Optimal Unemployment Insurance, with Human Capital Depreciation, and Duration Dependence*

Nicola Pavoni†

September 29, 2008

Abstract

This paper introduces the possibility of a deterioration in job opportunities during unemployment into the standard optimal unemployment insurance (UI) design framework, and characterizes the efficient UI scheme. The optimal program may display two novel features, which cannot be present in stationary models. First, UI transfers are bounded below by a minimal *assistance* level that arises endogenously in the efficient contract. Second, the optimal scheme implies a *wage subsidy* for long-term unemployed workers. Numerical simulations based on the Spanish and US economies suggest that both assistance transfers and wage subsidies should be part of the UI scheme in these countries.

Keywords: Unemployment insurance, human capital depreciation, duration dependence, wage subsidies, moral hazard, recursive contracts, envelope theorems.

JEL Classification: C61, D63, D82, D83, J24, J31, J38, J64, J65.

SHORTENED TITLE: Optimal UI and Human Capital Depreciation

*First submission: July 2006. Main revision: November 2007.

†I thank Antonio Cabrales, Per Krusell, Matthias Messner, Kjetil Storesletten, Fabrizio Zilibotti, and two anonymous referees for helpful comments. The paper benefits from a profitable discussion with Philippe Aghion and Gianluca Violante, during the preliminary stages of the work. I wish to thank conference and seminar participants at the Duke University, the LSE, the Northwestern University, the University of Iowa, the IIES in Stockholm, and the CERGE-EI in Prague. This research was supported by Marie Curie Fellowship MCFI-2000-00689.

1 Introduction

Unemployment insurance (UI) programs are an important ingredient of social welfare policies in developed economies. Public expenditures devoted to UI and assistance programs in the OECD countries exceed 2% of GDP (Martin, 2002). These programs have been widely criticized because of the adverse effects they can have on workers' incentives to search for a new job. This criticism has stimulated extensive research into optimal insurance schemes that take these perverse effects into account.

A series of papers use the dynamic moral hazard model to analyze the trade-off between (unemployment) insurance and (search) incentives.[1] In these models, the key features of the environment are that the probability of finding a new job depends *only* on the (unobservable) search effort exerted by the agent and that the available gross wages distribution is constant throughout unemployment spells. It is, however, well documented that job opportunities deteriorate during unemployment.[2] In fact, many OECD countries propose and apply active labor market policies such as wage subsidies for long-term unemployed workers, mainly because of this adverse change in job opportunities.[3] Human capital depreciation and hazard rate duration dependence are hence important elements that need to be included in the study of an optimal UI designing problem.

In this paper, we extend the basic model of unemployment insurance with moral hazard to allow for both the gross wages and the probabilities of re-employment to depend on the length of the worker's unemployment spell. Although we do not analyze how the UI designing problem interacts with other welfare or labor market policies, we do allow the planner to impose history-contingent wage taxes or to

[1] We briefly review this literature below in this section.

[2] Many authors consistently find that displaced US workers face a large and persistent earnings loss upon re-employment in the order of 10-25% compared with continuously employed workers (Bartel and Borjas, 1981; Ruhm, 1987; Jacobson, LaLonde and Sullivan, 1993; for a survey, see Fallick, 1996). Keane and Wolpin (1997) estimate structurally an annual human capital depreciation rate for white US males during unemployment of between 9.6% and 36.5% (for blue and white collars respectively). Note that in this literature, unemployment duration is often found to have an independent negative impact on re-employment earnings, beyond job-specific and occupation-specific skill losses (e.g., Addison and Portugal, 1989; Neal, 1995). Second, a common feature of the data is the negative duration dependence in the unemployment hazard (e.g., Machin and Manning, 1999, for a survey). In particular, several studies (e.g., van den Berg and van Ours, 1994 and 1996; Bover, Arellano and Bentolila, 2002) continue to find a rapidly declining hazard even after explicitly controlling for unobserved heterogeneity.

[3] There are at least eight OECD countries that have actually introduced major welfare-to-work programs: the United States, Canada, the United Kingdom, Ireland, Denmark, France, the Netherlands and Sweden.

pay wage subsidies upon re-employment.

First of all, our study confirms one important result of most previous studies with stationary models: Benefits should decrease with unemployment duration. In fact, this behavior characterizes virtually all existing UI schemes in OECD countries. Perhaps more interestingly, we also observe that very simple schemes, defined by a low constant benefit payment b and a higher time-invariant wage w, can never be optimal. This result holds for any reasonable range of parameters (no restrictions on the worker's utility function, other than concavity and additive separability, are required) and regardless of the characteristics of the human capital depreciation process. We emphasize, instead, a simple necessary characteristic of any optimal unemployment insurance program which can be used as a 'back-of-the-envelope' test of optimality. For each unemployment duration t, by using today and next period benefit payment levels b and b' and next period net wage w', it is always possible to draw a triangle such as that in Figure 1. Given a reliable point estimate of the hazard rate π_t associated with unemployment spell of length t, one can easily check whether the two parts that form the segment connecting w' to b' satisfy the conditions $A = w' - b = (1 - \pi)(w' - b')$ and $B = b - b' = \pi(w' - b')$. When this is not the case and logarithmic utility is a good approximation for the agent's risk preferences, the trade-off between insurance and incentives is not exploited optimally, and there is room for a budget-saving reform.

The introduction of human capital depreciation and duration dependence also generates two main novel features in the optimal program, which are not present in stationary models. First, provided that human capital depreciates sufficiently rapidly during unemployment, the optimal path for unemployment benefit payments is initially decreasing and then becomes completely flat. The idea is simple: For low levels of human capital, the planner does not find it worthwhile to induce the worker to search intensively for a new job. Hence, UI benefits eventually stop decreasing and remain constant forever since the (long-term) unemployed worker is fully insured. The presence of a minimal, *assistance* level of UI benefits generates an endogenous lower bound on a worker's expected discounted utility. This feature of the optimal contract creates an important link between the characteristics of the efficient UI scheme and the speed of skills depreciation in the economy, which can also be used for positive analysis.

Second, recall that in our model the planner can impose history-contingent wage taxes upon re-employment. A key finding of the stationary model is that - due to consumption smoothing - the optimal wage tax should *increase* with the length of the previous unemployment spell (Hopenhayn and Nicolini, 1997a-b). In our nonstationary model, at least three new forces are at work that contrast with the tendency towards the wage tax increase. First, since the planner tends to insure agents against wage depreciation, a reduction in the gross wage received by the worker decreases the optimal wage tax. Second, in our model, low human capital levels imply a low effectiveness of the search activity, which increases the incentive costs. This latter effect tends to widen the difference between unemployment benefits and *net wage*, decreasing both the UI transfers and the wage tax upon re-employment. As long as the optimal scheme contemplates the emergence of an assistance state, these characteristics are reinforced by a third effect, which is strictly related to the exogenous minimum bound analysis of Pavoni (2007). The (now endogenous) presence of a lower bound on a worker's expected discounted utility shortens the effective time-horizon of the problem. This reduces the possibility of giving dynamic incentives and forces the planner to design a scheme biased toward the 'static' component of the incentives. During the unemployment period before the assistance state, the planner optimally increases the within-period gap between the unemployment insurance benefit and the net wage, further reducing the wage tax.

In the second part of the paper, we perform two quantitative exercises. We calibrate our model to the Spanish and US labor markets and compute the optimal UI schemes. Our numerical simulations suggest that in both these economies, the optimal UI scheme should contemplate an assistance state for workers with sufficiently long unemployment spells. Moreover, during the transition period towards assistance, the optimal level of wage taxes decreases with the length of the worker's previous unemployment spell and becomes a *wage subsidy* for long-term unemployed workers.

In both countries, the efficient provision of dynamic incentives implies optimal UI payments that are more generous than existing replacement ratios. This finding is consistent with most previous studies (for stationary models) with history-dependent taxes. We also find that the optimal US scheme involves a more substantial use of the wage subsidy instrument with respect to the optimal scheme for Spain.

Interestingly, both assistance programs and wage subsidies are common characteristics of most OECD unemployment insurance and welfare programs.[4] The possibility of an efficient wage subsidy to long-term unemployed workers is the second key normative prediction of our nonstationary search model with

[4]Examples of social assistance programs are the Temporary Assistance for Needy Families (TANF) and Food Stamps in the US, the Renta Minima de Insercion (or Renta Minima) in Spain, and the Revenu Mininum d'Insercion (RMI) in France. The Earned Income Tax Credit (EITC) is a major wage subsidy program in the US. This program has a structure similar to the Working Families' Tax Credit (WFTC) in the UK. Spain and France have much smaller wage subsidy programs, named Renta Activa de Insercion and Prime pour l'Employ respectively. Canada, Australia, and the Scandinavian countries have similar income assistance and wage subsidy programs.

moral hazard. Notice that this policy advice does not come purely from equity considerations,[5] nor is it induced by the presence of 'institutional' labor market imperfections such as the minimum wage. In our model, firms pay the workers according to their productivity, and the optimality of a wage subsidy crucially arises from efficiency (and incentive) considerations. When a previously unemployed worker finds a job, the planner saves the cost associated with the UI benefits. Whenever the wage subsidy costs are considerably lower than the costs of paying the UI transfers, the planner finds it profitable to induce the agent to search actively for a job, which will then be subsidized.

In their seminal work on UI, Shavell and Weiss (1979) establish that, because of moral hazard, benefits must decrease throughout the unemployment spell and *approach zero* in the limit. Hopenhayn and Nicolini (1997a) extend the analysis of Shavell and Weiss by introducing the possibility of contingent wage taxes after re-employment, and confirm the forever decreasing benefits result. The analysis of Hopenhayn and Nicolini also suggests that the tax on the wage the agent receives after re-employment should typically increase with the length of the previous unemployment spell. Our analysis shows that both these features may disappear once the stationarity assumption is relaxed.

As a by-product of our analysis, we develop a systematic approach suitable for studying the properties of the value function associated with a wide range of dynamic moral hazard problems, and other models with similar characteristics.

The literature on optimal UI is relatively new, yet quite extended. The summaries by Karni (1999) and Fredriksson and Holmlund (2003) report most of the relevant literature.[6] Virtually all these papers consider stationary models. Usami (1983) proposes a finite-horizon model with moral hazard, where the probability of re-employment conditional on search depends on the previous *employment* history. Usami confirms the aforementioned decreasing benefits result and finds that the worker compensation should be nondecreasing during the employment period. Although our model permits employment history dependence, most of our findings are induced by skill depreciation during *unemployment*. Moreover, Usami studies the problem choosing an 'inconvenient' state variable which prevents a complete analysis. Our recursive formulation generates a manageable value function, which allows us to characterize in

[5]Diamond (1980) and more recently Saez (2002) use static models of adverse selection with workers of different productivity and derive conditions for the optimality of income subsidies to low-skilled employed individuals. Differently from us, in this framework the *sign* of the tax (i.e., whether there is a positive wage tax or a subsidy) is solely determined by the Pareto weight attached to low-skilled workers. Incentive effects only determine the magnitude of the subsidy/tax.

[6]Most of the remaining papers address questions and/or use approaches that cannot be directly related to our own.

detail the optimal scheme, both qualitatively and quantitatively.

The paper is organized as follows. The next section presents the general formulation of the model. In Section 2.1, we then summarize the main characteristics of the optimal scheme under stationarity. In Section 2.2, we introduce human capital depreciation in a partially parameterized specification of the model, and provide sufficient conditions for the optimality of assistance programs. In Section 3, we briefly describe our approach to solving the general problem and provide the key qualitative characteristics of the optimal contract. In Section 4, we calibrate the model with both the US and Spanish economies, and simulate the efficient schemes for these cases. Section 5 concludes.

2 Model

Consider a risk-neutral planner who must design an optimal unemployment compensation scheme for a risk-averse worker. Both the planner and the worker discount the future at a rate $\beta^{-1} - 1$, with $\beta \in (0, 1)$. The worker has time-invariant preferences (over flows) of the following separable form:

$$u(c) - v(a)$$

where c is consumption and a is the search effort. We assume $u(\cdot)$ to be strictly increasing, strictly concave and continuously differentiable, with bounded inverse function u^{-1}.[7] In any period, the worker can be either employed (e) or unemployed (u). If the worker is employed, he produces the quantity $S(h)$ (his gross wage) which is assumed to be a continuous, increasing and bounded function of the worker's human capital endowment, h. Moreover, we assume that h follows the following stochastic law of motion:

$$h' = m_z(h), \; z = u, e; \quad \text{with } m_u(h) \leq h \leq m_e(h); \quad \text{and } m_z(\cdot) \text{ continuous}, \tag{1}$$

where h' is the next-period human capital level and z is the worker's employment state. The idea is that during unemployment ($z = u$) h depreciates, while during employment there can be human capital accumulation due, for example, to on-the-job training.

While unemployed, the worker can either search ($a = 1$) or not search ($a = 0$) for a new job, i.e., $a \in A = \{0, 1\}$. The search activity is costly: $v(1) = v > v(0) = 0$, and a affects the transition probability between employment states according to a hazard rate function $\pi(a, h)$, with $\pi(0, h) := \hat{\pi}(h) \geq 0$ and $1 > \pi(1, h) := \pi(h) > \hat{\pi}(h)$.

[7] This latter assumption is merely a technical one. It allows us to simplify considerably the proof of Proposition 7.

The situation where the planner asks the agent to stop searching actively for a job can be interpreted as that of 'social assistance' (or 'early retirement'). The analysis that follows will be considerably simplified by the assumption that the assistance state is irreversible. In order to have a simple tractable recursive formulation, we also introduce an aggregate variable denoted 'human capital', h, which affects both the worker's productivity and his re-employment probabilities. In a more general setup, h can be multidimensional, capturing - for example - different degrees of human capital specificity. In fact, the variable h should be considered more broadly than mere skill or ability. Since we do not study remedies for skill depreciation (such as training programs), our analysis will not be affected by the specific nature of h.

The timing of the model during unemployment is as follows. At the beginning of period t, the worker receives the unemployment benefit b_t and is required to supply the costly job-search effort a_t. If at the beginning of the following period the worker is employed, he receives a net wage $w_{t+1} = S(h_{t+1}) - \tau_{t+1}$, where τ_{t+1} represents the wage tax (or subsidy if negative); otherwise, the worker receives the UI benefit transfer b_{t+1}; and so on.

The crucial assumption of the model that we keep throughout the paper is that the planner cannot observe the worker's search effort, a. Thus, during unemployment there is a moral hazard problem. This means that unemployment benefits are not paid only as insurance, but must also play the role of giving incentives for search. We assume there are no informational problems related to h.[8]

Following the recursive contracts literature, we characterize the optimal scheme by using the following formulation. Let U and h be the discounted utility promised to the agent in period t and his/her human capital endowment respectively. Given (U, h), the planner's value function in the unemployment state V is defined as follows:

$$V(U,h) = \max_{a \in \{0,1\}} \left\{ V_a(U,h) \right\}. \tag{2}$$

The function V_0 describes the planner's value in the case where the agent is not required to search

[8]Since the laws m_z are known, and the realized employment states z are perfectly observable, this is equivalent to assuming that the planner knows the initial endowment, h_0, i.e., the pre-displacement wage.

$(a = 0)$. In this case, the planner's value is defined as

$$\begin{aligned} V_0(U,h) &= \sup_{b,U^u,U^e} -b + \beta\left[\hat{\pi}(h)\, W(U^e,h') + (1-\hat{\pi}(h))V_0(U^u,h')\right] && (3) \\ \text{s.t.} &: \\ U &= u(b) + \beta\left[\hat{\pi}(h)\, U^e + (1-\hat{\pi}(h))U^u\right] && (4) \\ h' &= m_u(h). \end{aligned}$$

Equation (4) is commonly called the 'promise-keeping' constraint and requires the contract to deliver the promised level of discounted utility to the worker. It also plays the role of law of motion for the state variable U. It is easy to see that - because of its absorbing nature - during assistance we have $U^u = U$, and the planner fully insures the worker by paying a constant transfer forever.

The planner's value associated with the case where the worker is required to actively search for a job is defined by V_1, and it solves

$$\begin{aligned} V_1(U,h) &= \sup_{b,U^u,U^e} -b + \beta\left[\pi(h)W(U^e,h') + (1-\pi(h))V(U^u,h')\right] && (5) \\ \text{s.t.} &: \\ U &= u(b) - v + \beta\left[(1-\pi(h))\, U^u + \pi(h)U^e\right], && (6) \\ U &\geq u(b) + \beta U^u, \text{ and} && (7) \\ h' &= m_u(h). \end{aligned}$$

In addition to the promise-keeping constraint (6), the problem defining V_1 also includes the incentive compatibility constraint (7) which ensures the agent is willing to deliver the amount of effort called for in the contract.

The function $W(U,h)$ denotes the planner's net return in the employment state, when the worker is entitled to receive a level U of expected discounted utility and is endowed with human capital stock h. During this state, the planner satisfies the promise-keeping restriction

$$U = u(w) - l + \beta U^e \tag{8}$$

by transferring a net wage w, after imposing the tax (or paying the subsidy) $\tau = S(h) - w$ on the gross wage $S(h)$. The parameter $l \geq 0$ denotes the effort cost of working. In the model, jobs are permanent, and we assume that there are no informational asymmetries during employment. The assumption that employment is an absorbing state is made, as in Hopenhayn and Nicolini (1997a-b), to focus the analysis of the optimal dynamic contract to the unemployment experience. This assumption has no bearing on the qualitative characterization of the optimal program during unemployment.

While employed, the worker is hence fully insured, and from (8) one can easily verify that the planner's value is

$$W(U,h) = \frac{S_e(h) - w(U)}{1-\beta} = \frac{S_e(h) - u^{-1}((1-\beta)U + l)}{1-\beta}, \tag{9}$$

where $S_e(h)$ is the average discounted gross wage and can be computed recursively as follows:

$$S_e(h) = (1-\beta)S(h) + \beta S_e\left(m_e(h)\right).$$

The properties of u and S guarantee that W is bounded, strictly decreasing, strictly concave and continuously differentiable in U.

From (3), it is easy to see that the function V_0 takes a separable form similar to that of W, and it satisfies all properties we just mentioned for W.

2.1 The Stationary Benchmark

In order to understand the role played by human capital depreciation and duration dependence in shaping the optimal unemployment compensation scheme, it is useful to first analyze the stationary version of the model. Let's assume that the re-employment probabilities are constant $[\pi\,(h)\,, \hat{\pi}\,(h)] := [\pi, \hat{\pi}]$ with $\pi > \hat{\pi} \geq 0$ and that the gross wage $S\,(h) := S > 0$ does not depend on the level of human capital h. Similarly to the general case, the value of unemployment V is defined as

$$V(U) = \max\ \{V_1(U), V_0\,(U)\}\,,$$

where V_1, V_0 are the stationary analogs to (3)-(4) and (5)-(7), and $W(U) = \frac{S\, -\, u^{-1}((1-\beta)U+l)}{1-\beta}$.

Proposition 1 ***(The Stationary Case)*** *Assume that at U_0 in period $t = 0$ the planner decides to implement $a_0^* = 1$. Then* (i) *the agent will never be required to stop searching actively for new jobs, i.e., $a_t^* = 1$ for all t;* (ii) *U_t, b_t and w_t are all strictly decreasing during unemployment; and* (iii) *for any level of utility $\underline{U} > \frac{u(0)}{1-\beta}$, there exists a finite unemployment spell duration T such that $U_T < \underline{U}$.*

All proofs are reported in the Appendix.

Since u is concave, when U_0 is very large even compensating the agent for the search effort cost becomes too expensive, and the planner implements $a_0^* = 0$. For more moderate levels of initial lifetime utility, the worker is required to search intensively for a new job ($a_0^* = 1$). In this case, because of dynamic incentive provision, both the worker's UI payments and lifetime utility decrease during unemployment. Since lower lifetime utilities imply lower incentive and effort compensation costs, in a nontrivial stationary problem there is no role for social assistance: The worker is always required to search for a job, and the UI transfers never stop decreasing. Result *(i)*, the monotonicity of search effort during unemployment, has never been emphasized in the literature before. Result *(ii)* is more standard and constitutes a key finding in the analysis of Hopenhayn and Nicolini (1997a): Because of consumption smoothing, the wage tax $\tau_t = S - w_t$ upon re-employment increases with unemployment duration. This way, worker consumption is uniformly reduced during unemployment, in both employment and unemployment states.

The fact that, in order to spread out incentive costs, the planner reduces the agent's expected discounted utility U_t through time is a property always true in stationary models, and sometimes has unpleasant consequences. Result *(iii)* shows that the optimal contract implied by the repeated moral hazard model creates a weaker form of the 'immiseration result': When u is unbounded below, efficiency requires that the worker's expected discounted utility falls, with positive probability, below any arbitrary negative level. The infinite punishments result is questionable in some circumstances. For example, it may be impossible for the planner to enforce, ex-post, such punitive plans because these would imply excessive social conflict costs. Similarly, excessive punishments may induce the worker to opt out of the insurance scheme (e.g., see Pavoni, 2007). Below, we show that when human capital depreciates rapidly enough, misery is avoided since the optimal scheme generates an endogenous lower bound on payments, and hence on lifetime utilities.

2.2 Introducing Human Capital Depreciation: Efficient Policy of Assistance

Human capital depreciation clearly has the effect of reducing the planner's returns from the search activity for any *given* level of worker's lifetime utility. On the other hand, the analysis of the stationary model shows that because of dynamic incentive provision, the worker's expected discounted utility is decreasing during unemployment, and low lifetime utilities imply lower incentive and effort compensation costs. Obviously, which one of these two forces - human capital depreciation or decrease in lifetime utility - dominates in a nonstationary model is in part a quantitative issue.

We now partially parameterize the model and show a sufficient condition on the speed of human capital depreciation that guarantees the presence of an endogenous lower bound on worker's lifetime utility due to the emergence of an assistance state in the optimal program.

Consider the following relationship between gross wage and human capital endowment: $S(h) = \omega h > 0$.[9] Clearly, the *stationary* case can now be replicated by assuming that h remains constant (i.e., $m_u(h) = h = m_e(h)$). In this case, when the utility of the agent takes the logarithmic form (i.e., $u(c) = \ln(c)$) and $\hat{\pi} = 0$,[10] it can be shown that the following solutions to the functional equations (2)-(9) describe the 'true' planner value functions:[11]

$$
\begin{aligned}
V_1(U) &= \frac{\omega h \Lambda}{1-\beta} - \frac{\Gamma \exp\{(1-\beta)U\}}{1-\beta} \quad \text{if } U \leq M, && (11)\\
V_0(U) &= -\frac{\exp\{(1-\beta)U\}}{1-\beta} \qquad\qquad \text{if } U \geq M; \text{ and} && (12)\\
W(U) &= \frac{\omega h}{1-\beta} - \frac{\exp\{l\}\exp\{(1-\beta)U\}}{1-\beta} \text{ for all } U. && (13)
\end{aligned}
$$

It can be verified directly by computing the optimality conditions that during unemployment (when $a = 1$) the worker's lifetime utility decreases according to[12]

$$U^u(U) = U - \frac{\ln \Gamma}{\beta}, \qquad \text{with} \qquad \Gamma > 1. \tag{14}$$

Notice that consistently with Proposition 1 *(iii)*, since $\ln(0) = -\infty$ lifetime utility goes to minus infinity as $t \to \infty$ in the stationary version of the model.

Human Capital Depreciation. Now assume that, during unemployment, h depreciates according to

$$h_{t+1} = m_u(h_t) = (1-\delta)\, h_t, \qquad \text{with} \qquad \delta \in (0,1]\,.$$

[9]In a world where the aggregate production technology uses skill units h, which are perfect substitutes, ω is the skill's marginal product.

[10]A similar closed form can also be obtained for the case where $\hat{\pi} > 0$.

[11]The values for the constants are as follows: $M = \frac{\ln\left(\frac{\Lambda \omega h}{\Gamma - 1}\right)}{1-\beta}$, $\Lambda = \frac{\beta\pi}{1-(1-\pi)\beta}$, and Γ solves

$$\ln\left[\Gamma^{\frac{1}{\beta}} - (1-\pi)\Gamma\right] = \ln \pi + \frac{1-\beta}{\beta}\frac{v}{\pi} + l. \tag{10}$$

The reader can verify directly that the proposed functions are a solution of the functional equations. The proof that they actually represent the 'true' values for the planner is available upon request.

[12]We also have $U^e(U) = U + \frac{\frac{v}{\pi} - \ln\Gamma}{\beta}$ and $u\left(b(U)\right) = (1-\beta)U + \ln\Gamma$. Moreover, the employment state policy consists of a constant utility flow of $u\left(w(U)\right) = (1-\beta)U + l$.

We retain the assumptions that, during employment, h_t remains constant (i.e., $m_e(h) = h$) and that $\pi > \hat{\pi} = 0$ invariant in h.

Let a_t^* and U_t^* be the optimal search effort and continuation utilities in period t. It is easy to see that if $a_t^* = 1$ for all t (as would be the case for the stationary model), the planner would otherwise insure the agent against gross wage depreciation and the ratio between $V_1(U_t^*, h_t)$ and $V_0(U_t^*)$ would satisfy

$$\frac{V_1(U_t^*, h_t)}{V_0(U_t^*)} = \frac{\omega h_0 \Lambda \left(\frac{1-\delta}{1-\gamma}\right)^t - \Gamma \exp\{(1-\beta)U_0\}}{-\exp\{(1-\beta)U_0\}} \tag{15}$$

where we used (14) and $1-\gamma = \Gamma^{-\frac{1-\beta}{\beta}}$. Whenever $(1-\gamma) > (1-\delta)$, for $t \to \infty$ the right-hand side of (15) tends to $-\Gamma < -1$. Hence, there must exist a $T < \infty$ such that $V_0(U_T^*) > V_1(U_T^*, h_T)$, and the conjecture that $a_t^* = 1$ for all t is false. We have just shown the following result.[13]

Proposition 2 ***(Assistance)*** *Consider the log-utility specification of the model, with $\pi(h) = \pi$, $S(h) = \omega h$, $m_e(h) = h$ and $m_u(h) = (1-\delta)h$. If $(1-\delta) < \Gamma^{-\frac{1-\beta}{\beta}}$ then the optimal scheme contemplates an assistance program.*

The discrete effort choice is important for the previous result. This assumption is coherent with a long tradition in labor economics and macroeconomics which stresses the importance of fixed costs and the extensive margin in participation decisions (e.g., Cogan, 1981; Eckstein and Wolpin, 1989).

It turns out that the value of Γ depends very little on π, while it is quite sensitive to the parameters describing the effort costs. Assuming $v = l$, i.e., that work and search effort costs are of comparable magnitude, one can numerically compute the rate of human capital depreciation δ required to satisfy the condition of Proposition 2 for the least favorable value of π, within an empirically relevant range.[14] Depending on the effort cost,[15] we find that depreciation rates between 3% and 5% suffice to generate the presence of an assistance state.

In Section 4, we will discuss more extensively the empirical evidence for the Spanish and US labor markets. For comparison, note that virtually all empirical findings for the US report an annual rate of

[13]It is not difficult to extend the proposition to the case where $\hat{\pi} > 0$. Moreover, a similar sufficient condition can be derived in the case of time-changing hazard $\{\pi(h_t)\}$. Details are available upon request.

[14]In our computations, we interpret one period as one month and set $\beta = .996$ accordingly. We then consider values for the monthly hazard rate, π, between .4 and .03, corresponding to average unemployment durations of between 2.5 and 33 months.

[15]We consider effort cost levels between .5 and 1, which are well within the range of empirical relevance, as we will discuss in the quantitative section.

wage depreciation above 10%. The evidence on wage depreciation for Spain is rare and sometimes shaky. The lowest estimate for this country is as low as 3%, implying a violation of the sufficient condition for some values of the parameters. We postpone to Section 4 a more accurate quantitative analysis of the Spanish case.

3 Characterization of the Optimal Contract

We now study our model in its general specification. It is well known that hidden-action moral hazard models do not typically describe concave problems (Grossman and Hart, 1983; Phelan and Townsend, 1991).[16] There are three main reasons why a nonconcave problem may prove to be problematic to solve, especially in a dynamic environment. First of all, nonconcavity might also cause issues of nondifferentiability. Second, even assuming differentiability, first-order conditions may no longer be sufficient for local or global maxima. Finally, and more importantly, the usual envelope theorems cannot be applied,[17] and this may reduce considerably the usefulness of our recursive formulation.

In this paper, we develop a systematic approach that allows for these complications. First, we somewhat confirm the above-mentioned difficulties, since we find that 'in most cases' the associated value function is neither concave nor differentiable. However, we derive a positive result as well: The optimal contract can still be characterized to a great extent by using the familiar first-order conditions.

The idea of our approach is as follows.[18] The complication involved by the recursive study of the dynamic moral hazard problem comes from the incentive constraint. This prevents a direct approach to the study of the concavity and differentiability of the value function V. We thus first reformulate the problem to make it suitable for such analysis. We define a collection of concave and continuously differentiable functions (the *conditional functions*), of which the value function V is the upper envelope. We then apply the extended envelope theorem of Daskin (1967) to this problem to show that V is almost everywhere differentiable.

Our successive step is to study the 'switching points,' that is, the utility levels at which the upper envelope function V switches between two different conditional functions of the above-mentioned class.

[16] Clearly, these complications arise both in the case where effort is discrete and when there is a continuum of effort levels (see also Pavoni, 2000).

[17] By envelope theorems we mean theorems that describe conditions under which the value function of a parameterized optimization problem is a differentiable function of the parameter.

[18] The formal derivation of the results is in the Appendix.

Those points are indeed the only problematic ones. Given the characteristics of our class of functions, however, each switching point possesses a very nice characteristic: Either the V function is in fact differentiable at this point or the point is never reached in equilibrium. The fact that the points of nondifferentiability cannot be reached in equilibrium allows us to disregard them while characterizing the optimal contract.

The use of the Daskin envelope result is not new to the economic literature.[19] The contribution here is to show that it can be applied to study the properties of the dynamic moral hazard problem, and how. In particular, we demonstrate that some of the key characteristics of the value function generate important new advantages in dynamic recursive models, where the value function enters in the objective of the problem. The most useful one is perhaps the possibility of using the standard first-order conditions to characterize the optimal contract.

Sequence of Efforts Formulation. Consider the space $\mathcal{A}$ of all the sequences of efforts $\mathbf{a} = \{a_n\}_{n=0}^{\infty}$, $a_n \in A$, implementable during unemployment. For any human capital endowment $h \in \mathcal{H}$, effort sequence $\mathbf{a} \in \mathcal{A}$ and utility level $U \in \mathcal{U} \subset \mathrm{I\!R}$, we can define

$$\begin{aligned} V(\mathbf{a}, U, h) &= \sup_{b, U^e, U^u} -b + \beta \left[\pi(a,h) W(U^e, h') + (1-\pi(a,h)) V({}_1\mathbf{a}, U^u, h')\right] \\ & \quad \text{s.t. (1), and} \\ \text{if } a &= 1, \text{ (6) and (7);} \quad \text{if } a = 0, \ U = u(b) + \beta U^u. \end{aligned} \tag{16}$$

The functional $V(\mathbf{a}, U, h)$ represents the planner's optimal payoffs *conditional on* a given sequence of efforts, when the worker is unemployed. The symbol ${}_1\mathbf{a} = \{a_n\}_{n=1}^{\infty}$ stands for the 'one-step-ahead' continuation of $\mathbf{a}$.

It is easy to show that the value function of the original (fully-sequential) problem satisfies (16).[20] In Proposition 7 and 8 in the Appendix, we show that the converse is also true, that the (true) conditional function V is continuous, and that for all $\mathbf{a}$ and h the functions $V(\mathbf{a}, \cdot, h)$ are concave and differentiable in U. The problem (16) of implementing optimally (minimizing costs) a *given* sequence of efforts $\mathbf{a}$ is indeed concave, with linear constraints. The associated value function $V(\mathbf{a}, \cdot, h)$ is thus concave. Given

[19] It was first discovered by Kim (1993), Sah and Zhao (1998) and Milgrom (1999); and it has recently been extended by Milgrom and Segal (2002). They all consider static problems, however.

[20] Notice that we do not have measurability problems because there are only finitely many possible outcomes.

concavity, differentiability can be shown by applying the Benveniste and Scheinkman (1979) Lemma in the standard way.

The Shape of the Value Function. In the Appendix, we show that the value function $V(U,h)$ can be seen as the upper envelope of the collection of the conditional functions $V(\mathbf{a},U,h)$. The approach we propose exploits this interpretation for $V(U,h)$. In order to eliminate the possibility of complicated behaviors at the infinite due to the nonstationarity of the problem, we first make the following assumption:

> **Assumption A1** For any endowment h, there is a time-horizon $T(h) < \infty$ such that $\forall\, t \geq T(h)$ both $S\left(m_u^t\left(h\right)\right) = \underline{S}$ and $\pi\left(m_u^t\left(h\right)\right) = \underline{\pi}$, where m_u^t is the t^{th} iteration of m_u.

Assumption **A1** can be seen as a restriction on the law m, or on the functions $S(\cdot)$ and $\pi(\cdot)$, or on both. It should be noted that the above conditions allow $T(h)$ to be arbitrarily large (provided that it remains finite).

Proposition 3 *Consider a pair (U,h), with U interior,*[21] *and assume that* ***A1*** *is satisfied. Then $V(U,h)$ possesses both right and left derivatives in the first argument, with $V_+(U,h) \geq V_-(U,h)$. Moreover, $V(\cdot,h)$ is almost everywhere differentiable for all h, and where it is differentiable we have*

$$V'(U,h) = V'(\mathbf{a},U,h) \text{ for all } \mathbf{a} \in \mathcal{A}^*(U,h), \tag{17}$$

where $\mathcal{A}^*(U,h)$ is the nonempty set of optimal effort sequences when the initial conditions are (U,h) (i.e., the set of plans solving the maximization (28) defined in Proposition 9).

Proposition 3 is based on a version of the Daskin's envelope theorem for the case where the number of conditional functions is finite. This version of the theorem requires $V'(\mathbf{a},U,h)$ to exist and to be continuous in U. This is shown in the appendix in Proposition 8. In the Appendix, we also show that under Assumption **A1** our model can be reduced to the case where the *relevant* set of effort sequences is finite.

Remark 4 *The crucial step in the proof of the last proposition uses the fact that the set of relevant action plans $\mathcal{A}$ is finite. It is easy to show - using exactly the same lines of proof - that the results of*

[21] See Proposition 8 for details.

Propositions 3-5 remain valid for any finite horizon version of the dynamic moral hazard model, as long as there are finitely many outcomes and the set of feasible actions A is finite.

Using the First-Order Conditions to Characterize the Optimal Contract. Clearly, the characteristic $V'_+(U,h) \geq V'_-(U,h)$ we derived above is not a property of concave functions. In fact, when the directional derivatives differ, $V(U,h)$ cannot be concave in any interval containing U. However, this same characteristic implies that the 'kink' has a particular nature: It is an 'inward' one. This implies that the optimal choice of the continuation utility U^u will never be at a switching point. The problem (5) is hence differentiable at all relevant points, and the usual first-order conditions are still a necessary characteristic of an efficient contract. This is our main result and is presented in the following proposition:

Proposition 5 *Assume **A1** and interiority. The optimal contract necessarily satisfies (with the asterisks denoting optimality)*

$$-V'(U,h) = \frac{1}{u'(b^*)} \tag{18}$$

$$-W'\left(U^{e*},h'\right) = \frac{1}{u'(b^*)} + \mu\frac{\pi(a^*,h)-\pi(\hat{a},h)}{\pi(a^*,h)} \qquad \mu \geq 0 \tag{19}$$

$$-V'\left(U^{u*},h'\right) = \frac{1}{u'(b^*)} - \mu\frac{\pi(a^*,h)-\pi(\hat{a},h)}{1-\pi(a^*,h)}, \tag{20}$$

with $\mu = 0$ if either $a^* = 0$ or (7) is satisfied with strict inequality. Moreover, (18) can possibly fail only in the first period. In addition, we have

$$V'(U,h) = \pi(a^*,h)W'\left(U^{e*},h'\right) + (1-\pi(a^*,h))V'\left(U^{u*},h'\right).$$

The implications for the optimal scheme are not yet transparent. By rearranging the first-order conditions and using the envelope theorem, we get the following:

Corollary 6 *Under the conditions of the previous proposition, we have the following:*

$$\frac{1}{u'(b_t^*)} = \pi(a_t^*,h_t)\frac{1}{u'(w_{t+1}^*)} + (1-\pi(a_t^*,h_t))\frac{1}{u'(b_{t+1}^*)}. \tag{21}$$

Moreover, (i) $w_{t+1}^* \geq b_t^* \geq b_{t+1}^*$; *and* (ii) *either* $w_{t+1}^* > b_t^* > b_{t+1}^*$ *or* $w_{t+1}^* = b_t^* = b_{t+1}^*$.

Equation (21) is the same as that in Rogerson (1985), who uses a variational approach. Notice importantly that the first-order conditions (18)-(20) provide a stronger characterization of the optimal contract than (21). Incidentally, this becomes particularly evident in dynamic moral hazard models with more than two outcomes. In this case, equation (21) alone does not allow much to be said about the monotonicity of the payments. In contrast, the obvious generalization of our first-order conditions together with the envelope theorem (which one can easily show to be both true under very general conditions) will, for example, allow us to link punishments and rewards to the likelihood ratios, formally establishing in a multiperiod environment one of the key properties of the optimal contract in the static moral hazard model.[22]

Result *(i)* in Corollary 6 confirms, in a general, possibly nonstationary framework, one key finding of Shavell and Weiss (1979) and Hopenhayn and Nicolini (1997a) for the stationary model: The optimal unemployment insurance scheme requires the UI benefits to decrease with the duration of unemployment. However, notice that the wage tax behavior remains indeterminate. In Section 4, we take advantage of the recursive formulation to perform computer simulations of the optimal contract for the US and Spanish economies. We anticipate that in our numerical exercises we will find that an optimal scheme typically generates a plan of wage taxes $\tau_t^* = S_t - w_t^*$ that decrease with the length of the previous unemployment spell. This of course contrasts the result of the stationary model of Section 2.1, where the optimal re-employment wage tax, τ_t^*, increases during unemployment.

The second part of Corollary 6 has another interesting implication for policy. In many studies, unemployment insurance programs are modelled in a very simple way. Only two parameters are used to define the scheme. There, it is assumed to be a time-invariant unemployment benefit payment b, which - usually because of job-search incentives - is strictly lower than a time-invariant wage payment w. We

[22] Consider, for example, an extended version of our model with a more general wage distribution. Assume there are N potential jobs above the reservation value. If we denote by $\pi_i(h_t)$ the probability of finding job $i = 0, 1, 2, ..., N$ (with $i = 0$ meaning no job is found) when $a = a_t^* = 1$ and by $\hat{\pi}_i(h_t)$ the same conditional probability for $a = 0$, we obtain

$$\frac{1}{u'\left(c_{t+1}^i\right)} = \frac{1}{u'(b_t)} + \mu \frac{\pi_i(h_t) - \hat{\pi}_i(h_t)}{\pi_i(h_t)},$$

where c_{t+1}^i represents the worker's period $t+1$ consumption, conditional on the i^{th} realization. In the case of the worker finding a job (i.e., $i > 0$), his consumption corresponds to the net wage w_{t+1}^i in job i, while $c_{t+1}^0 = b_{t+1}$ denotes the consumption in the case of no job being found. The analogy to the classical 'modified Borch condition' for optimal risk sharing in the (static) moral hazard model is transparent.

can then ask the following question. Could this simple scheme be optimal, for some combination of wage depreciation and duration dependence? Part *(ii)* gives a clear negative answer to this question. The intuition is simple. If we had $w_{t+1} > b_{t+1}$ but $b_{t+1} = b_t$, the planner could reduce w_{t+1} and b_{t+1} by (different) small amounts so as to keep the dispersion in utilities in the two states, $u(w_{t+1}) - u(b_{t+1})$, constant. This uniform reduction in next-period utilities would continue to satisfy the incentive compatibility constraint. In addition, the expected discounted utility of the agent could be kept constant by an increase in b_t of magnitude $\frac{1}{\beta}$ times the reduction in b_{t+1}. The planner would gain by this perturbation, however, because the concavity of u implies that the reduction in w_{t+1} would be larger than β times the increase in b_t.

Intuitively, the reduction of w_{t+1} (compensated by an increase in b_t) is more attractive to the planner the higher the re-employment probability $\pi(a_t^*, h_t)$ is. This simple argument can be used to make our results ready to use for policy purposes. If we rearrange condition (21) and choose logarithmic utility ($u = \ln$), we obtain the following relationship between the optimal payments:

$$b_t^* - b_{t+1}^* = \pi(a_t^*, h_t) \left[w_{t+1}^* - b_{t+1}^* \right]. \tag{22}$$

According to (22), for small $\pi(a_t^*, h_t)$, optimality suggests almost flat UI schemes with a relatively large difference between net wage and unemployment insurance benefits; and vice versa for high hazard rates. This gives to (22) a quite appealing economic interpretation. According to this condition, workers facing relatively low hazard rates should be motivated to search for new jobs mainly through rewards: If they find a new job, they should receive a high net wage w_{t+1}^*. In contrast, those workers facing high probabilities of re-employment are mainly given search incentives by the use of punishments: If they fail to find a job they should face considerable drop in the unemployment benefit payment, b_{t+1}^*.[23]

Because of its graphical representation, shown in Figure 1, we may name the above equation as the *triangle rule*. Figure 1 reports an example of a possible existing UI scheme, where the lowest-level dotted line represents UI benefit payments, b_t, while the solid line represents a possible path for the net wage w_t, as a function of unemployment duration t. The triangle rule can be used as a very simple, 'back-of-the-envelope' test of optimality for an unemployment insurance scheme as follows: For each length t of unemployment, using today and next period benefit payment levels $b = b_t$ and $b' = b_{t+1}$ and next period net wage $w' = w_{t+1}$, one can easily draw a triangle such as that in Figure 1, where

[23] Notice that this simple argument about *relative* differences in consumption dispersion does not require knowledge of the likelihood ratio, which is an object quite difficult to observe since it includes the 'off-the-equilibrium' value $\hat{\pi}(h_t)$.

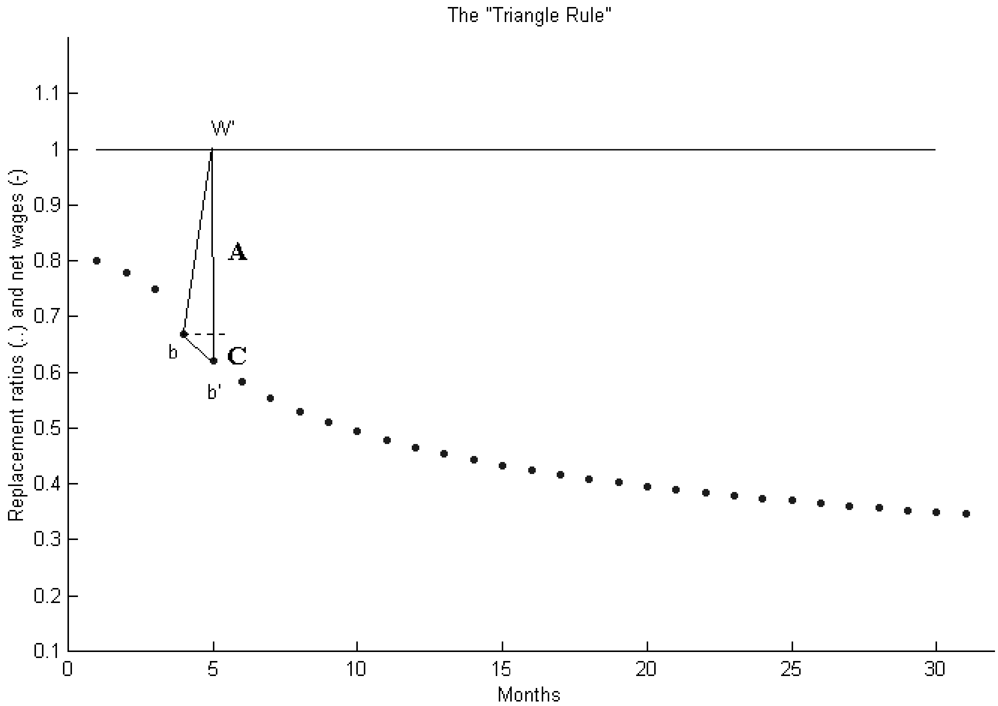

Figure 1: **An Example of the Triangle Rule.** In the figure, the lowest-level dotted line represents UI benefit payments, b_t, while the solid line represents the net wage, w_t, as a function of unemployment duration. Given a UI scheme, by using the level of UI benefit payments in two successive periods (say, t and $t+1$, denoted in the figure as $b = b_t$ and $b' = b_{t+1}$ respectively) and period $t+1$ net wage $w' = w_{t+1}$, it is always possible to draw a triangle such as that in the figure. Given an estimate of the hazard rate π_t associated to unemployment duration t, one can check whether the two parts that form the segment joining w_{t+1} to b_{t+1} have the proportions required by equation (22). Namely, if the UI scheme is efficiently designed the segments A joining b_t to w_{t+1} and C joining b_t to b_{t+1} should have lengths $(1-\pi_t)(w_{t+1} - b_{t+1})$ and $\pi_t(w_{t+1} - b_{t+1})$ respectively.

the segment jointing w_{t+1} to b_{t+1} has been divided in two parts by the perpendicular height (the short dotted horizontal line). The test also needs a reliable point estimate of the (equilibrium) hazard rate π_t associated with unemployment duration of length t. Then, by using π_t, one verifies whether the two parts that form the segment have the proportions required by equation (22). Namely, the segments A and C must have lengths $(1-\pi_t)(w_{t+1}-b_{t+1})$ and $\pi_t(w_{t+1}-b_{t+1})$ respectively. When this is not the case, the trade-off between insurance and incentives is not exploited optimally, and there is room for a budget-saving reform.

Alternatively, one could consider condition (22) as a valuable tool for deriving an optimal unemployment insurance scheme. In that case, however, the policymaker would need full knowledge of the hazard rate *functions*, $\pi(\cdot,\cdot)$. A further quantitative assessment concerns the degree of approximation - towards the fully optimal scheme - that use of the triangle rule involves for utility functions different from the logarithmic one. This analysis is left for future research.[24]

4 Quantitative Analysis

We now perform a couple of quantitative exercises. First, we calibrate our model with the Spanish economy and compute the optimal program. In order to have a more standard reference point for our quantitative results, we also perform a simpler calibration exercise based on the US economy.

We first of all aim to study the emergence of policies of social assistance and wage subsidies, for a coherently calibrated, empirically relevant set of parameters. Second, we attempt to characterize the quantitative aspects of such programs. In particular, we will measure the percentage of the wage subsidy and study how this number changes with the rate of human capital depreciation and the search effort cost.

Our numerical methodology is based on value function iteration. We approximate the value function with Chebyshev polynomials.[25] Value function iteration involves a maximization step. Although first-

[24]The log-utility case seems, however, to be a good approximation for an average level of relative risk aversion. For example, Attanasio and Weber (1993) use UK cohort data to estimate the intertemporal elasticity of substitution. Assuming CRRA preferences, their results imply a constant risk aversion parameter between 1.3 and 1.5 (where 1 corresponds to the log case). This is also consistent with several other previous studies.

[25]Chebyshev polynomials have several mathematical and practical advantages as discussed in Judd (1998). In order to capture the nonsmoothness of the value function, we considered a relatively high order of approximation (and several zeros). Given the simplicity of the model, the computational burden remained well within reasonable values (even for

order conditions were successfully used to characterize qualitatively the contract, given the nonconcavity of the problem, we use a global maximization algorithm in order to determine the optimal choices at each value function iteration.

4.1 The Spanish Case: Calibration

Preferences. We assume log-utility preferences over consumption (i.e., $u(c) = \ln c$) and we set the search cost equal to the cost of labor (i.e., $v = l$). Our calibration of the search cost is based on studies regarding the cost of participating to the labor market. Eckstein and Wolpin (1989) structurally estimate such a cost for women to be 62% in consumption equivalent terms, which in our formulation corresponds to $v = .97$. Keane and Wolpin (1997) estimate the same parameter for men to be 50% in consumption equivalent terms, which corresponds to $v = .69$.[26,27] Since these estimates are for the US economy, we implicitly assume stationarity of preferences across the two countries. We use a benchmark level of $v = l = .83$, which corresponds to the arithmetic average of the two values. The unit of time is set to one month. We hence pick a value for the discount factor of $\beta = .996$, in order to match an annual interest rate of 5%.

Wage and Wage Depreciation. Similarly to the analysis of Section 2.2, we consider a very simple linear relationship between human capital endowment and gross wage: $S(h_t) = \omega h_t$, and assume a geometric depreciation rule m_u for human capital: $h_{t+1} = (1-\delta)h_t$, where the pre-displacement value is normalized to $S(h_0) = 100$. Notice that in our specification, human capital and wage depreciation rates coincide; hence we can fix $\omega = 1$ without loss of generality. In order to calibrate the depreciation rate parameter δ, we follow Alba-Ramirez and Freeman (1990), one of the few studies that attempt such measurement for Spain. Their analysis is based on the ECVT labor force activity data-set (the Encuesta de Condiciones de Vida y Trabajo) for the period 1981-85. They find that a year of joblessness reduces

Matlab).

[26] The mapping between these estimates and our value for v is done according to the formula $\ln((1-x)c) = \ln c - v$.

[27] Eckstein and Wolpin (1989): see Table II for the sample means, Table IV for the estimated coefficients and equation (6) for the specification of the utility function. Keane and Wolpin (1997): see Tables 8 and 9 for estimated participation costs and Table 4 for sample means. Using a static model of labor supply, Cogan (1981) estimates a value for the participation cost as low as 41% in consumption equivalent terms. Static models, however, are likely to generate estimates for the participation cost that are biased downwards, since they ignore the cost from nonparticipation due to the forgone accumulation of experience.

workers' earnings by about 3%. It is very likely that 3% represents a lower bound on the true wage depreciation rate.[28] Given that our first target is to verify quantitatively the emergence of the assistance state and wage subsidies, we use such a precautionary rate for the benchmark case, and then consider higher rates in the sensitivity exercise. We furthermore assume that wage depreciation only lasts for three years: After 36 months, the gross wage remains constant. For simplicity, we assume that when the worker is employed, his human capital endowment remains constant (i.e., $m_e(h) := h$).

Hazard Rate. Our calibration of the hazard rate paths $\{\pi(a, h_t)\}_{t=0}^{T}$ for $a \in \{0, 1\}$ is based on Bover, Arellano and Bentolila (2002) (BAB). Since the search effort level a is not reported in the study, we performed the following simple 'identification' exercise. BAB use data from the Spanish Labor Force Survey for the period 1987-94 to estimate the Spanish hazard rates both for workers receiving unemployment insurance benefits and for those who do not receive any UI benefit transfer.[29] As expected, at all levels of unemployment duration, the former hazard rates (those relating to workers receiving UI benefits) are always lower than the latter ones. Moreover, the difference between the two functions decreases considerably with unemployment duration, and approaches zero after one year (see Figure 5 in BAB). Any reasonable structural model would typically imply that workers not receiving benefits supply a higher effort level than workers receiving benefits. We need to assume that this is the case at any level of unemployment duration. These considerations induce us to interpret, in our two-effort framework, the decrease in the difference between the hazard rates of the two groups of workers as a decrease in the effectiveness of the search activity. We hence calibrate the hazard rate paths $\{\pi(1, h_t)\}_{t=0}^{T}$ by using directly - month by month - their estimates, as reported by BAB in Table A6 and Figure 6, for workers not receiving benefits. We restrict attention to the group of unemployed aged 18-45, which is very homogeneous in the data. Since at the end of the period the hazard function has not yet levelled out, we linearly extrapolate two further periods, obtaining a final value of $\pi(1, h_T) = .03$ for $T = 17$ and over. Table 1 gives more detail.[30]

[28] In particular, the lack of data on pre-displacement wage prevents Alba-Ramirez and Freeman from estimating the initial wage loss. Rosolia and Saint-Paul (1998), for example, find that during unemployment, workers' wages might drop on average by up to 32%. Those authors, however, find that their results are not robust to different specifications of the statistical model.

[29] Notice that the pool of workers not receiving benefits includes those who received benefits in some previous period.

[30] Note that the hazard rate displays a nonmonotonic duration dependence. This phenomenon seems to be best explained by the so-called 'stock-flow approach' to search and matching (e.g., Coles and Smith, 1998; Gregg and Petrongolo, 2002).

Table 1: The Hazard Function (from BAB)

Duration (months)	$\pi(1, h_t)$	Duration (months)	$\pi(1, h_t)$
1	.1175	10	.1125
2	.16	11	.10
3	.225	12	.075
4	.20	13	.065
5	.15	14	.07
6	.14	15	.045
7	.144	16	.0375
8	.111	17 and over	.03
9	.108	-	-

For simplicity, we postulate a constant level for the 'passive' hazard rate for all unemployment durations, which is set to $\pi(0, h_t) := \hat{\pi} = .017$. This choice is consistent with both our interpretation of the data and the approximate stationarity of the estimated lower hazard rate function of the workers receiving benefits. The value has been chosen by looking at the hazard rates of the older generation (aged 45-60) since this is the group with the lowest search incentives (see BAB, Table 6 and Figure 6).

The Initial Level of Utility, U_0. The initial level of a worker's utility, U_0, is computed backwards in accordance with the existing scheme. The current insurance system can be represented by a contract that has no duration-dependent taxes or transfers when employed and pays a first benefit level b^1 of 70 for the first six months of unemployment; from the 7^{th} to the 24^{th} month, the benefit level b^2 is set equal to 60; and from the 25^{th} onwards, we assume the worker receives a subsistence level of benefits b^3 equal to 20.[31] The corresponding expected discounted utility value U_0 for an unemployed worker can

[31] In Spain, the replacement ratio is equal to 70% during the first six months of unemployment and 60% thereafter, subject to a floor of 75% of the minimum wage. Benefit duration is one-third of the last job's tenure, with a maximum of two years. The assistance system pays, for up to two years, 75% of the minimum wage to (unemployed) workers, with dependant, whose average family income is precisely below that amount. In 1998, the minimum wage was around 70,000 pesetas ($280) (Guia Laboral y de Asuntos Sociales, 1998). The amount of the noncontributory assistance level of transfers varies across different Autonomous Communities between 30,000 and 45,000 pesetas ($150-$180), is means-tested and has no fixed duration (see López, 1996). The Bulletin of Labor Statistics (1999) reports as 300,000 pesetas ($1,200) per month the 1998 average wage in nonagricultural activities. Following the common assumption that workers subject to

be calculated backwards as follows. According to the existing Spanish UI scheme, when a worker finds a job, his lifetime utility is $U_{work}(h) = \frac{\ln(S(h))-l}{1-\beta}$, which represents the utility of working forever and receiving the gross wage $S(h)$. Moreover, note that from period $T = 36$ onwards, the worker's problem is stationary. The unemployment benefits, the probability of finding a job and the gross wage are at their minimum levels. The value of his expected discounted utility U_T at period T can hence be computed as follows:

$$U_T = max\left\{\frac{\ln(b^3) - v + \beta\underline{\pi}U_{work}(\underline{h})}{1-\beta(1-\underline{\pi})}, \frac{\ln(b^3) + \beta\hat{\pi}U_{work}(\underline{h})}{1-\beta(1-\hat{\pi})}\right\}$$

where $b^3 = 20$ is the noncontributory assistance level of unemployment benefits and $\underline{\pi} = \pi(1, h_T) = .03$ is the bottom value for the high-effort hazard rate. For any $0 \leq t \leq T$, we can now define the value U_{T-t} recursively by

$$U_{T-t}(h_t) = \max_{a_t\in\{0,1\}} \ \ln(b_t) - v(a_t) + \beta\left[\pi(a_t,h_t)U_{work}(h_t) + (1-\pi(a_t,h_t))\,U_{T-(t-1)}(h_{t+1})\right],$$

where the period t benefit level $b_t \in \left\{b^1, b^2, b^3\right\}$ is computed according to the three-step scheme described above.

4.2 The Spanish Case: Results

The results of our benchmark simulations of the efficient scheme are reported in Figure 2. Three lines are displayed in this figure as a function of unemployment duration t: the UI benefit payments, b_t (represented by the dotted, lowest-level line); the gross wage, $S_t = h_t$ (represented by the homogeneously decreasing solid line); and the net wage, $w_t = S_t - \tau_t$ (the dash-dotted line).

The interested reader can first of all verify that the scheme obeys the triangle rule.[32] Second, observe that the optimal path for unemployment benefit payments is initially decreasing and then becomes completely flat. Note that the incentive constraint can be written as follows:

$$U^e \geq U^u + \frac{1}{\beta\left(\pi(h) - \hat{\pi}\right)}. \tag{23}$$

severe unemployment risk face a wage that is two-thirds of the average national wage, we consider the assistance level of benefits to be $\frac{40}{200} = .2$ (20%) of the pre-displacement wage, S_0.

[32]Using the payment levels of two consecutive periods of unemployment, one can draw a triangle such as that in Figure 1 and verify that the segments have the 'right' proportions. For example, consider the fourth and fifth periods of unemployment, which relate to $\pi(1, h_4) = .20$. Consistently with equation (22), the drop in UI payments $b_4 - b_5$ (roughly $.919 - .909 = .01$) is one-fifth of the gap between next-period payments in the two states: $w_5 - b_5$ (approximately $.962 - .909 = .053$).

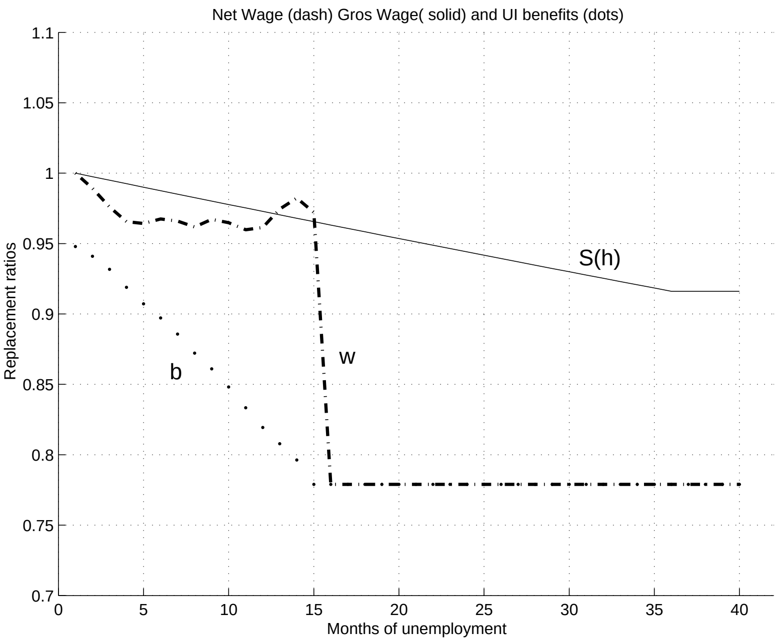

Figure 2: **The Spanish Benchmark Case.** The figure represents the simulation results for our benchmark calibration with $v = .83$ and an annual wage depreciation rate of 3%. The dotted, lowest-level line represents UI benefit payments, b_t, and the dash-dotted line represents net wage, $w_t = S(h_t) - \tau_t$, as a function of unemployment duration. In the 13^{th} 14^{th} and 15^{th} months, $w_t > S(h_t)$, i.e., the optimal program contemplates a wage subsidy upon re-employment.

When $\pi(h)$ decreases, incentive costs increase since, in order to widen the difference between U^e and U^u, the planner is forced to impose additional uncertainty on the worker's consumption. In addition, the expected returns from the search activity, $\pi(h)\frac{S(h)}{1-\beta}$, decrease as well. Hence, if human capital depreciates rapidly enough during unemployment, for long durations the planner does not find it profitable to induce the agent to search actively for a new job, and the UI payments eventually stop decreasing because the agent is fully insured. Third, notice that the net wage w_t is first decreasing, then roughly constant and then it actually increases with unemployment duration. After 14 months of unemployment, the worker enters the assistance program, and the net wage displays a remarkable drop since there is no more need for incentive provision.

The intuition for the behavior of the net wage before the assistance state is as follows. In a multiperiod setting, the optimal incentive scheme is shaped by the tension between within- and between-period consumption smoothing. The planner can improve within-period consumption insurance, i.e., reduce the difference between b_t and w_t, by relegating part of the punishment burden to the future. That is why the lifetime utility U_t and UI payments b_t decrease during unemployment. The emergence of a lower bound on a worker's lifetime utility shortens the effective time-horizon of the problem, forcing the planner to design a scheme biased toward the static component of the incentives. That is, the planner is forced to widen the difference between UI payments and net wage. Clearly, this effect tends to reduce the steepness of the net wage schedule w_t. However, it would also be present in a stationary model with an *exogenous* minimum bound on U. And, as is shown in Pavoni (2007), this force alone cannot induce the net wage schedule to increase with unemployment duration, i.e., it cannot induce a wage subsidy to long-term unemployed workers when S is constant.[33] Recall, however, that from (23) when $\pi(h)$ decreases the planner is forced to generate a larger wedge between U^e and U^u because of the incentive constraint. It is this additional effect on w_t, together with a moderate wage depreciation, that leads to the wage subsidy result in this case.

When unemployment duration is $12-15$ months, it is optimal to pay a wage subsidy to the worker upon re-employment. The planner finds it profitable to induce the agent to search actively for a subsidized job since the wage subsidy cost is considerably lower than the cost of paying the UI benefits. As described above, human capital depreciation and hazard rate duration dependence are key to generating

[33] The reason is consumption smoothing. Since U_t decreases with t and w_t is (weakly) increasing in U_t, the net wage cannot increase during the unemployment spell.

a policy paying a wage subsidy to long-term unemployed workers as part of an optimal UI program. In fact, we have already mentioned that this policy implication contrasts with that of stationary models à la Hopenhayn and Nicolini.[34] Although this qualitative phenomenon is of some importance *per se*, in the benchmark case the wage subsidy is relatively small, as it never goes beyond 2% of the gross wage.

In Figures 3 and 4, we report the results of our sensitivity analysis. Figure 3 is generated by simulating the optimal scheme for different rates of human capital depreciation. Notice that the consumption patterns are very similar, despite the relatively large variability in wage depreciation rates. There are indeed two contrasting effects. Human capital depreciation reduces planner returns at higher durations. At the same time, a high depreciation rate implies a lower initial utility U_0 for the worker under the existing scheme, and this reduces incentive and effort compensation costs at all durations. This latter 'wealth effect' is responsible for the emergence of the assistance state at long durations when the wage depreciation rate is actually high. When the depreciation rate is at 7%, for example (gross and net wages for this case are represented by the two solid lines), the assistance policy is postponed by one period with respect to the benchmark case of 3%. Moreover, since between durations of 14 and 15 months the hazard rate drops from .07 to .045 (see Table 1), the incentive constraint induces an important spike in the optimal net wage schedule for this case, implying a wage subsidy upon re-employment of slightly higher than 10%.

In Figure 4, we depart from our benchmark by considering two further levels of the search effort cost, v, which correspond to the estimated values for men and women we mentioned above. In this case, the wealth effect is negligible and the assistance state emerges at higher durations for lower search costs. The result is a nonmonotonic relationship between the level of the wage subsidy and v. For $v = .69$, for example, the wage subsidy reaches 5.1% at the 16 months duration.

Comparing the optimal program with the existing Spanish scheme we described in footnote 31, an important difference emerges: The optimal UI payments are much more generous than existing replacement ratios. This finding is consistent with most previous findings for stationary models and it is the consequence of the more efficient use of dynamic incentives. By partially back-loading punishments, the optimal program is able to reduce the within-period consumption dispersion imposed on the worker by the existing scheme.

We performed a number of further sensitivity exercises (which are available upon request). For

[34] For an application of the stationary version of our model to the Spanish economy, see Hopenhayn and Nicolini (1997b).

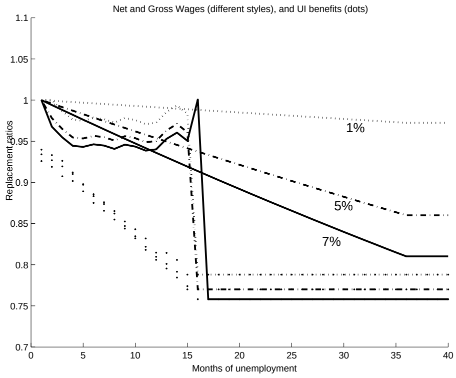

Figure 3: **Sensitivity Analysis I.** The optimal UI scheme for different human capital depreciation rates. All other exogenous parameters at their benchmark levels.

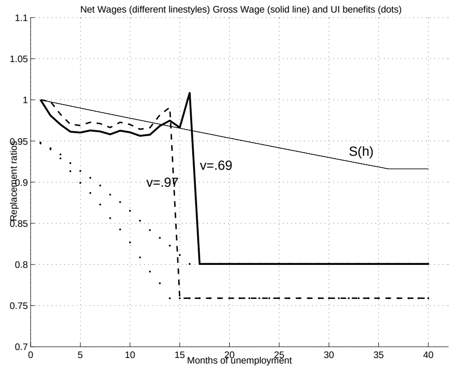

Figure 4: **Sensitivity Analysis II.** In this figure, we vary the search effort cost between .69 and .97. Wage depreciation is at the benchmark level of 3%.

example, we studied the effect of modifying the 'passive' hazard rate, $\hat{\pi}$. We find that increasing this parameter shortens the initial part of the program, i.e., the assistance state and wage subsidy emerge at lower unemployment durations. An increase in $\hat{\pi}$ also tends to increase the magnitude of the wage subsidy. This is very intuitive since from (23) a higher $\hat{\pi}$ increases incentive costs. For values of the parameter not too far from the benchmark case, however, the overall effect on the wage subsidy turns out to be relatively small.

4.3 The US Case

We conclude the quantitative section by performing a simple calibration exercise based on the US economy.

Calibration. We adopt the same preferences specification as that chosen for Spain: $u(c) - v(1) = \ln c - v$, with v ranging between .69 and .97. The unit of time is also set at one month, and $\beta = .996$.

Several estimates of the depreciation rate parameter δ are available for the US, with an important degree of dispersion. For example, Keane and Wolpin (1997) use NLSY data and (structurally) estimate an annual human capital (wage) depreciation rate for white US males of between 9.6% (for blue collars) and 36.5% (for white collars). Jacobson, LaLonde and Sullivan (1993) use plant closing data and find wage depreciation rates between 10% and 25%. We set our monthly benchmark level at $\delta = .0144$, which corresponds to an annual depreciation rate of 16%. We furthermore assume again a depreciation period of three years.

For simplicity, we assume that the hazard rate does not vary with unemployment duration, i.e., $\pi(1,h) = \pi$ and $\pi(0,h) = \hat{\pi}$ for all h. Using the results in Meyer (1990), we set the monthly US hazard rate corresponding to intensive search slightly higher than $\pi = .291$, so as to replicate a weekly value of 8.5%.[35] To identify the 'passive' hazard rate, we discount the average weekly rate by 7% as suggested by the estimates of Meyer reported in Table VI for the 45-54 age group, obtaining the value $\hat{\pi} = .039$. In summary, our benchmark choices for the US labor market are

$$\begin{bmatrix} \omega & \pi & \hat{\pi} & \delta \\ 1 & .291362 & .039 & .0144 \end{bmatrix}$$

Finally, we compute the initial level of a worker's utility, U_0, in a way similar to that of the previous example. According to Meyer (1990), the average level of UI benefits received in the sample is 66% of the average value of the pre-unemployment wage $S(h_0) = 100$, and lasted - again on average - 8.5 months (see Meyer, Table I). To have a finite value for U_0 with log utility, we assume that after 9 months of UI benefit payments, the worker continues to receive 10% of his gross wage. Such a figure is roughly consistent with the presence of basic assistance programs such as Food Stamps.[36]

Results. The results of our simulations are reported in Table 2, where we summarize our findings for three different levels of the search/work effort cost.

Qualitatively, the optimal path for unemployment benefit payments b_t presents the same characteristics as the Spanish case: It is initially decreasing and then it becomes completely flat. Notice that in this example both π and $\hat{\pi}$ are constant; hence, the spread in lifetime utilities required to implement the high effort level is constant. The planner's expected returns are, however, decreasing, since the gross wage, $S(h)$, decreases with unemployment duration. The resulting effect is similar to that of the previous case: Eventually, the planner 'releases' the agent from the search duty (i.e, the worker enters the assistance state). Also in this case, the optimal UI payments are more generous than the existing ones at all durations.

[35] Looking at Table II and Figure 3, the hazard rate seems pretty stable during the period 1978-83 considered by Meyer in his study.

[36] The maximum allotment of Food Stamps in 2005 was \$290 per month while, according to the US Bureau of Labor Statistics, the median male labor earnings in 2005 was $\$731 * 4 = \$2,924$.

Table 2: The US Case

Effort cost	$v = .69$	$v = .83$	$v = .97$
Months of UI before assistance	35	33	32
Initial replacement ratio (RR)	93.84%	93.78%	93.72%
RR after 1 year of unemployment	90.83%	89.91%	89.10%
RR after 2 years of unemployment	87.17%	85.73%	84.01%
RR at social assistance	79.22%	76.77%	74.52%
Wage subsidy after 1 year of unempl.	8.93%	8.33%	7.86%
Wage subsidy after 2 years of unempl.	24.93%	23.30%	21.75%

There are two main differences between the Spanish and US optimal schemes. First, in the latter the assistance state emerges at much higher durations than for Spain. Moreover, the optimal US scheme involves a considerable use of the wage subsidy instrument. The wage tax, $\tau_t = S(h_t) - w_t$, becomes negative (hence a subsidy) for unemployed durations as low as 6-8 months (depending on the effort cost). We also have large values for the optimal wage subsidy in this case. After approximately one year of unemployment, the worker should receive an average wage subsidy of 8.4%. Despite the sizeable depreciation rate for human capital, the assistance state only emerges for relatively high durations. This is so since π is quite high, both in absolute terms and relative to $\hat{\pi}$, implying high returns to search at relatively low incentive costs. At durations of around two years, this pattern of payments generates wage subsidy levels between 22% and 25% (depending on the effort cost).

Also in this case, we performed a number of sensitivity exercises, with parametric effects qualitatively similar to those for Spain. In the benchmark case (with $v = .83$), for depreciation rates around 8% and below, the optimal US program does not contemplate the emergence of an assistance state.

5 Conclusions

In the present paper, we extended previous studies on optimal unemployment insurance to incorporate the effects of human capital depreciation and duration dependence in the mechanism-design problem.

Our results partially confirm those obtained with stationary models, namely that benefits should (weakly) decrease with unemployment duration. The introduction of human capital depreciation and duration dependence also generates two key novel features of the optimal program. First, we derive analytical conditions on the speed of human capital depreciation during unemployment under which the planner eventually loses the incentives to induce the agent to supply high search effort. Consequently, unemployment benefit payments are initially decreasing and eventually become completely flat, since the long-term unemployed worker is fully insured by the planner. Second, we find that the increasing wage tax result of the stationary models à la Hopenhayn and Nicolini is not robust to this extension. Our simulations for the US and Spanish economies both show that although it is optimal to impose a wage tax after re-employment on short-term unemployed workers, the optimal level of wage tax should decrease with the length of the worker's previous unemployment spell, eventually becoming a wage subsidy for the long-term unemployed. The comparison with the standard (stationary) model emphasizes that the possibility of a transition from an insurance regime to an *assistance* regime, and the optimality of a *wage subsidy*, are the key policy implications of our nonstationary search model with moral hazard.

The results of our benchmark simulations for Spain - based on an annual wage depreciation rate of 3% - suggest that the optimal wage subsidy should never be above 2%. Wage subsidies above 10% can be generated as well, for human capital depreciation rates of around 7%. Our simulation results for the US case indicate a much more considerable use of the wage subsidy instrument for this country. This is for two main reasons: First, the US wage depreciation rate is higher than that of Spain. Moreover, since the US hazard rate is considerably higher than that of Spain, despite the higher depreciation the optimal scheme requires the worker to actively search for a job for much longer durations in the US. At durations of around two years, the wage subsidy can get as large as 25%.

Our analysis has an independent theoretical interest. We develop a new approach that allows us to study recursively the properties of the dynamic moral hazard model in a systematic way. We find that the associated value function is in general nonconcave and nondifferentiable. In spite of these nonsmoothness problems, we show that the optimal contract can be characterized by using the usual first-order conditions. The technique we developed in this paper uses Daskin's envelope theorem, and can be easily extended both to a more general class of moral hazard problems and to other problems with similar characteristics. Abreu, Pearce and Stacchetti (1990) and Spear and Srivastava (1987) discuss conditions under which the value function of the dynamic model with a continuum of outcome realizations is concave.[37] Our approach allows systematic study of the case with a finite number of

[37]The presence of a continuum of possible outcomes essentially convexifies the problem in a similar fashion to the use of lotteries (e.g., Phelan and Townsend, 1991).

output realizations, which is an inherent characteristic of the unemployment insurance designing problem studied in this paper.

The principal-agent framework permits detailed study of how dynamic incentive provision shapes the efficient UI scheme. The model has several weaknesses, however. For tractability, the framework presents a very stylized demand side of the labor market, and the analysis neglects general equilibrium effects. Although the qualitative characteristics of the pattern of payments should not be affected by these simplifications, the inclusion of search externalities in the analysis may influence the emergence of the social assistance program. *A priori*, it is, however, difficult to guess the direction of the general equilibrium effects. On the one hand, there are likely to be congestion effects in the aggregate, which might reduce the social value of workers' search effort. On the other hand, more intense job-search activity on the workers' side might enhance firms' incentives to post vacancies. In the presence of increasing returns, such a search externality gives additional social value to the search activity. These issue are all left for future research.

Finally note, that even assuming efficient demand (such as vacancy creation), our use of the worker gross compensation as a measure for the total net income surplus generated by the job is - in general - justified only when firms make zero profits (in expected discounted terms). If profits were positive, our analysis would tend to recommend too low search intensities.

Nicola Pavoni, University College London and Institute for Fiscal Studies. Department of Economics, UCL, Gower Street, London WC1E 6BT. E-mail: `n.pavoni@ucl.ac.uk`

6 Appendix

6.1 The Shape of the Value Function

We start by showing that the conditional value function $V(\cdot,\cdot,\cdot)$ is jointly continuous in its arguments.

Proposition 7 *The Bellman operator implied by (16) defines a contraction in the space of continuous and bounded functions with the sup norm. Thus V exists and is unique, and*

$$\|V\|_\infty = \sup_{y \in Y = \mathcal{A} \times \mathcal{U} \times \mathcal{H}} |V(y)| < \infty.$$

Proof. First of all, we should define a topology on $\mathcal{A}$. We do a bit more than that. We define the δ-metric on $\mathcal{A}$ as follows:

$$d_\delta(\mathbf{a}, \mathbf{a}') = \sum_{n=0}^{\infty} \delta^n |a_n - a_n'| \, , \ \delta \in (0,1) \, .$$

Second, let us simplify the notation by eliminating the h indexation. It will become clear below that the continuity of $m_u(\cdot)$, together with Lemma 9.5 of Stokey and Lucas with Prescott (1989) (SLP), allows us to make this simplification, at this stage.

Now consider a generic $s = (\mathbf{a}, U)$ with $\mathbf{a} = \{a_0, a_1, a_2, ...\} := \{a_{0,1}\mathbf{a}\} \in \mathcal{A}$. The distance between any two points is $\|s - s'\| = \|\mathbf{a} - \mathbf{a}'\|_\delta + |U - U'|$. Using the promise-keeping constraint, we rewrite the Bellman operator T as follows:

$$(TV)(s) = \sup_{U^u, U^e;\ s.t.\ (7)} -u^{-1}(U - v(a_0) - \beta[\pi(a_0)U^e + (1 - \pi(a_0))U^u]) + \beta[\pi(a_0)W(U^e) + (1 - \pi(a_0))V(_1\mathbf{a}, U^u)] \, .$$

Now we show that the operator T maps bounded and continuous functions into bounded and continuous functions. From the definition of continuity, we must verify that for each given point s and for each $\varepsilon > 0$, there exists a $\gamma > 0$ such that

$$\text{if } \ \|s - s'\| < \gamma \quad \text{then} \quad |(TV)(s) - (TV)(s')| < \varepsilon.$$

To this extent, we rewrite the previous condition using the definition of the Bellman operator:

$$\begin{aligned} &| \sup_{U^u, U^e;\ s.t.\ (7)} \{f(U, U^u, U^e, a_0) + \beta(1 - \pi(a_0))V(_1\mathbf{a}, U^u)\} \\ &- \sup_{U^u, U^e;\ s.t.\ (7)} \{f(U', U^u, U^e, a_0') + \beta(1 - \pi(a_0'))V(_1\mathbf{a}', U^u)\} \, | < \varepsilon \end{aligned} \tag{24}$$

where

$$f(U, U^u, U^e, a_0) = -u^{-1}(U - v(a_0) - \beta[\pi(a_0)U^e + (1 - \pi(a_0))U^u]) + \beta\pi(a_0)W(U^e)$$

and

$$W(U^e) = \frac{S - u^{-1}((1-\beta)U^e + l)}{1 - \beta}.$$

Now two easy steps. First of all, consider the following two cases.

Case 1: Suppose that $a_0 = a_0'$. In this case, we can assume a_0 as a parameter of the problem and apply the Maximum Theorem to the problem

$$\begin{aligned} F(U,_1\mathbf{a}) &= \sup_{U^u, U^e} f(U, U^u, U^e, a_0) + \beta(1 - \pi(a_0))V(_1\mathbf{a}, U^u) \\ s.t. &: (7),\ U^u, U^e \in \Gamma(U) \end{aligned} \tag{25}$$

to show continuity of F in $(U,_1\mathbf{a})$. The auxiliary constraint $\Gamma(U)$ is imposed in order to guarantee the constraint correspondence to be compact valued. A possibility is the following. The incentive compatibility constraint (7) can be expressed as $U^e \geq U^u + k(a_0)$, so we can always choose appropriately two constants $k_1, k_2 > 0$, and add to (7) the constraints $U^u \geq U - k_1$ and $U^e \leq U + k_2$. The continuity of F implies that we can always find a γ such that (24) is verified.

Case 2: Now suppose $a_0 \neq a_0'$. The idea here is that we do not check for continuity in this case, that is, we set γ such that, whenever $a_0 \neq a_0'$, then $\|s - s'\| > \gamma$. This can always be done since in this case $|a_0 - a_0'| = 1$. In summary, the choice of γ is done according to the continuity properties of F, with the restriction $\gamma \leq 1$.

Since u^{-1} is bounded, if we start from a bounded V, TV is bounded as well.[38]

Finally, one can check directly that the operator satisfies the Blackwell's sufficient conditions; thus T defines a contraction in the complete metric space of the bounded and continuous functions with the sup norm, in the 'reduced' space $S = \mathcal{A} \times \mathcal{U}$. The continuity of the law $m_u(\cdot)$ allows us to complete the proof by applying Lemma 9.5 and Theorem 9.6 of SLP, which guarantee that the contraction mapping result is still true in the original space $\mathcal{A} \times \mathcal{U} \times \mathcal{H}$, with h as exogenous state variable. **Q.E.D.**

We now show that the conditional value functions $V(\mathbf{a}, \cdot, h)$ are concave and differentiable with respect to the continuation utility U.

Proposition 8 *Consider a sequence of efforts $\mathbf{a} \in \mathcal{A}$ and an endowment level h, together with a law m.* (i) *The conditional function $V(\mathbf{a}, \cdot, h)$ is concave in U.* (ii) *Moreover, if we let*

$$V(\mathbf{a}, U_0, h) = -b_0 + \beta \left[\pi(a_0, h) W(U_0^e, h') + (1 - \pi(a_0, h)) V({}_1\mathbf{a}, U_0^u, h')\right]$$

with U_0 in the interior of the effective domain of $V(\mathbf{a}, \cdot, h)$, and with b_0 belonging to the interior of the domain of the agent's utility function u. Then $V(\mathbf{a}, \cdot, h)$ is continuously differentiable at any such U_0, and

$$V'(\mathbf{a}, U_0, h) := \frac{\partial V(\mathbf{a}, U_0, h)}{\partial U} = -\frac{1}{u'(b_0)} < 0. \tag{26}$$

Proof. *(i)* Again, the presence of h creates only notational complications, so we fix it and eliminate the h index in what follows. Following Grossman and Hart (1983) and changing the variable by defining $z := u(b)$, the problem becomes

$$\begin{aligned} V(\mathbf{a}, U) &= \sup_{z, U^u, U^e} -u^{-1}(z) + \beta \left[\pi(a) W(U^e) + (1 - \pi(a)) V({}_1\mathbf{a}, U^u)\right] \\ s.t. &: z - v(a) + \beta \left[(1 - \pi(a)) U^u + \pi(a) U^e\right] \geq z - v(\hat{a}) + \beta \left[(1 - \pi(\hat{a})) U^u + \pi(\hat{a}) U^e\right] \\ U &= z - v(a) + \beta \left[(1 - \pi(a)) U^u + \pi(a) U^e\right] \end{aligned} \tag{27}$$

where a is the first element in the sequence $\mathbf{a} = \{a_n\}$ Notice that the problem satisfies all the conditions required to apply Theorems 4.7 and 4.8 of SLP. To see why the problem is monotonic, use the promise-keeping constraint and notice that since u is increasing, the planner's objective function $-u^{-1}(z)$ - where $z = U + v(a) - \beta \left[(1 - \pi(a)) U^u + \pi(a) U^e\right]$ - is strictly decreasing in U. In particular, notice that interiority

[38] Note that the boundedness of u^{-1} is only used here and only in this proof.

is important here: It guarantees that z can indeed be modified to satisfy promise keeping without affecting incentive compatibility. Finally, note that if V is concave, the planner's objective function is concave (since u^{-1} is convex), and the constraints set is convex (linear); as a consequence, (27) is a concave problem. This proves concavity.

(ii) Differentiability can be shown as follows. Given that the value function is concave, we can use Lemma 2 in Benveniste and Scheinkman (1979). For a fixed level of promised utility U_0, we are looking for a differentiable and concave function $F(\mathbf{a},U)$ such that it is well defined in an interval I around U_0 and such that for any $U \in I$ we have $F(\mathbf{a},U) \leq V(\mathbf{a},U)$ and $F(\mathbf{a},U_0) = V(\mathbf{a},U_0)$. We claim that

$$F(\mathbf{a},U) = -u^{-1}\left(U + v(a) - \beta\left[(1-\pi(a))U_0^u + \pi(a)U_0^e\right]\right) + \beta\left[\pi\left(a\right)W\left(U_0^e\right) + (1-\pi\left(a\right))V\left({}_1\mathbf{a}, U_0^u\right)\right]$$

is the function we are looking for. Indeed, the optimal values U_0^e and U_0^u satisfy the incentive compatibility, and (by interiority) the promise-keeping constraint can always be satisfied by varying the benefit transfer b, so the function F is well defined and we have $F(\mathbf{a}, U) \leq V(\mathbf{a},U)\ \forall\ U \in I$, as required. The properties of u imply the concavity and differentiability of $-u^{-1}$. So F is concave and differentiable, and this implies that V is differentiable at U_0 and $V'(\mathbf{a},U_0) = -\frac{1}{u'(b_0)}$. Since V is concave, it is continuously differentiable. **Q.E.D.**

Finally, we show that the maximization with respect to a is always well defined. Hence, the value function $V(U,h)$ defined in (2) can be written as the upper envelope of the collections of conditional functions $V(a,U,h)$.

Proposition 9 *The set $\mathcal{A}$ of sequences of efforts is compact and $V(\cdot,U,h)$ is continuous in $\mathcal{A}$ for all (U,h). Thus, a maximum exists for any (U,h), and we can define*

$$V(U,h) = \max_{\mathbf{a}\in\mathcal{A}} V(\mathbf{a},U,h). \tag{28}$$

Proof. Given the continuity result we obtained in Proposition 7, to show the existence result it suffices to show the compactness of the set $\mathcal{A}$ in the topology we adopted above to show continuity.

Lemma 10 *$\mathcal{A}$ is compact in the topology induced by the metric d_δ for any $\delta \in (0,1)$.*

Proof of the Lemma. The set of all infinite sequences of zeros and ones corresponds to the Cantor set $\Delta := \{0,1\}^{\mathbb{N}}$, which is known to be compact in the topology induced by the metric d_δ for $\delta = \frac{1}{3}$.[39] From this, we can easily show that the Cantor set is topologically equivalent to the same set endowed with the topology induced by a d_δ with $\delta \in (0,1)$. Finally notice that the set $\mathcal{A}$ is a closed subset of such set of sequences; hence it is compact. **Q.E.D.**

We now combine the previous results, especially those of Proposition 7, to show the equivalence between the sequential and the recursive choice of efforts. We can only present a sketch of the proof since we never introduced the notation for the fully sequential specification of the problem. Further details are available upon request.

[39] See, for example, Aliprantis and Border (1994), page 93.

First of all, note that for each fixed $\mathbf{a} \in \mathcal{A}$ we have $V(U,h) \geq V(\mathbf{a},U,h)$. This inequality is easy to see since for each given sequence of effort levels, $V(\mathbf{a},U,h)$ solves a constrained version of the Bellman equation defining the unconstrained function $V(U,h)$ where at each node the choice of effort is restricted to be a given value, according to $\mathbf{a}$. This implies that $V(U,h) \geq \sup_{\mathbf{a}\in\mathcal{A}} V(\mathbf{a},U,h)$.

Now, from the previous Propositions we know that there is a contract starting from (U,h) that attains the value $V(U,h)$. Clearly, the sequence of actions implemented by this contract - call it $\mathbf{a}^*$ - is a feasible one.

We now show that $V(\mathbf{a}^*,U,h) \geq V(U,h)$. Note that, given $\mathbf{a}^*$, the payments associated to the optimal contract must be feasible according to problem (16). This is so since the feasibility correspondence is exactly the same in the two cases. If we let $(U_0,h_0) = (U,h)$, by iterating on the Bellman equation (16) conditional on $\mathbf{a}^*$, and using the fact that $\beta^T V\left({}_T\mathbf{a}^*, U_T^*, h_T\right)$ goes to zero as $T \to \infty$[40] we obtain the desired inequality, namely $V(\mathbf{a}^*,U_0,h_0) \geq V(U_0,h_0)$ because $V(\mathbf{a}^*,U_0,h_0)$ dominates any other feasible contract and $V(U_0,h_0)$ is just the sup value among all feasible contracts. **Q.E.D.**

6.2 Proofs of the Propositions in the Main Text

Proof of Proposition 1 *(i)* If at $t=0$ $a_0=1$ then we have that $V_1(U_0) \geq V_0(U_0)$. For any $n \geq 0$ let us now define the set of conditional functions as follows:

$$\begin{aligned}
V(U,n+1) &= \max_{b,U^e,U^u} -b+\beta\left[\pi W(U^e)+(1-\pi)V(U^u,n)\right] \\
U &= u(b)-v+\beta\left[\pi U^e+(1-\pi)U^u\right] \\
U &\geq u(b)+\beta U^u,
\end{aligned}$$

where the second index in the functions $V(\cdot,n)$ indicates the number of remaining periods of intensive job search, and $V(U,0) := V_0(U)$. Notice that all functions V are strictly concave and continuously differentiable in the first argument (for $U > \frac{u(0)}{1-\beta}$). Using the properties of a contraction, we obtain that also $V(U,\infty)$ is concave and continuously differentiable.

We now show the following lemma, which ranks the slope of the functions $V(\cdot,n)$.

Lemma 11 *$V'(U,n+1) \leq V'(U,n)$ for all U and n, with strict inequality at least for some n.*

[40] In the previous expression, ${}_T\mathbf{a}^*$ indicates the continuation of the sequence $\mathbf{a}^*$ after date T, and the limit result is true since in Proposition 7 we have shown that the conditional functions are bounded.

Proof of the Lemma. We will follow an inductive argument. Consider first the problem for $n = 0$. If we set $z := u(b)$, and denote by $g = u^{-1}$ the inverse function of u, we have

$$\begin{aligned} V(U,1) &= \max_{z,U^e,U^u} -g(z) + \beta\left[\pi W(U^e) + (1-\pi)V_0(U^u)\right] \\ U &= z - v + \beta\left[\pi U^e + (1-\pi)U^u\right] \\ U &\geq z + \beta U^u. \end{aligned}$$

It is easy to see that incentive compatibility is binding in this case (since $l \geq 0 = v(0)$ implies that $W'(U) \leq V_0'(U)$ for all U, and the functions are concave, if the incentive compatibility is slack, we have $U^u \geq U^e$). Hence from the first-order conditions we have

$$-g'\left(U - \beta U^u\right) = V'(U,1) > V_0'(U^u) = -g'\left((1-\beta)U^u\right),$$

which implies $U > U^u$. But then $U - \beta U^u > (1-\beta)U$. From the convexity of g, we have that $V_0'(U) = -g'\left((1-\beta)U\right) > V'(U,1)$.

Now, assume that $V'(U,n) \leq V'(U,n-1)$ for all U and consider the first-order conditions of the problems defining $V(U,n)$ and $V(U,n+1)$. We have

$$\begin{aligned} V'(U_n^u, n-1) &= -g'\left(U - \beta U_n^u\right) + \mu_n \frac{\pi}{1-\pi} \\ W'(U_n^e) &= -g'\left(U - \beta U_n^u\right) - \mu_n \\ V'(U_{n+1}^u, n) &= -g'\left(U - \beta U_{n+1}^u\right) + \mu_{n+1} \frac{\pi}{1-\pi} \\ W'(U_{n+1}^e) &= -g'\left(U - \beta U_{n+1}^u\right) - \mu_{n+1}, \end{aligned}$$

where μ_n and μ_{n+1} represent the multipliers associated with the problems defining $V(U,n)$ and $V(U,n+1)$ respectively. Similarly, U_n^u, U_n^e and U_{n+1}^u, U_{n+1}^e represent the optimal continuation utilities in the two cases. Now assume $\mu_{n+1} > \mu_n \geq 0$. Then by the induction argument we must have $U_n^u \geq U_{n+1}^u$ (just assume that $U_{n+1}^u > U_n^u$ and from the first-order conditions and the concavity of V one gets a contradiction), and by the incentive compatibility constraint we have $U_n^e \geq U_{n+1}^e$. But then $U - \beta U_n^u \leq U - \beta U_{n+1}^u$and $\mu_{n+1} > \mu_n$ implies that $-g'\left(U - \beta U_{n+1}^u\right) - \mu_{n+1} < -g'\left(U - \beta U_n^u\right) - \mu_n$, which is in contradiction to the fact that $W'(U_{n+1}^e) \geq W'(U_n^e)$ since W is a concave function. It must hence be that $\mu_n \geq \mu_{n+1} \geq 0$. But then by the induction argument and incentive compatibility we get $U_n^u \leq U_{n+1}^u$, and by envelope $V'(U,n+1) \leq V'(U,n)$. **Q.E.D.**

By the above inequalities and the fact that V is differentiable, the limit function $V(U,\infty)$ is such that $V'(U,\infty) \leq V'(U,n)$. Now notice that we have also shown that if $a_0 = 1$ then $U^u \leq U$. Lemma 11 hence implies that if $a_0 = 1$ then it is never optimal to choose $a_t = 0$ in the future, and $V_1(U) = V(U,\infty)$, as claimed.

(ii) The first-order conditions at all t such that $U_t^u > \frac{u(0)}{1-\beta}$ are

$$V_1'(U_t^u) = V_1'(U_t) + \mu \frac{\pi}{1-\pi}.$$

Since V_1 is concave and, as we argued above, the incentive constraint is binding (see also Pavoni, 2007, Lemma 1), we obtain that $U_t > U_t^u = U_{t+1}$, i.e., U_t is decreasing during unemployment. From the envelope condition $V_1'(U_t) = -\frac{1}{u'(b_t)}$ the UI transfer b_t decreases as well. From the incentive compatibility constraint we have that U_t^e decreases, and from the definition of W $u(w_{t+1}) = (1-\beta)U_t^e + l$, so w_t decreases as well.

(iii) If we let $\underline{U} > \frac{u(0)}{1-\beta}$ and $\delta := \min_{\underline{U} \leq U \leq U_0} \{U - U^u(U)\}$ then the minimization is well defined (by the Maximum Theorem), and $\delta > 0$. Hence the duration of unemployment T is any natural number satisfying $\infty > T \geq \frac{U_0 - \underline{U}}{\delta}$. **Q.E.D.**

Proof of Proposition 3 Our line of proof is based on a version of Daskin's envelope theorem. We first need a couple of definitions.

Definition 12 *For each U and h, define the nonempty set $\mathcal{A}^*(U,h) = \arg\max_{\mathbf{a}\in\mathcal{A}} V(\mathbf{a},U,h)$; moreover, we call $\mathcal{A}^*(h) = \bigcup_U \mathcal{A}^*(U,h)$ the set of all possible maximizers.*

Now note that, from Propositions 7 and 9, the conditional functions $V'(\mathbf{a},\cdot,h)$ are continuous in U, and the set $\mathcal{A}^*(h)$ is nonempty. Moreover, recall that from Proposition 8, $V'(\mathbf{a},U,h) := \frac{\partial V(\mathbf{a},U,h)}{\partial U}$ exists and is continuous in U.

We now show that $\mathcal{A}^*(h)$ is a finite set for each h. Assumption **A1** implies that after period $T(h)$ the problem becomes stationary. Hence from Proposition 1 and the absorbing nature of the assistance state we know that the optimal path of actions $\mathbf{a}^*$ can only take two forms: For all $t \geq T(h)$, we have either $a_t = 0$, or $a_t = 1$. In both cases the set of optimal efforts $\mathcal{A}^*(h)$ for the period-zero problem is a subset of the *finite* set of all the sequences of efforts which end either with the sequence $\mathbf{0} = \{0,0,0,0,...\}$ or with the sequence $\{1,1,1,1,...\}$ after $T(h) < \infty$. More precisely, the number of sequences of actions in $\mathcal{A}^*(h)$ is bounded by $2^{T(h)+1}$.

We are now ready for the crucial part of the proof.

Lemma 13 *Assume that $\mathcal{A}^*(h)$ is nonempty and finite, and that $V(\mathbf{a},\cdot,h)$ is continuous. Moreover, assume that $V'(\mathbf{a},U,h) := \frac{\partial V(\mathbf{a},U,h)}{\partial U}$ exists and is continuous in U. Then, for each given h, the value function $V(\cdot,h)$ has always both right and left derivatives, and these are given by the formulas*

$$\begin{aligned} V'_+(U,h) &= \max_{\mathbf{a}\in\mathcal{A}^*(U,h)} V'(\mathbf{a},U,h) \\ V'_-(U,h) &= \min_{\mathbf{a}\in\mathcal{A}^*(U,h)} V'(\mathbf{a},U,h); \end{aligned}$$

moreover, $V(\cdot,h)$ is almost everywhere differentiable in U, and whenever the derivative exists we have

$$V'(U,h) = V'(\mathbf{a}^*,U,h) \quad \textit{for all} \quad \mathbf{a}^* \in \mathcal{A}^*(U,h).$$

Proof of the lemma. In order to simplify the notation we again neglect the dependence to h. Hence $\mathcal{A}^*(U,h)$ becomes $\mathcal{A}^*(U)$ and $\mathcal{A}^*(h)$ becomes $\mathcal{A}^*$.

Let us show first the formula for the right-hand derivative. From our assumptions $\mathcal{A}^*(U)$ is nonempty and finite, so for each U we can take $\mathbf{a}(U) \in \arg\max_{\mathbf{a}\in\mathcal{A}^*(U)} V'(\mathbf{a},U)$. Now consider $U' > U$ and write the incremental ratio

$$\frac{V(U')-V(U)}{U'-U} = \frac{V(\mathbf{a}(U'),U')-V(\mathbf{a}(U),U)}{U'-U} \geq \frac{V(\mathbf{a}(U),U')-V(\mathbf{a}(U),U)}{U'-U}; \tag{29}$$

the last inequality comes from the fact that $\mathbf{a}(U') \in \mathcal{A}^*(U')$ so any other choice will reduce the value of $V(\mathbf{a}(U'),U') = V(U')$. Now, since the *conditional* functions are differentiable, as $U' \to U$ with $U' > U$ the far right-hand side of (29) converges to the derivative of the conditional function. In other terms, we have that

$$\liminf_{\substack{U'\to U\\ U'>U}} \frac{V(U')-V(U)}{U'-U} \geq V'(\mathbf{a}(U),U).$$

We now want to show that

$$\limsup_{\substack{U'\to U\\ U'>U}} \frac{V(U')-V(U)}{U'-U} \leq V'(\mathbf{a}(U),U)$$

where recall that $\mathbf{a}(U)$ is the maximizer of the partial derivative $V'(\mathbf{a},U)$ over the nonempty and finite set $\mathcal{A}^*(U)$. So suppose instead that

$$\limsup_{\substack{U'\to U\\ U'>U}} \frac{V(U')-V(U)}{U'-U} > V'(\mathbf{a}(U),U)$$

Then - by the definition of $\limsup$ - there is a decreasing sequence $\{U_n\} \to U$, $U_n > U$, and a real number ε such that

$$\frac{V(U_n)-V(U)}{U_n-U} \geq V'(\mathbf{a}(U),U)+\varepsilon \text{ for all } n \in \mathbb{N}.$$

Recall that definition of the unconditional function V and that since each $V(\mathbf{a},.)$ is continuous and the set $\mathcal{A}^*$ is finite and non empty, for each n there is at least one $\mathbf{a}_n$ such that $V(U_n) = V(\mathbf{a}_n,U_n)$. The sequence of $\mathbf{a}_n's$ has a cluster point, call it $\bar{\mathbf{a}}$. Hence, there must exist a converging subsequence $\{U_{n_k},\mathbf{a}_{n_k}\} \longrightarrow (U,\bar{\mathbf{a}})$. Since $\mathcal{A}^*$ is a finite set, we can assume without loss of generality that $\mathbf{a}_{n_k} = \bar{\mathbf{a}}$ for all n_k sufficiently large, hence for all U_{n_k} in the subsequence with sufficiently large index we must have

$$V(U_{n_k}) = V(\bar{\mathbf{a}},U_{n_k}).$$

In the limit we hence obtain the contradiction

$$V'(\bar{\mathbf{a}},U_{n_k}) \geq V'(\mathbf{a}(U),U)+\varepsilon,$$

(indeed recall that $\mathbf{a}(U) \in \arg\max_{\mathbf{a}\in\mathcal{A}^*(U)} V'(\mathbf{a},U)$). The proof for the left hand derivative is symmetric.

We now show that V is almost everywhere differentiable. Consider the following function $g(U) := \max_{\mathbf{a}\in\mathcal{A}^*} |V'(\mathbf{a}, U)|$. Note that since V' exists and is continuous in U and $\mathcal{A}^*$ is finite the function g is continuous. Consider now the following sequence of inequalities

$$|V(U'') - V(U')| \leqq \max_{\mathbf{a}\in\mathcal{A}^*} |V(\mathbf{a}, U'') - V(\mathbf{a}, U')| \leqq \max_{\mathbf{a}\in\mathcal{A}^*} \left| \int_{U'}^{U''} V'(\mathbf{a}, U) dU \right| \leqq \int_{U'}^{U''} \max_{\mathbf{a}\in\mathcal{A}^*} |V'(\mathbf{a}, U)| \, dU \leqq \int_{U'}^{U''} g(U) dU,$$

where the first inequality uses the definition of V as the max over the conditional functions, the second one uses the fact that the conditional functions are differentiable, the penultimate inequality is obvious, while the last one is obtained by the definition of g above. We have hence shown that V is absolutely continuous hence almost everywhere differentiable. **Q.E.D.**

Proof of Proposition 5 It is immediate to see that (18), (19) and (20) are the first-order conditions for the proposed problem. Moreover, notice that the existence of $V'(U, h)$ is justified by Proposition 8. However, we must show *(I)* that the differentiability conditions for taking the first-order conditions are indeed satisfied, and *(II)* that $\mu \geq 0$, as claimed. Before finishing the proof we need a simple lemma.

Lemma 14 *Assume that f is a continuous function that admits both right and left derivatives in an interior point U_0. If U_0 maximizes f we must have $f'_-(U_0) \geq f'_+(U_0)$.*

Proof of the Lemma. The proof is standard. For completeness, we briefly show the result. First, if U_0 is optimal we must have $f'_-(U_0) \geq 0$. Indeed, whenever $f'_-(U_0) < 0$, the incremental ratio for the left derivative implies that for U sufficiently close to U_0 (but still $U - U_0 < 0$), we have $f(U) > f(U_0)$, which contradicts the optimality of U_0. A similar argument can be used to show that optimality of U_0 implies $f'_+(U_0) \leq 0$. **Q.E.D.**

(I) Since the case with $a^* = 0$ is obvious, we will consider only $a^* = 1$. When $a^* = 1$, the incentive constraint (7) can be rewritten as follows:

$$U^e - U^u \geq \frac{v}{\beta\left[\pi(h) - \hat{\pi}(h)\right]}. \tag{30}$$

We can have two cases.

Case 1: At the optimum the incentive constraint (30) is satisfied with equality. If we rewrite the objective function using (30) with equality and use (6), we can rewrite the problem as a function of U^u alone:

$$\sup_{U^u} -u^{-1}\left(U - \beta U^u + \frac{\hat{\pi}(h)v}{\pi(h) - \hat{\pi}(h)}\right) + \beta\left[\pi(h) W\left(U^u + \frac{v}{\beta\left[\pi(h) - \hat{\pi}(h)\right]}\right) + (1 - \pi(h))\, V(U^u, h')\right].$$

The problem is now a free maximization whose objective function is a weighted sum between the differentiable functions u^{-1} and W, and the function $V(U^u, h')$. We can directly apply Lemma 14 to this problem and obtain the desired result.

Case 2: The optimum is such that the incentive constraint (30) is slack. In this case, we can use (6) and rewrite the problem as a function of both U^u and U^e as follows:

$$\sup_{U^u,U^e} -u^{-1}\left(U + v - \beta U^u - \beta\pi(h)\left(U^e - U^u\right)\right) + \beta\left[\pi(h)W(U^e,h) + (1-\pi(h))\,V(U^u,h')\right].$$

Notice that in the objective function the two choice variables U^u and U^e interact in a very peculiar way. Either they are part of a linear mapping into a differentiable function (this is the case of the first term of the objective function, the term inside u^{-1}) or they enter into two different functions which are linearly related to each other. This feature guarantees that when taking the directional derivative for optimality we can separate the two variables. The choice of U^e is clearly well defined since both u^{-1} and W are differentiable everywhere. Moreover, for any given choice of U^e, the optimal level U^u is now computed by solving again a *free* maximization over a weighted sum between the differentiable function u^{-1} and the function $V(U^u,h')$; thus Lemma 14 also applies to this case.

(II) Consider again the $a^* = 1$ case. Once we have shown that the problem must be differentiable at the optimum, we can use the (local) Kuhn-Tucker theorem. For this, notice that the incentive constraint (30) is linear, hence satisfies the constraint qualification requirement needed to apply the Kuhn-Tucker theorem. Hence, if μ is the multiplier associated with the incentive constraint, μ is nonnegative, as claimed. **Q.E.D.**

Proof of Corollary 6 The first part of the corollary is easily derived from the last result of Proposition 5. It suffices to use Proposition 8 and rewrite $V'(U,h) = -\frac{1}{u'(b_t^*)}$, $W'(U^{e*},h') = -\frac{1}{u'(w_{t+1}^*)}$ and $V'(U^{u*},h') = -\frac{1}{u'(b_{t+1}^*)}$. To show the second part, notice that since $\pi(1,h) > 0$, both results (i) and (ii) can be easily derived from (18), (19), (20), $\mu \geq 0$ and the strict concavity of u. Obviously, if $a_t^* = 0$ then $w_{t+n}^* = b_t^* = b_{t+n}^*$ for $n \geq 1$. **Q.E.D.**

References

[1] Abreu, D., D. Pearce and E. Stacchetti, "Towards a Theory of Discounted Repeated Games with Imperfect Monitoring," Econometrica, 58(5), (1990), 1041-1064.

[2] Addison, J., and P. Portugal, "Job Displacement, Relative Wage Changes and Duration of Unemployment," Journal of Labor Economics, 7(3), (1989), 281-302.

[3] Alba-Ramirez, A., and R. B. Freeman,"Jobfinding and Wages when Longrun Unemployment Is Really Long: The Case of Spain," NBER Working Paper 3409, 1990.

[4] Aliprantis, C. D., and K. C. Border, Infinite Dimensional Analysis, Second Edition, (Berlin: Springer-Verlag, 1994).

[5] Attanasio, O. P., and G. Weber, "Consumption Growth, the Interest Rate and Aggregation," Review of Economic Studies, 60(3), (1993), 631-649.

[6] Bartel, A. P., and G. J. Borjas, "Wage Growth and Job Turnover: An Empirical Analysis," in S. S. Rosen (Ed.), Studies in Labor Markets, University of Chicago Press for National Bureau of Economic Research, (1981), 65-90.

[7] Benveniste, L. M., and J. A. Scheinkman, "On the Differentiability of the Value Function in Dynamic Models of Economics," Econometrica, 47(3), (1979), 727-732.

[8] Bover, O., M. Arellano and S. Bentolila, 'Unemployment Duration, Benefit Duration and the Business Cycle," The Economic Journal, 112(479), (2002), 223-265.

[9] Cogan, J. F.,"Fixed Costs and Labor Supply," Econometrica, 49(4), (1981), 945-963.

[10] Coles, M., and J. Smith, "Marketplaces and Matching," International Economic Review, 39(1), (1998), 239-254.

[11] Daskin, J. M., The Theory of Max-Min and Its Applications to Weapon Allocation Problems, (New York: Springer-Verlag, New York, 1967).

[12] Diamond, P., "Income Taxation with Fixed Hours of Work," Journal of Public Economics, 13(1), (1980), 101-110.

[13] Eckstein and Wolpin, "Dynamic Labour Force Participation of Married Women and Endogenous Work Experience,"Review of Economic Studies, 56(3), (1989), 375-390.

[14] Fallick, B. C., "A Review of Recent Empirical Literature on Displaced Workers," Industrial and Labor Relations Review, 50(1), (1996), 5-16.

[15] Fredriksson, P., and B. Holmlund, "Improving Incentives in Unemployment Insurance: A Review of Recent Research," IFAU Working Paper 2003:5, Uppsala University, 2003.

[16] Gregg, P., and B. Petrongolo, "Stock Flow Matching and the Performance of the Labor Market," IZA Discussion Paper 723, IZA, 2002.

[17] Grossman, S. J., and O. D. Hart, "An Analysis of the Principal-Agent Problem," Econometrica, 51(1), (1983), 7-45.

[18] Hopenhayn, H., and J. P. Nicolini, "Optimal Unemployment Insurance," Journal of Political Economy, 105(2), (1997a), 412-438.

[19] Hopenhayn, H., and J. P. Nicolini, "La Reforma de Seguro de Desempleo en España: ¿ Hay Algo que Aprender de la Teoría ?" Moneda y Credito, 204, (1997b), 263-303.

[20] Jacobson, L. S., R. J. LaLonde, and D. G. Sullivan "Earning Losses of Displaced Workers," American Economic Review, 83(4), (1993), 695-709.

[21] Judd, K. L., Numerical Methods in Economics, (Cambridge: MIT Press, 1998).

[22] Karni, E., "Optimal Unemployment Insurance: A Survey," Southern Economic Journal, 66(2), (1999), 442-465.

[23] Keane, M. P., and K. I. Wolpin, "The Career Decisions of Young Men," Journal of Political Economy, 105(3), (1997), 473-522.

[24] Kim, T., "The Differentiability of the Value Function: A New Characterization," Seoul Journal of Economics, 6(3), (1993), 257-265.

[25] Lopez Lopez, M. T., "La Proteccion Social a la Familia en España y en los Demas Estados Miembros de la Union Europea," Working Paper, Fundacion BBV Documenta, Centro de Estudios de Economia sobre el Sector Publico, 1996.

[26] Machin, S., and A. Manning, "The Causes and Consequences of Long-Term Unemployment in Europe," in O. Ashenfelter and D. Card (Eds), Handbook of Labor Economics, volume 3C, (Amsterdam: Elsevier, 1999).

[27] Martin, J. P., "What Works among Active Labour Market Policies: Evidence from OECD Countries' Experiences," OECD Economic Studies, 30, (2002), 79-113.

[28] Meyer, B. D., "Unemployment Insurance and Unemployment Spells," Econometrica, 58(4), (1990), 757-782.

[29] Milgrom, P., "Envelope Theorems," Working Paper 99-016, Stanford University, 1999.

[30] Milgrom, P., and I. Segal,"Envelope Theorems for Arbitrary Choice Sets," Econometrica, 70(2), (2002), 583-602.

[31] Neal, D., "Industry-Specific Human Capital: Evidence from Displaced Workers," Journal of Labor Economics, 13(4), (1995), 653-677.

[32] Pavoni, N., Three Essays on Dynamic Contracts, with Applications to the Labor Market, Ph.D. dissertation, Universitat Pompeu Fabra, 2000.

[33] Pavoni, N., "On Optimal Unemployment Compensation," Journal of Monetary Economics, 54(6), (2007), 1612-1630.

[34] Phelan, C., and R. M. Townsend, "Computing Multi-Period, Information-Constrained Optima," Review of Economic Studies, 58(5), (1991), 853-881.

[35] Rogerson, W., "Repeated Moral Hazard," Econometrica, 53(1), (1985), 69-76.

[36] Rosolia, A., and G. Saint-Paul, "The Effect of Unemployment on Subsequent Wages in Spain," UPF Economics and Business Working Paper 295, Universitat Pompeu Fabra, 1998.

[37] Ruhm, C. J., "The Economic Consequences of Labor Mobility," Industrial and Labor Relations Review, 41(1), (1987), 30-49.

[38] Saez, E., "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses," Quarterly Journal of Economics, 117(3), (2002), 1039-1073.

[39] Sah, R., and J. Zhao, "Some Envelope Theorems for Integer and Discrete Choice Variables," International Economic Review, 39(3), (1998), 623-634.

[40] Shavell, S., and L. Weiss, "The Optimal Payment of Unemployment Insurance Benefits over Time," Journal of Political Economy, 87(6), (1979), 1347-1362.

[41] Spear, S. E., and S. Srivastava, "On Repeated Moral Hazard with Discounting," Review of Economic Studies, 54(4), (1987), 599-617.

[42] Stokey, N. L., and R. E. Lucas, with E. C. Prescott, Recursive Methods in Economic Dynamics, (Cambridge: Harvard University Press, 1989).

[43] Usami, Y., Payroll-Tax Financed Unemployment Insurance with Human Capital, Ph.D. dissertation, Massachusetts Institute of Technology, 1983.

[44] Van den Berg, G. J., and J. C. van Ours, "Unemployment Dynamics and Duration Dependence in France, the Netherlands and the United Kingdom," Economic Journal, 104(423), (1994), 432-443.

[45] Van den Berg, G. J., and J. C. van Ours, "Unemployment Dynamics and Duration Dependence," Journal of Labor Economics, 14(1), (1996), 100-125.