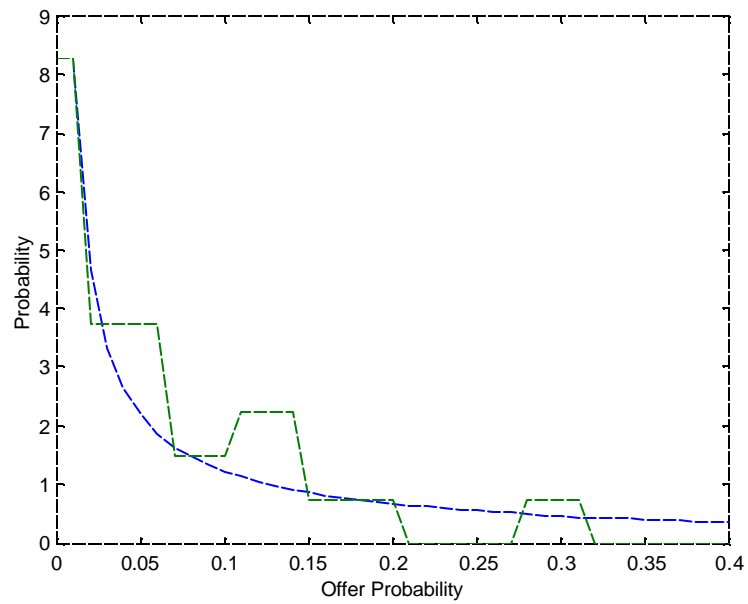


Table 1
Values used to construct Literature-Based Priors

<u>Author(s)/Year</u>	<u>b</u>	<u>p</u>
1. Kiefer and Steel (1998)	.9209, .8720, .9414	.2555, .3115, .6043
1984 SIPP Public Use Files, 16-19yrs., non-white males, not disabled, not self-employed, not in agriculture, $N=44$, $b=1$, $\bar{w} = \$3.67$, $w_{min} = \$2.18$, $T=746$		
2. Mortensen and Pissarides (1994)	.99923	.0065
Values chosen for the purpose of simulation		
3. Wolpin (1992)	.99907	.11922, .17424
Assumed value of parameters		
4. Wolpin and Eckstein (1990)	.999	.96125 (avg.), .0625, .0150
Kaplan-Meier Hazards rates and assumed value for discount factor. Other parameters are model estimates. Data are from NLS of Labor Market Experience Youth Cohort, high school graduates, from 1979-1981, $N=1339$.		
5. van den Berg (1990a)	.99754	.0095, .0112, .0143
Values were estimated from a model. Data is from 400 males in Amsterdam.		
6. van den Berg (1990b)	.99688, .99798, .99901	.012
Data are from the Netherlands Socio-Economics Panel, 223 men, ages 17-65 who were unemployed when first interviewed. The discount rates were assumed values.		
7. Wolpin (1987)	.999	.0110 (avg.), .01, .05, .10
The Kaplan-Meier hazards are from the 1979 NLS Youth Cohort, white males subsample, while the other values were estimated using his models.		
8. Ridder and Gorter (1986)	.99798	.0281
Netherlands Institute for Public Opinion, 1983, Amsterdam male population, 30-55yrs old, 428 workers.		
9. Narendranathan, et al. (1985)	.98615, .99811	.1194, .0940

Estimated values, Cohort Study of Unemployed, 1978, 2332 men

10. Flinn and Heckman (1982) .99798, .99901 .0510, .0325

Assumed values for discount rate and estimated values for the offer probability. NLS of Young Men, 1966, 2915 observations

11. Wolpin (1995) .99874

School and Home Parameters taken from Lecture Notes

12. Mortensen (1986) .99923

Assumed values of parameters.

13. Kiefer and Neumann (1979) .99808

Trade Adjustment Assistance Program, 517 male workers in 14 states, October 1975, who experienced permanent layoffs, accepted wages observed for 327 workers.

14. Holzer (1986) .99444

Estimated values, NBER Survey of Inner City Black Youth, 1979-1980, 2400 males, 16-24 years.

15. Devine (1988) .0465 (avg.), .14 (avg.)

Survey of Income and Program Participation, 1984, 5214 workers

16. Khanker (1988) .0375

EOPP Survey, 1980, 12-14yrs of schooling, 1285 observations, people in their 20s.

Table 2
Beta Parameters for Literature-Based Priors

Prior	Beta Distribution Parameters (a, b)	Mean	Standard Deviation
<i>Discount Factor</i>			
LB-1	(12.1655, .1904)	.9846	.0337
LB-2	(2.7307, .0427)	.9846	.0634
LB-3	(1.0159, .0159)	.9846	.0865
<i>Offer Probability</i>			
LB-1	(.1847, 1.2244)	.1311	.2174
LB-2	(.1511, 1.0019)	.1311	.2300
LB-3	(.0986, .6536)	.1311	.2550

Table 3
Values of a , b , c , d , e , and f for Literature-Based Priors

	LB-1	LB-2	LB-3
a	0.1847	0.1511	0.0986
b	1.2244	1.0019	0.6536
c	11.0000	11.0000	11.0000
d	0.0549	0.0549	0.0549
e	12.1655	2.7307	1.0159
f	0.1904	0.0427	0.0159

Table 4
Simulation Results for Simple Alternative Model
Sample Means

Natural Conjugate Priors

			T_0				
			200	2000	4000	6000	8000
\bar{w}_0	\$2.50	p	0.9960	0.9832	0.9520	0.7227	0.5576
		g	0.7742	1.1520	1.2855	1.2304	1.3975
		x	2.1660	2.1576	2.1759	2.1772	2.1693
	\$3.00	p	0.9911	0.8650	0.5269	0.3846	0.2968
		g	0.7063	1.0596	1.0202	1.0854	1.0892
		x	2.1164	2.1661	2.1467	2.1682	2.1724
	\$3.50	p	0.9739	0.4575	0.2634	0.1860	0.1436
		g	0.6664	0.7601	0.7149	0.7526	0.7879
		x	2.1225	2.1445	2.1192	2.1499	2.1261
	\$4.00	p	0.8842	0.3119	0.1813	0.1268	0.0986
		g	0.5795	0.5853	0.5442	0.5848	0.5973
		x	2.1131	2.1235	2.1145	2.1224	2.0949
	\$4.50	p	0.7127	0.2465	0.1429	0.1007	0.0798
		g	0.4778	0.4825	0.4358	0.4818	0.4832
		x	2.1079	2.1115	2.0970	2.1058	2.0632

Noninformative and Literature-Based Priors

	Non-Inf.	LB-1	LB-2	LB-3
p	0.2575	0.2469	0.2497	0.2471
g	0.6646	0.6542	0.6559	0.6571
x	2.1473	2.0608	2.0616	2.0603

Table 5
Simulation Results for Optimal Model
Sample Means

Natural Conjugate Priors

			T_0				
			200	2000	4000	6000	8000
\bar{w}_0	\$2.50	p	0.9956	0.9828	0.9567	0.7222	0.5211
		\mathbf{g}	0.7760	1.1721	1.3644	1.3765	1.3750
		\mathbf{x}	1.8161	1.7803	1.6658	1.5812	1.5058
		\mathbf{b}	0.7127	0.8728	0.8870	0.8887	0.8916
	\$3.00	p	0.9900	0.8569	0.5425	0.3845	0.2855
		\mathbf{g}	0.7089	1.0473	1.0787	1.0769	1.0776
		\mathbf{x}	1.8852	1.8308	1.6934	1.5922	1.5207
		\mathbf{b}	0.7001	0.8669	0.8828	0.8858	0.8901
	\$3.50	p	0.9715	0.4520	0.2637	0.1861	0.1429
		\mathbf{g}	0.6461	0.7526	0.7502	0.7506	0.7489
		\mathbf{x}	1.9348	1.8884	1.7241	1.6254	1.5486
		\mathbf{b}	0.6801	0.8553	0.8726	0.8809	0.8841
	\$4.00	p	0.8733	0.3085	0.1797	0.1267	0.0986
		\mathbf{g}	0.5690	0.5760	0.5742	0.5738	0.5766
		\mathbf{x}	1.9850	1.9489	1.7672	1.6629	1.5849
		\mathbf{b}	0.6630	0.8423	0.8637	0.8747	0.8796
	\$4.50	p	0.7003	0.2419	0.1414	0.1007	0.0708
		\mathbf{g}	0.4703	0.4715	0.4667	0.4672	0.4723
		\mathbf{x}	2.0421	2.0016	1.8279	1.7127	1.5625
		\mathbf{b}	0.6455	0.8321	0.8587	0.8699	0.8793

Noninformative and Literature-Based Priors

	Non-Inf.	LB-1	LB-2	LB-3
p	0.2571	0.2295	0.2381	0.2445
\mathbf{g}	0.6927	0.6735	0.6635	0.6616
\mathbf{x}	1.2407	1.8489	1.6387	1.4472
\mathbf{b}	0.4599	0.8838	0.7879	0.6410

Table 6
Simulation Results for First Approximately Optimal Model
Sample Means

Natural Conjugate Priors

			T_0				
			200	2000	4000	6000	8000
\bar{w}_0	\$2.50	p	0.9957	0.9829	0.9567	0.7227	0.5437
		\mathbf{g}	0.7763	1.1725	1.3633	1.3736	1.3744
		\mathbf{x}	1.8220	1.7795	1.6503	1.5836	1.5129
		\mathbf{b}	0.7157	0.8737	0.8837	0.8901	0.8921
	\$3.00	p	0.9903	0.8578	0.5412	0.3824	0.2839
		\mathbf{g}	0.7081	1.0466	1.0777	1.0790	1.0786
		\mathbf{x}	1.8806	1.8308	1.6749	1.5807	1.5094
		\mathbf{b}	0.6987	0.8664	0.8797	0.8846	0.8883
	\$3.50	p	0.9704	0.4539	0.2644	0.1864	0.1431
		\mathbf{g}	0.6461	0.7516	0.7508	0.7498	0.7489
		\mathbf{x}	1.9360	1.8836	1.7171	1.6199	1.5504
		\mathbf{b}	0.6819	0.8543	0.8714	0.8801	0.8845
	\$4.00	p	0.8737	0.3708	0.1799	0.1259	0.0986
		\mathbf{g}	0.5698	0.5774	0.5742	0.5740	0.5758
		\mathbf{x}	1.9848	1.9473	1.7661	1.6652	1.5923
		\mathbf{b}	0.6638	0.8435	0.8642	0.8759	0.8817
	\$4.50	p	0.6986	0.2416	0.1413	0.1006	0.0695
		\mathbf{g}	0.4705	0.4721	0.4663	0.4678	0.4732
		\mathbf{x}	2.0317	1.9988	1.8212	1.7047	1.5486
		\mathbf{b}	0.6439	0.8329	0.8579	0.8694	0.8784

Noninformative and Literature-Based Priors

	Non-Inf.	LB-1	LB-2	LB-3
p	0.2545	0.2271	0.2384	0.2452
\mathbf{g}	0.6837	0.6721	0.6647	0.6605
\mathbf{x}	1.2862	1.8417	1.6256	1.5178
\mathbf{b}	0.5053	0.8857	0.7870	0.6935

Table 7
Simulation Results for Second Approximately Optimal Model
Sample Means

Natural Conjugate Priors

			T_0				
			200	2000	4000	6000	8000
\bar{w}_0	\$2.50	p	0.9955	0.9824	0.9576	0.7204	0.5561
		g	0.7766	1.1717	1.3670	1.3760	1.3754
		x	1.8111	1.7762	1.6481	1.5728	1.5100
		b	0.7148	0.8741	0.8846	0.8886	0.8898
	\$3.00	p	0.9910	0.8587	0.5418	0.3812	0.2821
		g	0.7091	1.0457	1.0784	1.0784	1.0780
		x	1.8701	1.8307	1.6909	1.5835	1.5141
		b	0.6996	0.8681	0.8821	0.8856	0.8908
	\$3.50	p	0.9707	0.4521	0.2633	0.1858	0.1433
		g	0.6463	0.7506	0.7502	0.7488	0.7481
		x	1.9279	1.8844	1.7171	1.6152	1.5439
		b	0.6830	0.8559	0.8730	0.8795	0.8833
	\$4.00	p	0.8743	0.3081	0.1796	0.1267	0.0987
		g	0.5695	0.5771	0.5753	0.5740	0.5755
		x	1.9758	1.9390	1.7568	1.6522	1.5439
		b	0.6645	0.8433	0.8646	0.8729	0.8833
	\$4.50	p	0.6967	0.2415	0.1417	0.1005	0.0713
		g	0.4700	0.4721	0.4666	0.4664	0.4718
		x	2.0283	0.9970	1.8107	1.6936	1.5618
		b	0.6453	0.8337	0.8573	0.8687	0.8787

Noninformative and Literature-Based Priors

	Non-Inf.	LB-1	LB-2	LB-3
p	0.2593	0.2292	0.2406	0.2443
g	0.6892	0.6713	0.6619	0.6607
x	1.3394	1.8379	1.6455	1.5337
b	0.5193	0.8891	0.7986	0.6946

Table 8
Bayes Factors – Exact Optimality versus the Alternative

		T_0				
		200	2000	4000	6000	8000
\bar{w}_0	\$2.50	2.84e+30	3.40e+25	2.08e+30	6.40e+30	8.40e+16
	\$3.00	8.37e+11	2.50e+20	3.45e+20	1.50e+13	4.13e+7
	\$3.50	59.14	3.14e+9	2.51e+10	8.65e+6	14352.00
	\$4.00	0.02	140.71	7.44e+5	14594.00	156.38
	\$4.50	8.19e-011	0.03	1147.30	1327.00	6.74e+16
			LB-1	LB-2	LB-3	
			15.77	5.14e+20	6.49e+26	

Table 9
Bayes Factors – Non-optimal versus Approximately Optimal 1 Models

		T_0				
		200	2000	4000	6000	8000
\bar{w}_0	\$2.50	1.27e+43	2.08e+34	4.64e+42	5.50e+42	7.83e+23
	\$3.00	3.86e+23	1.17e+27	1.26e+26	5.70e+17	1.53e+10
	\$3.50	3.21e+12	2.97e+12	1.38e+13	1.67e+8	99491.00
	\$4.00	3898.90	34649.00	3.62e+6	52007.00	2234.90
	\$4.50	1.25e-8	.09	3854.00	2999.40	5.03e+17
			LB-1	LB-2	LB-3	
			2.27	1.01e+19	8.29e+31	

Table 10
Bayes Factors – Non-optimal versus Approximately Optimal 2 Models

		T_0				
		200	2000	4000	6000	8000
\bar{w}_0	\$2.50	3.09e+65	5.45e+47	1.05e+66	2.04e+53	1.26e+32
	\$3.00	5.06e+38	7.22e+38	1.22e+38	1.65e+26	4.31e+13
	\$3.50	2.76e+31	1.08e+18	5.70e+16	7.11e+11	4.56e+5
	\$4.00	1.92e+9	2.00e+6	4.37e+9	2.45e+6	1457.00
	\$4.50	.0034	1.41	91003.00	27374.00	8.50e+16
			LB-1	LB-2	LB-3	
			5.56e+12	2.78e+24	4.14e+21	

Table 11
Bayes Factors – Exactly Optimal versus Approximately Optimal 1 Models

		T_0				
		200	2000	4000	6000	8000
\bar{w}_0	\$2.50	4.48e+12	6.14e+8	2.23e+12	8.60e+11	9.31e+6
	\$3.00	4.61e+11	4.67e+6	3.66e+5	38014.00	371.12
	\$3.50	5.43e+10	946.81	551.47	19.38	6.93
	\$4.00	1.82e+5	246.25	4.87	3.56	14.29
	\$4.50	153.02	3.07	3.36	2.26	7.47
		Non-Inf.	LB-1	LB-2	LB-3	
		0.36	0.14	0.02	1.27e+5	

Table 12
Bayes Factors – Exactly Optimal versus Approximately Optimal 2 Models

		T_0				
		200	2000	4000	6000	8000
\bar{w}_0	\$2.50	1.08e+35	1.60e+22	5.05e+35	3.19e+22	1.50e+15
	\$3.00	6.05e+26	2.87e+18	3.53e+17	1.10e+13	1.04e+6
	\$3.50	4.66e+29	3.45e+8	2.27e+6	82207.00	31.82
	\$4.00	9.03e+10	14218.00	5872.80	168.47	9.32
	\$4.50	4.09e+7	48.57	79.32	20.63	1.26
		Non-Inf.	LB-1	LB-2	LB-3	
		1.18e-6	3.52e+11	5415.40	6.38e-6	

Table 13
Bayes Factors – Approx. Optimal 1 vs. Approx. Optimal 2 Models

		T_0				
		200	2000	4000	6000	8000
\bar{w}_0	\$2.50	2.42e+22	2.61e+13	2.26e+23	3.71e+10	1.61e+8
	\$3.00	1.31e+15	6.16e+11	9.62e+11	2.89e+8	2812.40
	\$3.50	8.58e+18	3.65e+5	4118.30	4242.00	4.59
	\$4.00	4.93e+5	57.74	1205.20	47.28	0.65
	\$4.50	2.67e+5	15.83	23.61	9.13	0.17
		Non-Inf.	LB-1	LB-2	LB-3	
		3.29e-6	2.44e+12	2.75e+5	4.99e-11	