Learning from Unemployment*

Francisco Gonzalez
Department of Economics
University of Calgary

Shouyong Shi
Department of Economics
University of Toronto

June 2005

Abstract

In this paper we analyze unemployed workers' learning during search about the aggregate matching efficiency in the market. Each worker chooses whether to participate in the market and which submarket to search in. A submarket is described by a wage and a job finding probability, with a higher wage being accompanied by a lower job finding probability. After each period of search, an unemployed worker updates his belief about the matching efficiency. We show that, as the number of past search failures increases, a worker's potential wage falls and the likelihood for the worker to exit the market rises. We also find that an increase in the unemployment benefit can *simultaneously* account for (i) a lower flow from unemployment to employment; (ii) a lower flow from unemployment to out of the labor force; (iii) a higher unemployment duration; and (iv) a lower probability for the workers who are out of the labor force to re-enter the labor force. These results are useful for explaining the differences in unemployment and labor force participation between the US and continental European countries.

* Very preliminary. Please do not quote.

1. Introduction

It is well known that continental European countries have significantly higher unemployment rates and longer unemployment duration than the US. Also, labor market participation is lower in continental Europe than in the US. These facts are puzzling. If European workers' high valuation of leisure accounts for the low participation, then it should also increase the opportunity cost of job search, thus reducing the unemployment rate and unemployment duration. If generous unemployment benefits in Europe account for the high unemployment rate and duration, then they should also induce higher participation. To understand the facts, it seems important to analyze workers' job search decisions.

In this paper we focus on unemployed workers' learning during search about the matching efficiency in the aggregate matching function. Workers are ex ante identical and unemployment histories are private information. At the beginning of each period, a non-employed worker receives a shock to the value of leisure. These shocks are *iid* across agents and over time. After the shock, a non-employed worker decides whether to search in the period or to enjoy leisure and stay out of the labor force. A worker who decides to search chooses the submarket in which to search. A submarket is described by a wage and a job finding probability, with a higher wage being accompanied with a lower job finding probability. After each period of search, an unemployed worker updates his belief about the matching efficiency. Consequently, beliefs will diverge as the workers' search outcomes differ. Also, workers who stay out of the labor force do not receive new information about the matching efficiency, and they do not receive the unemployment benefit.

We show that in the competitive search equilibrium, as the number of past search failures increases, a worker becomes more pessimistic about the matching efficiency. This will make the worker more likely to choose to stay out of the labor force in the future.

Also, if the worker continues to search in the next period, he will be more likely to search in a submarket in which the job finding rate will be higher. Because such a submarket comes with a lower wage, a worker's permanent income decreases with the duration of unemployment.

We also find that an increase in the unemployment benefit has the following effects on workers' decisions in the labor market. First, a worker is more likely to stay unemployed to search for a job. Second, a worker will search for higher wages which have lower job finding probabilities. Third, when workers exit the labor force from unemployment, their beliefs about the matching efficiency will be lower than if the unemployment benefit is lower, and hence they will be less likely to re-enter the labor force in the future. These effects imply that a higher unemployment benefit leads to (i) a lower flow from unemployment to employment; (ii) a lower flow from unemployment to out of the labor force; (iii) a higher unemployment duration; and (iv) a lower probability for the workers who are out of the labor force to re-enter the labor force.

These results are useful for explaining the differences between the US and continental European countries in unemployment and labor market participation. In particular, the theory suggests that more generous unemployment benefits in European countries can account *simultaneously* for higher unemployment, longer unemployment duration, and lower labor force participation in these countries. Learning during employment is important for this theory because, without learning, high unemployment benefits would lead to high unemployment but also high labor force participation.

There are other theories that also explain why workers' wages and participation fall during non-employment. For example, workers' skills may deteriorate during non-employment and, as the skills deteriorate, the workers will obtain lower wages and will be less likely to participate in the labor force. Although not exclusive to this alternative theory, our theory's testable implications are distinguishable. In particular, the alternative theory does not distinguish an unemployed worker from a worker out of the labor force in the sense that both workers' skills deteriorate during non-employment, while the distinction between the two is important in our theory. Only when a worker searches for a job does he gain new information about the matching efficiency and only when he fails to find a match will he become more pessimistic in the future search. Thus, in our theory, a worker who has been searching for a number of periods will behave very differently from a worker who has stayed out of the labor force for these periods. Moreover, the unemployment benefit affects workers' learning process in our theory by affecting their decisions to participate and to search in particular submarkets. In the alternative theory, the deterioration of a worker's skills does not have any link to the unemployment benefit, provided that the worker is non-employed.

2. The Model

2.1. Agents, Markets and Matching

Time is discrete and all agents discount the future at a rate $r > 0$. There are a large number of workers. A worker is either employed or non-employed. A non-employed worker is either unemployed or out of the labor force. When employed, a worker produces $y > 0$ units of goods. When unemployed, a worker gets an unemployment benefit $b > 0$ and searches for a job. A worker who is out of the labor force enjoys leisure, which yields a benefit $l > 0$. The value of leisure is a random variable whose value lies in $[\underline{l}, \bar{l}]$, where $0 < \underline{l} < \bar{l}$. At the beginning of each period, a non-employed worker draws a value of l from the distribution $F(.)$. Assume that the corresponding density function, F', is continuously differentiable over $[\underline{l}, \bar{l}]$. The draws are identically and independently distributed across agents and over time. After observing the value of l, an unemployed decides whether to stay unemployed or to get out of the labor force, while a worker out of the labor force decides whether to stay out of the labor force or to enter unemployment to search.

Let u be the number of unemployed workers and v the total number of vacancies in a period. The aggregate matching function is $muv/(u+v)$, where $m > 0$ is the "efficiency" of matching. Denote

$$\lambda = \frac{1}{1 + u/v}.$$

We will refer to λ as the tightness of the market, where a higher λ means a tighter market for the firms. Then, the job finding probability is λm, and the probability with which a vacancy finds a match is $(1-\lambda)m$. Let the domain of λ be $[0, \bar{\lambda}]$, where $\bar{\lambda} \in (0, 1)$.

The matching efficiency, m, is unknown to individual workers and firms. It has two possible values, m_H and m_L, where $m_H \in (0, 1)$ and $m_L \in (0, m_H)$. One way to interpret the problem is as follows. The economy consists of a continuum of markets, a fraction p of which exhibit $m = m_H$ and the remaining fraction of which have $m = m_L$. Agents are assigned to one of these markets at random. They can choose whether or not to participate in the market, but they cannot switch markets. Accordingly, agents initially have a common prior on their type. That is, an agent's initial belief is that he is assigned to an H market with probability p and to an L market with probability $1 - p$.[1]

In each of the markets described above, there are submarkets. A submarket is characterized by a tightness, λ, and a wage level $W(\lambda)$. The wage function $W(.)$ describes wages across the submarkets, and it is public knowledge. Firms and workers can choose which submarket to enter. The equilibrium wage in a submarket clears the submarket in the sense that the induced entry of firms and workers is consistent with the tightness in

[1] Note that there is some heterogeneity, ex post, since workers are assigned to one of the two markets. In this sense, observing equilibrium wage dispersion may not be that surprising. However, the pattern of wages which a given worker can obtain potentially in different periods is interesting.

that submarket. We will determine the equilibrium wage in each submarket later, which will satisfy the following assumption:

Assumption 1. *For all* $\lambda \in [0, \bar{\lambda}]$, *the function* $W(\lambda)$ *is twice continuously differentiable and it has the following properties: (i)* $0 < W(\lambda) \leq y$; *(ii)* $W'(\lambda) < 0$; *and (iii)* $W''(\lambda) \leq 2W'(\lambda)/(1-\lambda)$.

Part (i) of the assumption is evident. Part (ii) requires that the wage in a submarket that is tighter for firms should have a lower wage. This is necessary in order to induce firms to enter the submarket. Part (iii) requires that the wage function be sufficiently concave, and it facilitates some technical aspects of the analysis.

For a worker, entering a submarket (to search) for a period makes the worker entitled to the unemployment benefit in that period, $b > 0$. At the same time, the worker foregoes the benefit of leisure in that period, l. For a firm, entering a submarket for a period requires the firm to incur the vacancy cost, $c > 0$, which is the same for all submarkets. As usual, the expected profit of entry into any submarket is zero.

Matched workers and firms will stay matched forever and they will be replaced by crones with the initial beliefs.

2.2. Updating Beliefs

Agents update their beliefs on m after observing whether or not they have a match. The updating depends on the particular submarket into which the agent just searched. To describe the updating process, it is convenient to express the beliefs in terms of their expected type. Let the initial prior expectation of m be $\mu_0 \in (m_L, m_H)$, for all agents (workers and firms). We examine the updating process of a worker.

Let μ denote a worker's mean beliefs at the beginning of an arbitrary search period.

Accordingly, the prior probability distribution of m is $(P(m_H), 1 - P(m_H))$, where

$$P(m_H) = \frac{\mu - m_L}{m_H - m_L}.$$

Note that the prior distribution of m is Bernoulli, and

$$E(m) = \mu,$$

$$Var(m) = (m_H - \mu)(\mu - m_L).$$

Let $z \in \{0, 1\}$ denote a worker's matching outcome in the current period, where $z = 1$ if the worker finds a job and $z = 0$ if he does not find a job. Recall that a worker search in the submarket with a tightness λ finds a match with probability $m\lambda$. Then, the conditional distribution of m is given by

$$P(m \mid z = 0) = \frac{1 - m\lambda}{1 - \mu\lambda} P(m), \quad m = m_L, m_H,$$

$$P(m \mid z = 1) = \frac{m}{\mu} P(m), \quad m = m_L, m_H.$$

Because the conditional distribution of m is Bernoulli, then for $z \in \{0, 1\}$, the mean and variance of m conditional on z are:

$$E(m \mid z) = m_H P(m_H \mid z) + m_L(1 - P(m_H \mid z))$$

$$Var(m \mid z) = (m_H - m_L)^2 P(m_H \mid z)(1 - P(m_H \mid z)).$$

Note that, since $\lambda \leq \bar{\lambda} < 1 < 1/m_H$, we have $P(m_H \mid z = 0) > 0$ for all $\mu > m_L$. Thus, if the initial mean belief μ_0 exceeds m_L, then $E(m|z) > m_L$ for both $z = 0$ and $z = 1$.

This updating process has two preliminary properties. First, the sequence $\{E(m)\}$ is a Markov process. Second, a worker's mean beliefs $E(m)$ are a sufficient statistic for a worker's unemployment history. One implication of these properties is as follows. Consider two unemployed workers who draw the same value of leisure for the current period and who have the same beliefs, μ. One has switched between unemployment and out of the labor force many times in the past, while the other worker entered unemployment a few periods ago and has remained in unemployment continuously in the last few periods. Although these two workers followed different paths to reach the same beliefs, they will choose the same submarket to enter and update their beliefs in the same way.

Search is more informative than staying out of the labor market. This is evident from the fact that search introduces variance in the worker's posterior beliefs. Moreover, conditional on search, a higher λ is more informative in the sense that it causes a mean-preserving spread in the distribution of the posterior expectation $E(m \mid z)$. To see this, note that ex ante z is a random variable, and so is the posterior expectation $E(m \mid z)$. The mean of this posterior expectation is $E(m \mid z) = E(m) = \mu$, which is unaffected by λ. The variance of the posterior expectation is:

$$Var(E(m \mid z)) = (\mu - m_L)^2 \left[\frac{\lambda m_H^2}{\mu} + \frac{(1 - m_H \lambda)^2}{1 - \mu\lambda} - 1 \right].$$

This variance increases with λ. Thus, conditional on search, a higher λ generates outcomes that are more informative about the matching efficiency.

The informational content of λ is asymmetric with respect to the matching outcome. For a worker who succeeds in finding a match, the posterior, $P(m \mid z = 1)$, is not a function of λ. The posterior mean belief in this case is $E(m|z = 1) = m_H + m_L - m_H m_L/\mu$, which is also independent of λ. Therefore, a worker's choice of submarket, λ, does not affect the information contained in successful search outcomes. In contrast, for a worker who fails to find a match, the posterior, $P(m \mid z = 0)$, decreases with λ. That is, the higher λ of a submarket in which a worker searches for a job, the more the worker will reduce the posterior on the matching efficiency after he fails to find a match. This is because finding a match in a submarket with a higher λ is supposed to be easier, and a failure to find a match there should induce the worker to revise the beliefs downward sharply.[2]

Given the above asymmetry, it is useful to separate the updating process of a worker who fails to find a job. Denote

$$H(\mu,\lambda) \equiv E(m \mid z=0) = m_H - \frac{1-m_L\lambda}{1-\mu\lambda}(m_H-\mu). \tag{2.1}$$

Then, the updating process of an unemployed worker is $\mu_{+1} = H(\mu,\lambda)$, where the subscript $+1$ indicates the next period. We can verify the following lemma:

Lemma 2.1. *The function H has the following properties: (i) $H_1 > 0$; (ii) $H_2 < 0$; (iii) $H_{11} = \frac{2\lambda}{1-\mu\lambda}H_1 > 0$, $H_{22} = \frac{2\mu}{1-\mu\lambda}H_2 < 0$; (iv) $\mu(1-\mu\lambda)H_{12} - H_2 - \mu^2 H_1 = -m_H m_L$.*

The property (i) states simply that, for any given λ, a worker with higher mean expectations will also have higher posterior expectations. The property (ii) states that a higher λ reduces the worker's posterior expectations after the worker fails to find a match, which is discussed above. Properties (iii) and (iv) will be useful later. Notice that property (ii) implies that $\mu_{+1} < \mu$ for all $\lambda > 0$. Thus, a worker's mean beliefs about the matching efficiency decreases over time as the number of search failures increases.

2.3. The Value of Search

Consider a non-employed worker who enters a period with the mean beliefs, μ. Let $J(\mu)$ be his value function before he observes the value of leisure in the period, l. When l is low, the worker will search and hence will be unemployed in the period. Because search foregoes the value of leisure and because the value of leisure is independent over time, the value function of search is independent of l. Denote this value function as $V(\mu)$. Then,

$$V(\mu) = max_{\lambda\in[0,\bar{\lambda}]}U(\mu,\lambda), \tag{2.2}$$

[2] This asymmetry of the role of λ in the posterior will remain as long as the job finding probability has the form $\lambda g(m)$, where g is an increasing function.

where

$$U(\mu,\lambda) \equiv b + \mu\lambda\frac{W(\lambda)}{r} + (1-\mu\lambda)\frac{J(H(\mu,\lambda))}{1+r}. \tag{2.3}$$

In this payoff function, λ is the submarket in which the worker chooses to search in the period, and $\mu\lambda$ is the expected job finding probability. Notice that, if a worker finds a job, he will stay employed forever, and so the present value is W/r. Denote the worker's optimal decision as $\lambda^* = \Lambda(\mu)$.

If the worker stays out of the labor force in the period, his value function is $l + \frac{J(\mu)}{1+r}$. Thus, the worker chooses to search in the period if and only if $l \leq V(\mu) - \frac{J(\mu)}{1+r}$. Define l^* as follows:

$$l^* = L(\mu) \equiv V(\mu) - \frac{J(\mu)}{1+r}. \tag{2.4}$$

Then, a non-employed worker chooses to search in the period if and only if $l \leq l^*$.

Before observing l in the period, the value function is given as follows:

$$J(\mu) = \int_{\underline{l}}^{\bar{l}} max\left\{l + \frac{J(\mu)}{1+r}, V(\mu)\right\} dF(l). \tag{2.5}$$

We can define a mapping T as follows:

$$T(J)(\mu) = \int_{\underline{l}}^{\bar{l}} max\left\{l + \frac{J(\mu)}{1+r}, V(\mu)\right\} dF(l). \tag{2.6}$$

Notice that J appears on the right-hand side in both arguments inside the maximum operator. Then, the value function J is a fixed point of T. That is,

$$J(\mu) = T(J)(\mu). \tag{2.7}$$

It is easy to verify that T satisfies the sufficient conditions for contraction mapping (see Theorem 3.3 in Stokey and Lucas, p.54). The standard argument shows that there is a unique solution J to (2.7) and that J is bounded and continuous on $[m_L, m_H]$ (see Theorem 4.6 in Stokey and Lucas, p.79). Similarly, there is a unique function V satisfying (2.2) and V is bounded and continuous on $[m_L, m_H]$.

2.4. Firms' Behavior and the Definition of an Equilibrium

Firms also choose which submarket to enter to post vacancies and, after observing the matching outcome, they update their beliefs. For simplicity, let us assume that all firms have the same initial mean belief about the matching efficiency and that this initial belief is the same as the workers', μ_0. The updating process of a firm is similar to that of the workers'. Let λ^v be the submarket which a firm enters. If the firm finds a match, its posterior expectation of m does not depend on λ^v. If the firm fails to find a match, then the posterior expectation of m decreases in λ. Let μ_1^v denote the firm's expectations of m after the firm fails to find a match with one period's search. Because the firm's matching probability is $m(1-\lambda^v)$, then

$$\mu_1^v = H(\mu_0, 1-\lambda^v).$$

Here, H is defined in (2.1). Because $H_2 < 0$ and $H(\mu, 0) = \mu$, then $\mu_1^v < \mu_0$ for all $\lambda^v < 1$.

Let $J^v(\mu^v)$ be the value function of a vacancy given that the firm's mean belief at the beginning of a period is μ^v. With free entry, $J^v(\mu_0) = 0$. Because $\mu_1^v < \mu_0$ for all $\lambda^v < 1$, as explained above, then $J^v(\mu_1^v) < 0$. That is, a firm will always exit the market after one period of search if the search fails to find a match in the period. This result allows us to simplify a firm's value function as follows:

$$J^v(\mu_0) = \max_{\lambda^v} \left[-c + \mu_0(1-\lambda^v)\frac{y - W(\lambda^v)}{r} \right].$$

The first-order condition for the firm's optimal choice, λ^{v*}, involves the wage function and its derivative. This can be alternatively viewed as a differential equation for the wage function. Without an initial condition, there are a continuum of solutions to the differential equation. This indeterminacy simply says that there are many level of λ^v that are optimal for the firm. Put differently, a firm is willing to enter into any submarket, provided that the wage in the submarket is consistent with the free-entry condition.

To determine the wage function, we use the free-entry condition, $J^v(\mu_0) = 0$. Using the above expression for the firm's value function, we can write this condition as follows:

$$W(\lambda) = y - \frac{rc/\mu_0}{1-\lambda}. \tag{2.8}$$

It is easy to verify that this function indeed satisfies Assumption 1. In fact, part (iii) of the assumption holds as equality.

We can now define the equilibrium.[3] An equilibrium consists of workers' decisions (l^*, λ^*), firms' decision λ^{v*}, and a wage function $W(\lambda)$, that meet the following requirements: (i) Given the worker's belief at the beginning of a period and given the wage function, a worker's participation decision obeys the optimal rule $l^* = L(\mu)$ and the choice of the submarket obeys the optimal rule $\lambda^* = \Lambda(\mu)$; (ii) Given the initial belief μ_0 and the wage function, a firm's choice λ^{v*} is optimal; (iii) A worker updates the beliefs according to $\mu_{+1} = H(\mu, \lambda^*)$ and a firm according to $\mu_1^v = H(\mu_0, \lambda^{v*})$; (iv) Consistency: For every λ, the sum of all workers who choose $\Lambda(\mu) = \lambda$ divided by the sum of all firms who choose $\lambda^{v*} = \lambda$ is equal to $\lambda^{-1} - 1$; (v) Free-entry: The wage function satisfies (2.8).

3. Worker's Decisions

We examine a worker's decisions in detail. There are two decisions. One is whether to participate in the labor market and the other is which submarket to search for a match. To analyze these decisions, we impose the following assumption:

Assumption 2. *The condition, $W(\lambda) > \frac{r}{1+r} J(H(\mu, \lambda))$, is satisfied for all $\mu \in [m_L, m_H]$ and $\lambda \in [0, \bar{\lambda}]$. Moreover, workers and firms commit to accepting all successful matches.*

The first part of the assumption ensures that accepting a successful match gives the worker a higher present value of utility than rejecting the match and revising the beliefs

[3] The steady state conditions on worker flows need to be incorporated here.

according to H. However, this condition may not be sufficient to guarantee that the worker will accept the match. This is because a worker who succeeds in finding a match does not update the beliefs according to H, and so the value of rejecting the match and continuing to search next period is not given by $J(H)/(1+r)$. The second part of the assumption guarantees that the worker and the firm will always accept the match.

3.1. The Participation Decision

As analyzed above, a non-employed worker chooses to search if and only if the value of leisure in the period satisfies $l \leq l^* = L(\mu)$. To know how this reservation value of leisure depends on the beliefs, μ, we first establish the following lemma:

Lemma 3.1. *V is strictly increasing and J is non-decreasing. If $L(\mu) > \underline{l}$ for all μ, then J is strictly increasing.*

Proof. Let $C'[m_L, m_H]$ be the set containing all bounded, continuous and non-decreasing functions on $[m_L, m_H]$. Let J_0 be any function in $C'[m_L, m_H]$ and use it to serve the role of J in (2.3) and (2.6). Because $H_1 > 0$, it is easy to show that $V(.)$ is *strictly* increasing under Assumption 2. Thus, T maps functions in $C'[m_L, m_H]$ into functions in $C'[m_L, m_H]$. By the argument of contraction mapping, the fixed point of T, J, is non-decreasing. Using this fixed point in the expression of U, we find that V is strictly increasing. Then, we can write the mapping T as follows:

$$T(J)(\mu) = F(L(\mu))V(\mu) + [1 - F(L(\mu))]\frac{J(\mu)}{1+r}.$$

Suppose $L(\mu) > \underline{l}$ for all μ, so that $F(L(\mu)) > 0$. Because V is strictly increasing and J is non-decreasing, then $J = T(J)$ is strictly increasing. QED

Next, we use (2.2) and (2.5) to express the value functions explicitly as follows:

$$J(\mu) = \frac{1+r}{r}\left[F(l^*)l^* + \int_{l^*}^{\bar{l}} l dF(l)\right], \tag{3.1}$$

$$V(\mu) = \left[1 + \frac{F(l^*)}{r}\right] l^* + \int_{l^*}^{\bar{l}} l dF(l). \tag{3.2}$$

Notice that the right-hand side of (3.2) is strictly increasing in l^*. Because V is strictly increasing in μ, (3.2) solves l^* as a strictly increasing function of μ. The following proposition is evident:

Proposition 3.2. *$L(\mu)$ is a strictly increasing function. Moreover, $L(\mu)$ is (twice) continuously differentiable if and only if $V(\mu)$ is (twice) continuously differentiable.*

Because $L(\mu)$ is increasing, then a worker will be more likely to choose to search if he has higher mean beliefs about the matching efficiency in the market. As the number of periods in which the worker has been unemployed increases, the mean beliefs deteriorate and so the worker becomes more likely to exit from the labor force.

3.2. The Choice of the Submarket

Once a worker chooses to search in the period, the worker chooses the submarket λ^* in which to search. The choice of the submarket involves two considerations. The first is the familiar trade-off between the wage and the matching probability, because a submarket with a high λ has a high matching probability but a low wage. The second consideration is about the information content of the search outcome. To see this, recall that a high λ induces a mean-preserving spread in the distribution of the worker's posterior beliefs. In this sense, a high λ makes the matching outcome more informative about the matching efficiency. This forthcoming information will be useful for the search decisions in the future when (and only when) the worker fails to find a match in the current period.

The idea that a more informative search outcome is valuable to the worker can be captured formally by value functions that are strictly convex in the mean beliefs. The following lemma establishes the convexity (see Appendix A for a proof):

Lemma 3.3. *V is strictly convex and J is convex. Thus, both V and J are almost everywhere twice differentiable with almost everywhere continuous first derivative. If $L(\mu) > \underline{l}$ for all $\mu \in [m_L, m_H]$, then J is strictly convex.*

With strictly convex value functions, a worker who chooses to enter a submarket with a high λ makes a trade-off of the benefits of a higher matching probability and a more informative search outcome against the low wage. It is not clear, a priori, how the optimal choice λ^* depends on the worker's beliefs. To find this dependence, we need to characterize the optimal choice explicitly, e.g., using the first-order condition.

However, an explicit characterization may encounter potential difficulties created by the convexity of the value functions. First, because H is also convex in λ, the function $U(\mu, \lambda)$ may not necessarily be concave in λ. However, concavity of U in λ is a desirable property because it ensures that the optimal choice λ^* is unique for each given μ, and hence that λ^* is continuous in μ. Second, although the value functions are twice differentiable almost everywhere, it is well known in the literature on optimal learning that they may fail to be differentiable at certain points (e.g., Easley and Kiefer (1988), Kiefer (1989), Balvers and Cosimano (1993)). To facilitate the analysis, we impose the following assumption:

Assumption 3. *For all μ, (i) $U(\mu, \lambda)$ is strictly concave in λ, and (ii) the solution λ^* is interior.*

Under this assumption, we can establish the following lemma (see Appendix B for a proof):

Lemma 3.4. *Under Assumptions 1, 2 and 3, the functions $V(\mu)$, $J(\mu)$ and $L(\mu)$ are all differentiable on $[m_L, m_H]$. In particular,*

$$V'(\mu) = U_1(\mu, \lambda^*(\mu)). \tag{3.3}$$

With differentiability of the value functions and Assumption 3, the optimal choice of λ^* obeys the first-order condition:

$$U_2(\mu, \lambda^*) = 0. \tag{3.4}$$

The following proposition describes how the optimal choice depends on the beliefs:

Proposition 3.5. *The optimal choice* $\lambda^* = \Lambda(\mu)$ *is strictly decreasing in* μ. *As a result, the wage* $W(\Lambda(\mu))$ *is strictly increasing in* μ.

Proof. Differentiating (3.4), we have $\Lambda'(\mu) = -U_{12}/U_{22}$. Because $U_{22} < 0$ by Assumption 3, then $\Lambda'(\mu) < 0$ if and only if $U_{12} < 0$. Writing (3.4) explicitly and using it to substitute for $(W + \lambda W')$, we can compute:

$$U_{12} = \frac{J'(H)}{1+r}\left[\mu(1-\mu\lambda)H_{12} - H_2 - \mu^2 H_1\right] + (1-\mu\lambda)\frac{J''(H)}{1+r}H_1 H_2.$$

Because $J' > 0$, $J'' > 0$, $H_1 > 0$ and $H_2 < 0$, we can use part (iv) of Lemma 2.1 to verify $U_{12} < 0$. Since $W(\lambda)$ is strictly decreasing, then $W(\Lambda(\mu))$ is strictly increasing in μ. QED

Recall that a high λ means that the submarket is tight for the firms, i.e., the ratio of searching workers to vacancies is low. The above proposition states that, when a worker is more optimistic about the matching efficiency in the market, he will choose to enter a submarket which offers a higher wage and which is more congested with workers. As the number of periods in which the worker has remained in unemployment increases, the worker will become more pessimistic about the matching efficiency, and hence will search for lower wages and in less congested submarkets.

3.3. Distributions of Worker's Search Experiences and Beliefs

We can now characterize the dynamics of a worker's search and beliefs. Consider an arbitrary period. Let τ denote the duration of past searches, i.e., the number of periods in which the worker has searched without finding a match up to (but not including) the current period. Note that the worker may not have searched continuously for τ periods in the past. Given μ, the worker will choose to search in the period with probability $F(L(\mu))$. If he chooses to search, then he will enter the submarket $\lambda = \Lambda(\mu)$ and will find a match with probability $1 - \mu\lambda$. Thus, τ_{+1} will be distributed as follows:

$$\tau_{+1} = \begin{cases} \tau + 1, & \text{with prob. } F(L(\mu))\,[1 - \mu\Lambda(\mu)] \\ \tau, & \text{with prob. } 1 - F(L(\mu))\,[1 - \mu\Lambda(\mu)]\,. \end{cases}$$

Similarly, the distribution of μ_{+1}, conditional on the outcome that the worker fails to find a match in the current period, will be:

$$\mu_{+1} = \begin{cases} H(\mu, \Lambda(\mu)), & \text{with prob. } \frac{F(L(\mu))[1-\mu\Lambda(\mu)]}{1-F(L(\mu))\mu\Lambda(\mu)} \\ \\ \mu, & \text{with prob. } \frac{1-F(L(\mu))}{1-F(L(\mu))\mu\Lambda(\mu)}. \end{cases}$$

Starting with a given μ_0 and $\tau_0 = 0$, the above processes induce the distribution of (τ, μ) in any period.

4. The Effects of the Unemployment Benefit

The unemployment benefit, b, affects the unemployment duration and labor force participation. To do so, let us modify the notation $U(\mu, \lambda)$ to $U(\mu, \lambda, b)$, $J(\mu)$ to $J(\mu, b)$, and $V(\mu)$ to $V(\mu, b)$. Notice that the unemployment benefit does not directly affect the updating process of the beliefs, which is still described by $\mu_{+1} = H(\mu, \lambda)$. Modifying (2.3), (2.5) and (2.2), we have:

$$U(\mu, \lambda, b) = b + \mu\lambda \frac{W(\lambda)}{r} + (1 - \mu\lambda) \frac{J(H(\mu, \lambda), b)}{1 + r}, \tag{4.1}$$

$$V(\mu, b) = max_{\lambda} U(\mu, \lambda, b), \tag{4.2}$$

$$J(\mu, b) = \int_{\underline{l}}^{\bar{l}} max \left\{ l + \frac{J(\mu, b)}{1 + r}, V(\mu, b) \right\} dF(l). \tag{4.3}$$

Similarly, we can modify (2.4) to obtain the reservation value of leisure as $l^* = L(\mu, b)$. A worker's optimal choice of the submarket is $\lambda^* = \Lambda(\mu, b)$, which satisfies the following first-order condition:

$$U_2(\mu, \lambda^*, b) = 0. \tag{4.4}$$

Let us first examine how the unemployment benefit affects a worker's decision on participation. An increase in the unemployment benefit increases the likelihood of participation if the reservation value of leisure increases in the unemployment benefit, i.e., if $L_2(\mu, b) > 0$. Because (3.1) and (3.2) are valid after obvious modifications of the notation, differentiating those conditions yields:

$$L_2(\mu, b) = \frac{\partial l^*}{\partial b} = \frac{r}{r + F^*} V_2(\mu, b) = \frac{r/F^*}{1 + r} J_2(\mu, b), \tag{4.5}$$

where $F^* = F(l^*)$. Thus, the reservation value of leisure increases in the unemployment benefit if and only if the value functions do. The following proposition states the intuitive result on participation:

Proposition 4.1. $V_2(\mu, b) > 0$, $L_2(\mu, b) > 0$, *and so an increase in the unemployment benefit increases a worker's likelihood of participation in the labor force. Also,* $J_2(\mu, b) \geq 0$, *with strict inequality if* $L(\mu, b) > \underline{l}$ *for all* μ.

Proof. From (4.2) we have:

$$V_2(\mu, b) = 1 + (1 - \mu\lambda^*) \frac{F^*}{r + F^*} V_2(H, b). \tag{4.6}$$

Here we have used (4.5) to substitute $J_2(H, b)$ with a function of $V_2(H, b)$. The above equation implies that, if $V_2(H, b) \geq 0$, then $V_2(\mu, b) > 0$. By induction, $V_2(\mu, b) > 0$. Then, (4.5) implies $L_2(\mu, b) > 0$ and $J_2(\mu, b) \geq 0$. Moreover, if $F^* > 0$, then $J_2(\mu, b) > 0$. QED

The unemployment benefit also affects the submarket in which a worker searches for a job. An intuitive outcome is that an increase in the unemployment benefit encourages a worker to enter a submarket with a high wage and a low job finding probability. That is, the optimal choice, $\lambda^* = \Lambda(\mu, b)$, decreases in b for given beliefs. To see whether this intuitive outcome occurs in the model, differentiate the first-order condition (4.4) to obtain:

$$\Lambda_2(\mu, b) = -U_{23}/U_{22}.$$

Under Assumption 3, $U_{22} < 0$. Thus, λ^* decreases in b if and only if $U_{23} < 0$. We verify this feature in Appendix C and summarize the result as follows:

Proposition 4.2. $\Lambda_2(\mu, b) < 0$. *Thus, for given beliefs, an increase in the unemployment benefit induces workers to search for high wages.*

In the above analysis, we have fixed the worker's beliefs. In equilibrium, however, a worker's beliefs are also affected by the unemployment benefit. To account for the total effects of the unemployment benefit on a worker's decisions of search and participation, let us index (μ, λ^*, l^*) by τ – the number of past periods in which an individual worker has searched without finding a match. Starting with a fixed μ_0, we can compute the sequence of (μ, λ^*, l^*) for the worker as follows:

$$\mu_\tau(b) = H(\mu_{\tau-1}(b), \lambda^*_{\tau-1}(b)),$$

$$\lambda^*_\tau(b) = \Lambda(\mu_\tau(b), b), \quad l^*_\tau(b) = L(\mu_\tau(b), b).$$

We describe the total effects of the unemployment benefit on the above sequence in the following proposition:

Proposition 4.3. $\mu'_\tau(b) > 0$ *for all* $\tau \geq 1$. *Also,* $\lambda^{*\prime}_\tau(b) < 0$ *and* $l^{*\prime}_\tau(b) > 0$ *for all* $\tau \geq 0$.

Proof. Start with $\tau = 0$. Using Propositions 4.1 and 4.2, we have:

$$\lambda^{*\prime}_0(b) = \Lambda_1 \mu'_0(b) + \Lambda_2 = \Lambda_2 < 0,$$
$$l^{*\prime}_0(b) = L_1 \mu'_0(b) + L_2 = L_2 > 0.$$
$$\mu'_1(b) = H_1 \mu'_0(b) + H_2 \lambda^{*\prime}_0(b) = H_2 \lambda^{*\prime}_0(b) > 0.$$

Thus, the proposition holds for $\tau = 0$. Suppose that it holds for an arbitrary $\tau \geq 0$. Because $\Lambda_1 < 0$, $\Lambda_2 < 0$, $L_1 > 0$, $L_2 > 0$, $H_1 > 0$ and $H_2 < 0$, we have:

$$\lambda^{*\prime}_{\tau+1}(b) = \Lambda_1 \mu'_\tau(b) + \Lambda_2 < 0,$$
$$l^{*\prime}_{\tau+1}(b) = L_1 \mu'_\tau(b) + L_2 > 0.$$
$$\mu'_{\tau+1}(b) = H_1 \mu'_\tau(b) + H_2 \lambda^{*\prime}_\tau(b) > 0.$$

That is, the proposition holds for $\tau + 1$. By induction, the proposition holds for all $\tau \geq 0$. QED

To phrase the effects in the above proposition in a different way, consider two economies, A and B, that differ only in the unemployment benefit. Suppose $b_A > b_B$. Take two workers, one from each economy, who have had the same number of past search failures. Then, relative to the worker in economy B, the worker in economy A has more optimistic beliefs about the matching efficiency, is more likely to participate in the labor market, and is more likely to search in a submarket with a high wage and a low job finding probability.

This comparison has several implications. First, an increase in the unemployment benefit reduces the flow from unemployment to out of the labor force. In the two economies described above, workers who are searching for jobs in economy A are less likely to choose to stay out of the labor market. Second, an increase in the unemployment benefit reduces the flow from unemployment to employment. This is because workers who search for jobs in economy A choose to search for higher wages which necessarily come with lower job finding probabilities. Third, the unemployment benefit increases the unemployment duration. This is a consequence of the first two effects, because unemployed workers in economy A are more likely to stay unemployed than in economy B, and because they choose to search for jobs that are more difficult to find. More precisely, the distribution of τ in economy A dominates the distribution in economy B in the sense of the first-order stochastic dominance.

However, the economy with a higher unemployment benefit (i.e., economy A) does not necessarily have a higher rate of labor force participation in the steady state, despite the fact that the flow from unemployment to out of the labor force is lower in such an economy. The reason is that the probability of re-entering the labor force in economy A is not necessarily higher than in economy B. In fact, the opposite may be true. To explain this feature, notice that the worker's beliefs are indexed by τ and that these beliefs deteriorate as τ increases. Because the distribution of τ in economy A dominates that in economy B, the workers in economy A who are out of the labor force have had, on average, a longer unemployment duration than those in economy B, and hence they are more pessimistic about the matching efficiency. These pessimistic beliefs discourage such workers in economy A to re-enter the labor force, despite the higher unemployment benefit.

We summarize the above discussions as follows:

Theorem 4.4. *(Conjectured) Consider two economies that differ only in the unemployment benefit. Then the economy with a higher benefit will exhibit (i) a lower flow from unemployment to employment; (ii) a lower flow from unemployment to out of the labor force; (iii) a higher unemployment duration; and (iv) a lower probability for the workers who are out of the labor force to re-enter the labor force.*

These effects indicate the difference in the unemployment benefit between continental European countries and the US may explain the differences in the labor market performances between these countries. In particular, higher unemployment benefits may be important for explaining why continental European countries have *both* a higher duration of unemployment and lower labor force participation.

It is important to emphasize the role of information in our analysis. If the matching efficiency were public knowledge, then a high unemployment benefit would increase, rather than decrease, the probability with which workers who are out of the labor force choose to re-enter the labor force. In this case, labor force participation, as well as unemployment, would be higher in an economy with a high unemployment benefit. In contrast, when the matching efficiency is unknown and workers update their beliefs about it, workers will be more likely to search for a long time but, once they get out of the labor force, they will also be more pessimistic and hence more reluctant to re-enter.

5. Conclusions

In this paper we analyze unemployed workers' learning during search about the aggregate matching efficiency in the market. Workers are ex ante identical and unemployment histories are private information. Each worker decides whether to participate in the market and to choose the submarket in which to search. A submarket is described by a wage and a job finding probability, with a higher wage being accompanied with a lower job finding probability. After each period of search, an unemployed worker updates his belief about the matching efficiency. We show that, as the number of past search failures increases, a worker becomes more pessimistic about the matching efficiency and will be more likely to search for low wages in the future.

We also find that an increase in the unemployment benefit leads to (i) a lower flow from unemployment to employment; (ii) a lower flow from unemployment to out of the labor force; (iii) a higher unemployment duration; and (iv) a lower probability for the workers who are out of the labor force to re-enter the labor force. Thus, more generous unemployment benefits in European countries can account *simultaneously* for higher unemployment, longer unemployment duration, and lower labor force participation in these countries.

Appendix

A. Proof of Lemma 3.3

We first prove the following lemmas.

Lemma A.1. *If J is strictly convex, then $U(\mu,\lambda)$ is strictly convex in μ for any given λ.*

Proof. Let μ_1 and μ_2 be two arbitrarily values in $[m_L, m_H]$, with $\mu_1 < \mu_2$. Let $\theta \in (0,1)$ be a number. Denote $\mu_\theta = \theta\mu_1 + (1-\theta)\mu_2$. We show that

$$U(\mu_\theta,\lambda) < \theta U(\mu_1,\lambda) + (1-\theta)U(\mu_2,\lambda).$$

Because $\partial H/\partial\mu > 0$, then $H(\mu_1,\lambda) < H(\mu_\theta,\lambda) < H(\mu_2,\lambda)$. Let

$$\sigma = \frac{H(\mu_2,\lambda) - H(\mu_\theta,\lambda)}{H(\mu_2,\lambda) - H(\mu_1,\lambda)}.$$

Then, $\sigma \in (0,1)$ and $\sigma H(\mu_1,\lambda) + (1-\sigma)H(\mu_2,\lambda) = H(\mu_\theta,\lambda)$. If J is strictly convex, then

$$J(H(\mu_\theta,\lambda)) < \sigma J(H(\mu_1,\lambda)) + (1-\sigma)J(H(\mu_2,\lambda)).$$

By the definition of U in (2.3), we have:

$$\begin{aligned} U(\mu_\theta,\lambda) &< b + \mu_\theta\lambda\tfrac{W(\lambda)}{r} + \tfrac{1-\mu_\theta\lambda}{1+r}\left[\sigma J(H(\mu_1,\lambda)) + (1-\sigma)J(H(\mu_2,\lambda))\right] \\ &= \theta U(\mu_1,\lambda) + (1-\theta)U(\mu_2,\lambda) + \tfrac{J(H(\mu_1))}{1+r}\Delta_1 + \tfrac{J(H(\mu_2))}{1+r}\Delta_2, \end{aligned}$$

where

$$\begin{aligned} \Delta_1 &= (1-\mu_\theta\lambda)\sigma - \theta(1-\mu_1\lambda), \\ \Delta_2 &= (1-\mu_\theta\lambda)(1-\sigma) - (1-\theta)(1-\mu_2\lambda). \end{aligned}$$

For $i, j \in \{1, 2, \theta\}$, we use (2.1) to compute:

$$\sigma = \frac{(\mu_2-\mu_\theta)(1-\mu_1\lambda)}{(\mu_2-\mu_1)(1-\mu_\theta\lambda)} = \frac{\theta(1-\mu_1\lambda)}{(1-\mu_\theta\lambda)}.$$

Now it is easy to see that $\Delta_1 = 0 = \Delta_2$. Therefore, U is strictly convex. QED

Lemma A.2. *If J is convex (not necessarily strict), then V is strictly convex.*

Proof. We prove the lemma separately for the case where J is strictly convex and for the case where J is weakly convex. For both cases, let μ_1 and μ_2 be two arbitrarily values in $[m_L, m_H]$, with $\mu_1 < \mu_2$. Let $\theta \in (0,1)$ be a number. Denote $\mu_\theta = \theta\mu_1 + (1-\theta)\mu_2$. We need to show that

$$V(\mu_\theta) < \theta V(\mu_1) + (1-\theta)V(\mu_2).$$

To simplify notation, denote $J_i = J(\mu_i)$ and $V_i = V(\mu_i)$, where $i \in \{1, 2, \theta\}$.

Suppose first that J is strictly convex. By the above lemma, $U(\mu, \lambda)$ is strictly convex in μ for given λ. Let λ_i^* be the solution to $max_\lambda U(\mu_i, \lambda)$, $i \in \{1, 2, \theta\}$. That is, $V(\mu_i) = U(\mu_i, \lambda_i^*)$. We have:

$$\begin{aligned} V(\mu_\theta) &= U(\mu_\theta, \lambda_\theta^*) \\ &< \theta U(\mu_1, \lambda_\theta^*) + (1-\theta)U(\mu_2, \lambda_\theta^*) \\ &\leq \theta U(\mu_1, \lambda_1^*) + (1-\theta)U(\mu_2, \lambda_2^*) \\ &= \theta V(\mu_1) + (1-\theta)V(\mu_2). \end{aligned} \tag{A.1}$$

The first (strict) inequality follows from the fact that U is strictly convex in μ and the second inequality from the fact that $U(\mu_i, \lambda) \leq U(\mu_i, \lambda_i^*)$ for all λ. Thus, V is strictly convex.

Now suppose that J is only weakly convex, i.e., J has some linear segments. If any two of $J(H(\mu_\theta, \lambda_\theta^*))$, $J(H(\mu_1, \lambda_\theta^*))$ and $J(H(\mu_2, \lambda_\theta^*))$ do not lie on the same linear segment of J, then the first inequality in (A.1) is still strict. Suppose that $J(H(\mu_\theta, \lambda_\theta^*))$, $J(H(\mu_1, \lambda_\theta^*))$ and $J(H(\mu_2, \lambda_\theta^*))$ all lie on the same linear segment of J. Temporarily denote this linear segment as $J(H) = A + BH$, where $B > 0$ by Lemma 3.1. Using (2.1), we can see compute:

$$(1-\mu\lambda)J(H) = (1-\mu\lambda)(A + Bm_H) - B(1 - m_L\lambda)(m_H - \mu).$$

This is linear in μ and λ. It is also differentiable in μ and λ. Restricting the values of μ to such that $J(H(\mu, \lambda))$ lies on the linear segment described above, then $U(\mu, \lambda)$ is strictly concave in λ under Assumption 1. Thus, the solution λ^* is unique and satisfies the following first-order condition:

$$0 = U_2(\mu, \lambda) = \mu \left[\frac{W + \lambda W'}{r} - A - Bm_H \right] + Bm_L(m_H - \mu).$$

Furthermore, we can differentiate this first-order condition to find that the solution, $\lambda^* = \Lambda(\mu)$, satisfies

$$\Lambda'(\mu) = -\frac{m_H}{m_H - \mu} < 0.$$

Thus, $\lambda_1^* \neq \lambda_\theta^*$ and $\lambda_2^* \neq \lambda_\theta^*$. Because the solutions are unique, then $U(\mu_1, \lambda_\theta^*) < U(\mu_1, \lambda_1^*)$ and $U(\mu_2, \lambda_\theta^*) < U(\mu_2, \lambda_2^*)$. The second inequality in (A.1) is strict, and so V is strictly convex. QED

Now we prove Lemma 3.3. Let $C''[m_L, m_H]$ contain the functions in $C'[m_L, m_H]$ which are convex (not necessarily strict). For any $J_0 \in C''[m_L, m_H]$, V is strictly convex. Because the maximum of two convex functions is convex, then $T(J_0)$ is convex. The contraction mapping argument shows that the fixed point of T is convex, i.e., J is convex. By the previous lemma, V is strictly convex. Because a convex function is almost everywhere twice differentiable with almost everywhere continuous first derivative (see Lemma 3.2 in Rader, 1973), then V and J have these properties.

For strict convexity of J, it suffices to show that T maps convex functions into strictly convex functions. Let $J \in C''[m_L, m_H]$. Using $L(\mu)$ defined in (2.4), we rewrite (2.6) as follows:

$$T(J)(\mu) = F(L(\mu))V(\mu) + \int_{L(\mu)}^{\bar{l}} l dF(l),$$

Let μ_1, μ_2 and μ_θ be described in the proofs of the above two lemmas. If $L(\mu) > \underline{l}$, then strict convexity of V and convexity of J imply:

$$\begin{array}{rl} & T(J)(\mu_\theta) \\ < & F(L(\mu_\theta))\left[\theta V_1 + (1-\theta)V_2\right] + \int_{L(\mu_\theta)}^{\bar{l}} \left[l + \frac{1}{1+r}\left(\theta J_1 + (1-\theta)J_2\right)\right] dF(l) \\ = & \theta T(J)(\mu_1) + (1-\theta)T(J)(\mu_2) + \theta\Delta_1 + (1-\theta)\Delta_2, \end{array}$$

where

$$\Delta_i = \left[F(L(\mu_\theta)) - F(L(\mu_i))\right] V_i + \int_{L(\mu_\theta)}^{L(\mu_i)} \left(l + \frac{J_i}{1+r}\right) dF(l), \quad i = 1, 2.$$

Using the definition of $L(\mu)$, we can rewrite Δ_i as follows:

$$\Delta_i = \int_{L(\mu_\theta)}^{L(\mu_i)} \left(l + \frac{J_i}{1+r} - V_i\right) dF(l) = \int_{L(\mu_\theta)}^{L(\mu_i)} \left[l - L(\mu_i)\right] dF(l).$$

Because $L(\mu)$ is increasing, $\Delta_i < 0$ for both $i = 1, 2$. Thus, $T(J)$ is strictly convex. QED

B. Proof of Lemma 3.4

To establish the differentiability of V, we apply the Benveniste-Scheinkman Theorem (see Stokey and Lucas, Theorem 4.10, p84). However, that theorem cannot be directly applied to (2.2) because $U(\mu, \lambda)$ is convex in μ for any given λ (see the proof of Lemma 3.3). To circumvent this problem, let us change the state variable in (2.2) from μ to x where x is defined as follows:

$$x = x(\mu) \equiv \frac{\mu - m_L}{m_H - \mu}.$$

This is well defined for all $m < m_H$. We write $\mu = \mu(x)$. Then, $\mu'(x) > 0$ and $\mu'' = -2\mu'/(1+x) < 0$. Moreover, the level of x in the next period can be written as:

$$H(\mu(x), \lambda) = x\frac{1 - \lambda m_H}{1 - \lambda m_L}.$$

Inverting this relationship to express λ as follows:

$$\lambda = g(H/x) \equiv \frac{1 - H/x}{m_H - m_L H/x}.$$

We change the choice variable from λ to H. The domain of H is $[\underline{H}(x), x]$, where $\underline{H}(x) = H(\mu(x), \bar{\lambda})$. Define

$$\widehat{U}(x, H) = U(\mu(x), g(H/x))$$
$$\widehat{V}(x) = max_H \widehat{U}(x, H).$$

Then, $V(\mu) = \widehat{V}(x(\mu))$. The function $V(\mu)$ is differentiable if and only if $\widehat{V}(x)$ is.

To show that $\widehat{V}$ is differentiable, we show that it is concave. In turn, we show that $\widehat{U}(x, H)$ is concave in x for any given H. We start with the following claim:

Claim 1. $[2\mu' g_x + \mu g_{xx}] < 0$, *and so* $[\mu(x)g(H/x)]$ *is concave in* x *for any given* H. *Also, under Assumption 1,* $W(g(H/x))$ *is concave in* x.

Proof. Compute:

$$[\mu(x)g(H/x)]_{xx} = \mu'' g + [2\mu' g_x + \mu g_{xx}],$$

where the subscript x indicates the derivative with respect to x. Because $\mu'' < 0$, then the condition $[2\mu' g_x + \mu g_{xx}] < 0$ is sufficient for $[\mu(x)g(H/x)]$ to be concave in x. Notice that $g' < 0$ and $g'' < 0$. Also,

$$g_x \equiv \tfrac{\partial g}{\partial x} = -\tfrac{H}{x^2} g' > 0,$$
$$g_{xx} \equiv \tfrac{\partial^2 g}{\partial x^2} = -\tfrac{2m_H}{xm_H - Hm_L} g_x < 0.$$

Then, the condition $[2\mu' g_x + \mu g_{xx}] < 0$ can be verified as follows:

$$2\mu' g_x + \mu g_{xx} = -2\mu' g_x \frac{(xm_H + m_L)^2 + (H+1)m_L(m_H - m_L)}{(m_H - m_L)(xm_H - Hm_L)} < 0.$$

To show that $W(g(H/x))$ is concave in x, compute:

$$[W(g(H/x))]_{xx} = W' g_x \left[\frac{g_{xx}}{g_x} + \frac{W''}{W'} g_x\right].$$

Because $W' < 0$ by Assumption 1, and $g_x > 0$, then $W(g(H/x))$ is concave in x if and only if $\frac{g_{xx}}{g_x} + \frac{W''}{W'} g_x \geq 0$. By part (iii) of Assumption 1, we have $W''/W' \geq 2/(1-g)$. Then, it can be verified that

$$\frac{g_{xx}}{g_x} + \frac{W''}{W'} g_x \geq \frac{g_{xx}}{g_x} + \frac{2}{1-g} g_x = \frac{2(1-m_H)}{(1-g)(xm_H - Hm_L)} > 0.$$

This completes the proof of the claim. QED

We continue the proof of Lemma 3.4. For any given H, the function $\widehat{U}(x, H)$ is twice differentiable in x. Computing the second-order derivative, we have:

$$\widehat{U}_{xx} = \left[\mu(x)g(\frac{H}{x})\right]_{xx} \left(\frac{W}{r} - \frac{J(H)}{1+r}\right) + \frac{W' g_x}{r}\left[2(\mu g)_x + \mu g \left(\frac{g_{xx}}{g_x} + \frac{W''}{W'} g_x\right)\right].$$

It is easy to verify that $(\mu g)_x > 0$. Under Assumption 2, $W > rJ(H)/(1+r)$. Then, the above claim implies $\widehat{U}_{xx} < 0$. That is, $\widehat{U}(x, H)$ is strictly concave in x.

We now show that $\widehat{V}(x)$ is strictly concave. Let x_1 and x_2 be two arbitrary values of x that are admissible, with $x_1 < x_2$. Let $\theta \in (0,1)$ be a number, and $x_\theta = \theta x_1 + (1-\theta)x_2$. Let H_i^* be the solution for H when $x = x_i$, where $i \in \{1, 2, \theta\}$. We have $\widehat{V}(x_\theta) = \widehat{U}(x_\theta, H_\theta^*)$, and so

$$\begin{aligned}\widehat{V}(x_\theta) \;\; &\geq \theta \widehat{U}(x_\theta, H_1^*) + (1-\theta)\widehat{U}(x_\theta, H_2^*) \\ &> \theta \left[\theta \widehat{U}(x_1, H_1^*) + (1-\theta)\widehat{U}(x_2, H_1^*)\right] \\ &\quad + (1-\theta)\left[\theta \widehat{U}(x_1, H_2^*) + (1-\theta)\widehat{U}(x_2, H_2^*)\right] \\ &= \theta \widehat{V}(x_1) + (1-\theta)\widehat{V}(x_2) - \theta(1-\theta)\Delta\end{aligned}$$

where

$$\Delta = \left[\widehat{U}(x_1, H_1^*) - \widehat{U}(x_1, H_2)\right] + \left[\widehat{U}(x_2, H_2) - \widehat{U}(x_2, H_1)\right].$$

The first inequality follows from the fact that H_θ^* is the optimal choice under x_θ, the second inequality from the fact that $\widehat{U}$ is strictly concave in x, and the equality from re-arranging the terms. It is easy to see that $\Delta \geq 0$. Thus, $\widehat{V}$ is strictly concave.

Under Assumption 3, the solution λ^* is interior, and so is the solution H^*. Because $\widehat{V}(x)$ and $\widehat{U}(x, H)$ are concave in x, the Benveniste-Scheinkman theorem applies, which shows that $\widehat{V}$ is differentiable. Moreover, $\widehat{V}'(x) = \widehat{U}_1(x, H^*)$.

As stated above, because $V(\mu) = \widehat{V}(x(\mu))$, then differentiability of $\widehat{V}$ implies that V is differentiable. By (3.1) and (3.2), $J(\mu)$ and $L(\mu)$ are also differentiable. These facts and Assumption 3 imply that the optimal choice λ^* is given by the first order condition, $U_2(\mu, \lambda^*) = 0$. Moreover, the envelope theorem implies $V'(\mu) = U_1(\mu, \lambda^*(\mu))$. QED

C. Proof of Proposition 4.2

As stated in the main text, $\Lambda_2(\mu, b) < 0$ iff $U_{23} < 0$. We prove $U_{23} < 0$ by induction. That is, supposing $U_{23} \leq 0$, we show $U_{23(-1)} < 0$, where the subscript (-1) indicates the previous period. Then, induction shows that $U_{23} < 0$ for all periods. Compute:

$$U_{23} = \frac{1}{1+r}\left[-\mu J_2(\mu, b) + (1-\mu\lambda)H_2 J_{12}(H, b)\right].$$

Suppose $U_{23} \leq 0$. Because $H_2 < 0$, then

$$J_{12}(H, b) \geq \frac{\mu}{(1-\mu\lambda)H_2} J_2(\mu, b). \tag{C.1}$$

To show $U_{23(-1)} < 0$, use (4.5) to obtain:

$$\frac{J_2(\mu, b)}{1+r} = \frac{F^*}{r+F^*} V_2(\mu, b) = \frac{F^*}{r+F^*}\left[1 + (1-\mu\lambda^*)\frac{J_2(H, b)}{1+r}\right]. \tag{C.2}$$

Differentiating the first equation with respect to μ and using (4.5) to substitute $L_2(\mu, b)$, we have:

$$\frac{J_{12}(\mu,b)}{1+r} = \frac{F^*}{r+F^*}V_{12}(\mu,b) + \frac{r^2 F^{*\prime}}{(r+F^*)^3}\left[V_2(\mu,b)\right]^2. \tag{C.3}$$

Differentiating (4.6) with respect to μ yields:

$$V_{12}(\mu,b) = -(\lambda^* + \mu\Lambda_1)\frac{J_2(H,b)}{1+r} + (1-\mu\lambda^*)\frac{J_{12}(H,b)}{1+r}\left(H_1 + H_2\Lambda_1\right).$$

Because $H_1 > 0$, $H_2 < 0$ and $\Lambda_1 < 0$, we can use (C.1) and the above equation to obtain:

$$V_{12}(\mu,b) \geq \left(\frac{\mu H_1}{H_2} - \lambda^*\right)\frac{J_2(H,b)}{1+r}.$$

Then, (C.3) implies:

$$\frac{J_{12}(\mu,b)}{1+r} \geq \frac{F^*}{r+F^*}\left(\frac{\mu H_1}{H_2} - \lambda^*\right)\frac{J_2(H,b)}{1+r} + \frac{r^2 F^{*\prime}}{(r+F^*)^3}\left[V_2(\mu,b)\right]^2.$$

Substituting this result and (C.2) into the formula of $U_{23(-1)}$ and noting $H_{2(-1)} < 0$, we get:

$$\begin{aligned} U_{23(-1)} \leq\ & \tfrac{F^*}{r+F^*}\left[(1-\mu\lambda^*)\left(\Delta - \mu_{-1}\right)\tfrac{J_2(H,b)}{1+r} - \mu_{-1}\right] \\ & +(1-\mu_{-1}\lambda^*_{-1})H_{2(-1)}\tfrac{r^2F^{*\prime}}{(r+F^*)^3}\left[V_2(\mu,b)\right]^2, \end{aligned} \tag{C.4}$$

where

$$\Delta = H_{2(-1)}\frac{1-\mu_{-1}\lambda^*_{-1}}{1-\mu\lambda^*}\left(\frac{\mu H_1}{H_2} - \lambda^*\right).$$

Recall that $\Lambda_1(\mu, b) < 0$. Because $H(\mu, \lambda^*) \leq \mu$, then $\lambda^* \geq \lambda^*_{-1}$. Computing H_1 and H_2, we get:

$$\frac{\mu H_1}{H_2} - \lambda = -\frac{(1-\mu\lambda^*)(\mu - \lambda^* m_H m_L)}{(m_H-\mu)(\mu-m_L)} \geq -\frac{(1-\mu\lambda^*)(\mu - \lambda^*_{-1} m_H m_L)}{(m_H-\mu)(\mu-m_L)}.$$

Because $H_{2(-1)} < 0$, then

$$\begin{aligned} \Delta\ & \leq -H_{2(-1)}\tfrac{(1-\mu_{-1}\lambda^*_{-1})(\mu-\lambda^*_{-1}m_Hm_L)}{(m_H-\mu)(\mu-m_L)} \\ & = \tfrac{(m_H-\mu_{-1})(\mu_{-1}-m_L)(\mu-\lambda_{-1}m_Hm_L)}{(1-\mu_{-1}\lambda_{-1})(m_H-\mu)(\mu-m_L)} = \mu_{-1}. \end{aligned}$$

The first equality follows from substituting $H_{2(-1)}$ and the second equality from substituting $\mu = H(\mu_{-1}, \lambda^*_{-1})$. Using the above result in (C.4) and noting $H_{2(-1)} < 0$, we have $U_{23(-1)} < 0$. QED

References

[1] Cosimano, 1993.

[2] Easley, D., 1989,

[3] Easley, D. and N.M. Kiefer, 1988, "Controlling a stochastic process with unknown parameters," Econometrica 56, 1045-1064.

[4] Rader, T., 1973, "Absolutely continuous constrained maximizers," Journal of Optimization Theory and Applications 12, 107-128.

[5] Stokey, N. and R.E. Lucas, Jr., with E.C. Prescott, 1989, Recursive Methods in Economic Dynamics, Harvard University Press, Cambridge, Massachusetts.