On optimal unemployment compensation

Nicola Pavoni

Department of Economics, Institute for Fiscal Studies, University College London, Gower Street, London WC1E 6BT, UK

Received 20 October 2005; received in revised form 22 June 2006; accepted 27 June 2006
Available online 29 November 2006

Abstract

The design of an optimal unemployment compensation scheme is analyzed, using a dynamic principal–agent relationship between a risk-neutral planner (the principal) and risk-averse workers (the agents), where the planner’s inability to observe workers’ job-search efforts creates a moral hazard problem. To design an implementable scheme, we require that each agent is guaranteed a minimum level of expected discounted utility, regardless of his past history. In contrast with previous studies, we find that the optimal contract is quite close to actual unemployment compensation schemes, both qualitatively and quantitatively.

JEL classification: D63; D74; D82; D83; H55; I38; J65

Keywords: Unemployment insurance; Recursive contracts; Moral hazard; Equity and social programs

1. Introduction

Unemployment compensation programs are an important ingredient of social welfare policies in developed economies. These programs have been widely criticized because of the adverse effects they can have on worker’s incentives to search for a new job. This criticism...
has stimulated extensive research into optimal insurance schemes that take these perverse effects into account.

This paper analyzes the design of an optimal unemployment compensation scheme within a dynamic principal–agent relationship between a risk-neutral planner (the principal) and risk-averse workers (the agents), where the planner’s inability to observe workers’ job-search efforts creates a moral hazard problem.

In order to obtain an implementable scheme, we consider the possibility that, when designing the optimal program, the planner must respect a lower bound on the expected discounted utility that the agent can have ex post regardless of the previous history.

This restriction on the contract space can be rationalized in several ways. First, it may be impossible for the planner to enforce, ex post, extremely punitive plans, for example, because these would imply excessive social conflict costs. Second, workers may have ways of opting out of the insurance scheme. This could be the case if unemployed workers could find a job in the informal economy, if there are migration possibilities, or if the worker can simply leave the program opting for self-sufficiency or relying on family support.1

The resulting optimal contract turns out to be quite close to actual unemployment compensation schemes, both qualitatively and quantitatively. First, benefit payments decrease steeply with time, and involve a “last jump” to a minimum. We refer to this minimum level of benefits as the long-term or subsistence level. When the unemployment benefits reach this level, the transfers stay constant during the remaining unemployment spell and the level of payments depends only on the lower bound on the agent’s expected discounted utility. Second, once a long-term unemployed worker finds a job, he pays a fixed tax (which can be equal to zero) that does not depend on the length of the long-term unemployment period.2

A quantitative analysis is also performed, using data for the Spanish economy. We find that the actual scheme is not too far from the optimal one. However, according to our model, the Spanish unemployment compensation system is too generous during intermediate periods of unemployment, and it provides benefit payments which are too low for both very-short-term and long-term unemployed workers, compared to the optimal ones.

A series of papers use the dynamic moral hazard model to analyze the trade-off between (unemployment) insurance and (search) incentives (see, the Literature Review section). All these models predict that unemployment insurance benefits should decrease with unemployment duration, a qualitative characteristic that seems to be shared by most

---

1The possibility of introducing a minimum utility bound has been considered (and rationalized) by other authors in different contexts. For example, Atkeson and Lucas (1995) use a dynamic interpretation of the infinite horizon model. Under this interpretation, an arbitrarily low utility level after some date means that a huge burden is imposed on future generations since ancestors could, and would, sell the consumption of their heirs without limit.

2Most OECD countries’ unemployment compensation schemes have two types of benefits. The unemployment insurance system and the assistance system. The unemployment assistance system grants supplementary income to workers who have exhausted the insurance system benefits or who do not qualify for receiving them. Unemployment insurance systems in EFTA countries have shorter duration periods than other countries. For EFTA countries, the average duration is around 0.5/1 year, whereas in many other OECD countries the transfers may continue for an indefinite number of periods. The behavior of the unemployment benefit payments always presents a downward jump after a relatively short period of time (from 6 months to 3 years) and an almost fixed level of transfers thereafter (see Reissert and Schmid, 1994; Kalisch et al., 1998; OECD, 1994; Layard et al., 1994, Chapter 1; Blöndal and Pearson, 1995).
existing schemes. However, the optimal unemployment insurance programs derived from the standard dynamic moral hazard model also have certain features that contrast with what we observe in the real world. First, the optimal contract implies that benefit payments decrease slowly with time and never stop decreasing. Second, whenever reemployment wage taxes are allowed the optimal program imposes a tax on the wage the worker receives when he finds a job, which is increasing in the length of worker’s previous unemployment spell. In contrast, most OECD countries’ unemployment compensation schemes pay a decreasing level of benefits for a fixed period of time and a constant minimum level thereafter. Moreover, none of the observed schemes presents duration-dependent wage taxes. Third (partially given by the combination of the first two points), the difference between the after-tax wage and the unemployment benefit implied by the standard dynamic moral hazard model is rather small compared with existing unemployment compensation schemes.

The discrepancy between optimal and actual programs has obvious implications for government policies. For example, Hopenhayn and Nicolini (1997a) (HN hereafter) argue that by switching from the existing policy to the optimal transfer scheme, the US government could save approximately between 15% and 30% of the overall spending on unemployment compensations. Hopenhayn and Nicolini (1997b) find costs reduction of the same magnitude for Spain.

In this paper we argue that most of the discrepancies between the optimal schemes derived as a solution of the standard moral hazard model and those implemented throughout most developed countries arise from the assumption in the former that the planner can, and will, inflict infinite punishments on workers. In particular, we show that the optimal contract derived from the standard dynamic moral hazard model implies a weaker form of what is known as the immiserization result: if the worker’s utility function is unbounded below then efficiency requires that worker’s expected discounted utility falls, with positive probability, below any arbitrary negative level.

We now use a graphical representation of our results to explain the reason why relaxing the infinite punishment assumption leads to an optimal contract which is quite close to existing unemployment compensation schemes. Fig. 3 displays a parametrized version of the optimal unemployment compensation scheme implied by the unrestricted model (dotted lines), i.e. the standard dynamic moral hazard model, and compares it with the restricted model (in solid lines), i.e. the model with utility bounds that we propose in this paper. In both cases, upper lines represent the net wage $w_t$—i.e. the after-tax wage a worker should receive once he finds a new job—and lower lines represent the replacement ratios for unemployment compensation benefits $b_t$. Both $w_t$ and $b_t$ are drawn as a function of the unemployment period, in percentage terms of the gross wage. One should easily notice that in the unrestricted case the two lines are very close to each other. This is so since the use of dynamic incentives allows the planner to reduce the within-period difference between unemployment benefits and after-tax wage, by back-loading part of the punishments. In the restricted model, this mechanism works only at the beginning of the unemployment spell. For long-term unemployed workers the lower bound constraint on expected discounted utility is binding. As a consequence, the scope for dynamic incentives vanishes (since $b_t$ cannot be lowered anymore), and the difference between the

---


4In the figure, this occurs after 9 months of unemployment.
unemployment benefits and the net wage must be much larger than that in early stages of unemployment. Thus, an optimal contract implies that the difference between after-tax wage and unemployment benefits increases during unemployment spell. Since \( w_t \) cannot grow with \( t \) because of consumption smoothing, the optimal way to increase the difference between \( w_t \) and \( b_t \) is to design a flat net-wage schedule (implying a constant wage tax) together with a steeply decreasing unemployment benefits path. Finally, when the unemployment benefits reach their minimum level, the transfers stay constant forever, due to the presence of the minimum utility bound.

1.1. Literature review

In their seminal work on unemployment insurance, Shavell and Weiss (1979) establish that, because of moral hazard, benefits must decrease throughout the unemployment spell, approaching zero in the limit. In an influential paper, HN, extend the analysis of Shavell and Weiss, increasing the number of policy instruments available to the government. Together with the sequence of benefits paid to the unemployed workers, they introduce the possibility of contingent wage taxes after reemployment. They show that in the optimal program the reemployment tax is increasing in the length of worker’s previous unemployment spell. Other authors extend this model in several dimensions (e.g. Hopenhayn and Nicolini, 2002; Zhao, 2001; Pavoni, 2003). The literature on optimal unemployment insurance is relatively new, yet quite extended. However, most of the remaining papers address questions and use approaches that cannot be directly related to our own. The interested reader can refer to the summary by Karni (1999).

A minimum bound on expected discounted utility has already been introduced by other authors. Atkeson and Lucas (1995) characterize the optimal contract in a pure adverse-selection setup with temporary (one-period) job offers. They are mainly interested in income distribution, and their approach is closely related to that in Hansen and Imrohoroglu (1992), and Wang and Williamson (2002) where the goal is to quantify the welfare effects on unemployment insurance in general equilibrium. Phelan (1995) modifies the repeated insurance problem with adverse selection of Green (1987) and Thomas and Worrall (1990) and shows the existence of a non-degenerate long-run distribution of consumption. Phelan (1993) analyzes a repeated moral hazard problem between firms and workers and shows that (efficiency) wages should increase with tenure. Finally, Wang and Williamson (1996) provide calibrated dynamic OLG models with moral hazard associated with search effort and job retention. Following Phelan (1994), they assume that each new labor-force entrant obtains a prespecified level of ex ante utility.

The methodology we use is often called “recursive contracts”, and some references related to our approach are Spear and Srivastava (1987), Abreu et al. (1990), Fudenberg et al. (1990) and Phelan and Townsend (1991).

The paper is organized as follows. In the next section we briefly introduce the model and its recursive formulation. In Section 3 we present the unrestricted version of the model and show the immiserization result. In Section 4 we propose the restricted model with utility bounds, and we characterize qualitatively the optimal contract. In Section 5 we study the quantitative features of the restricted model and compare them both with those of the unrestricted model and with those of the existing programs. Section 6 concludes.
2. Model

We consider a search model with informational problems and our starting point is HN. Suppose that a risk-neutral planner must design an optimal unemployment compensation scheme for a risk-averse worker. In any given period, the worker can be either employed or unemployed. Jobs are permanent: if a worker is employed, he produces $y$ forever. If the worker is unemployed, he can engage in costly job search, and the higher the search effort $a$, the higher his probability $p(a)$ of being employed in the next period (Table 1).

The optimal unemployment compensation would be straightforward if the planner could observe the search effort. In this case, the worker would be fully insured against employment risk, would enjoy constant consumption and it would be possible to implement the action that maximizes total welfare, i.e. the first-best action.

Suppose however that the planner cannot observe $a$. In this case, unemployment benefits are not paid only as insurance, but must also play the role of giving incentives for search. The problem can be formalized using the following contractual terminology. At time zero, the principal planner offers a contract $W$ to the agent. This contract is optimal in the sense that it guarantees an initial utility level $U_0$ to the worker, minimizing the expected discounted value of the net transfers to the agent. For each date $t$ the contract $W$ specifies a net transfer $c_t$ and a recommended action $a_t$ as a function of the realized history $h_t = (h_t, \ldots, h_{t-1})$. The (public) history $h_t$ is a vector of zeros and ones, where we denote $h_t = 1$ if the worker is employed at the beginning of period $t$ and $h_t = 0$ if he is unemployed. Thus, the principal planner designs the unemployment compensation scheme $W$ choosing optimally a pair of functions $\{c_t(h_t), a_t(h_t)\}_{t=0}^\infty$ taking into account the moral hazard problem in the unemployment state. The planner knows that the worker orders the implied stochastic processes of consumption and search efforts $\{c_t, a_t\}$ according to

$$\mathbb{E} \sum_{t=0}^\infty \beta^t [u(c_t) - v(a_t)],$$

(1)

where the probability distribution, with respect to which the expectation in (1) is taken, depends on the implied sequences of search efforts. We assume that the utility function $u(\cdot)$ is strictly increasing and strictly concave while the effort cost function $v(\cdot)$ is assumed to be strictly increasing in $a$. We indicate by $\beta \in (0, 1)$ the intertemporal discount factor. We use the symbol $b_t$ to denote the unemployment transfer, i.e. $c_t = b_t$ if $h_t = 0$ and we use $w_t$ to indicate the net wage (if the worker is employed), i.e. $c_t = w_t$ if $h_t = 1$.

Following the recursive contracts literature, the contract can be formulated within the dynamic programming framework using current employment situation $h_t$ and worker’s expected discounted utility $U_t$ as state variables. Consider first the unemployment state case ($h_t = 0$), and let $U$ be the discounted utility promised to the agent at the beginning of the period. Given a utility level $U$ and present state $h = 0$, recursivity means that the problem can be stated in terms of three functions $[a(U), b(U), U'(0, U, h')]$ determining the current action $a_t = a(U_t)$, the current net transfer $b_t = b(U_t)$ and a promised future utility $U_{t+1} = U'(0, U_t, h_{t+1})$, $h_{t+1} \in \{0, 1\}$, which is contingent on the search process outcome.

---

5In the previous version of the paper we allow for a positive firing probability. The assumption of permanent jobs does not change any of our qualitative results.

6Pavoni (1999) contains a formal proof of this recursivity result.
If we define \( U_u(0, U)/C_1 \) \( U(0, U)/C_1 \) and \( U_e(0, U)/C_1 \) \( U(0, U)/C_1 \), the choice of the functions \( a(U), b(U), U^e(0, U), U^u(0, U) \) must satisfy the following two sets of constraints:

\[
U \geq u(b(U)) - v(\hat{a}) + \beta[p(\hat{a})U^e(0, U) + (1 - p(\hat{a}))U^u(0, U)] \quad \forall \hat{a}. \tag{2}
\]

\[
U = u(b(U)) - v(a(U)) + \beta[p(a(U))U^e(0, U) + (1 - p(a(U)))U^u(0, U)]. \tag{3}
\]

Constraint (2) is the incentive compatibility constraint ensuring the agent is willing to deliver the amount of effort called for in the contract. Eq. (3) requires the contract to deliver the promised level of discounted utility to the worker, and is called the promise keeping constraint and plays the role of law of motion for the state variable \( U \). The equivalent promise keeping equation in the employment state is

\[
U = u(w(U)) - l + \beta U^e(1, U), \tag{4}
\]

where \( w(U) \) is the net wage the worker receives after tax \( (\tau = y - w) \) is paid, and we assume a fixed effort cost of working equal to \( l \geq 0 \). The starting value \( U_0 \) will be given by the time-zero participation constraint of the contract \( U^e \).\(^7\)

\(^7\)To clarify further the law of motion interpretation of the promise keeping constraints, let us be redundant in notation and rewrite (3) and (4) as follows:

\[
U_t^u = u(b_t) - v(a_t) + \beta[p(a_t)U_t^{e} + (1 - p(a_t))U_t^{u}],
\]

\[
U_t^e = u(w_t) - l + \beta U_{t+1}^{e}.
\]

It should now be easier to see that each equation implicitly maps values of the endogenous state variable in period \( t \) into values of the state variable in period \( t + 1 \), as a function of today’s controls.
Recursivity uses the fact that, under some conditions, second-best contracts have the property that any continuation of the contract is still on the constrained utility possibility frontier. Our unemployment compensation problem has two different utility possibility frontiers which are precisely the planner’s value functions in the two states. The planner’s value function in the unemployment state $V$ is defined as

$$V(U) = \sup_{a,b,U^u,U^e} - b + \beta[p(a)W(U^e) + (1-p(a))V(U^u)]$$

s.t. (2) and (3).

The problem in the employment state is much simpler since there are no incentive problems. The associated planner’s value function $V_e$ is

$$W(U) = \sup_{w,U^e} y - w + \beta W(U^e)$$

s.t. (4).

3. Unrestricted case: the immiserization result

In what follows we solve the problem when there are only two levels of effort. We assume $a \in \{0,1\}$, i.e. the worker can either “search” ($a = 1$) or “not search” ($a = 0$). We define $p(1) = p \in (0,1)$ and $p(0) = 0$, and normalize the search costs by setting $v(1) = v > 0$ and $v(0) = 0$. Finally, we assume it is always optimal to induce the agent to supply $a = 1$ while unemployed.\footnote{Intuitively, this assumption will always be met for a sufficiently high gross wage level $y$. Equivalently, as we will see below, the exact condition specifies a range of utilities as a function of the whole space of parameters.}

Analogous to the general case, the value functions $V$ and $W$ for a planner facing, respectively, an unemployed and an employed worker to whom a level $U$ of utility is promised are defined by

$$V(U) = \sup_{b \geq 0, U^u, U^e} - b + \beta [pW(U^e) + (1-p)V(U^u)]$$

s.t. $U^e \geq U^u + \frac{v}{bp}$, $U = u(b) - v + \beta [pU^e + (1-p)U^u]$ (5)

and

$$W(U) = \sup_{w \geq 0, U^e} y - w + \beta W(U^e)$$

s.t. $U = u(w) - l + \beta U^e$, (8)

where (6) and (7) are the incentive compatibility and promise keeping constraints, respectively. One can easily show that, since jobs are permanent, $W$ takes the following form:

$$W(U) = \frac{y - u^{-1}((1-\beta)U + l)}{1-\beta}.$$
By the properties of $u$, $W$ is strictly decreasing and concave. Moreover, if $u$ is continuously differentiable so is $W$. In the technical appendix (Pavoni, 2006) we show that these properties are inherited by $V$ as well.

Before stating the main result of this section we need two Lemmas.

**Lemma 1.** The incentive compatibility constraint (6) is always binding at the optimum.

The idea behind this result is the following. As long as incentive compatibility implies a net wage which is larger than the UI-benefit payment, the planner tends to insure the agent minimizing the difference between $U^e$ and $U^u$. The assumption $l \geq 0$ implies this property.9

**Lemma 2.** For all $U$ in the interior of the effective domain of $V$, we have $U^u(0, U) < U$.

This result summarizes the key property of all models of dynamic moral hazard. Intuitively, Lemma 2 together with consumption smoothing implies that unemployment compensation benefit payments should decrease during unemployment.

We now use this property of the optimal contract to show that if no bound is imposed on the lifetime utility promised to the agent, worker's utility may fall with positive probability below any arbitrarily low level.10

**Proposition 3.** Assume $u$ is unbounded below, and fix an arbitrary level of utility $U > -\infty$ and an initial level of utility $U_0 > U$. Then there is a positive probability that the worker's expected discounted utility falls below $U$.

Thomas and Worrall (1990), in a model with adverse selection and i.i.d. shocks, prove a similar but much stronger result. They show that if the utility is bounded above, unbounded below and displays non-increasing absolute risk aversion, any arbitrary lifetime utility level $U$ is reached with probability one. This feature of the optimal contract is known in the literature as the immiserization result.

**Remark 4.** HN, in their Proposition 1, show that as long as $V$ is concave then $U^u(0, U) < U$. Notice that to show Proposition 3 we only used Lemma 2. This observation implies that under concavity our immiserization result still holds with an arbitrary number of actions.

### 3.1. A closed form example

It can be verified directly that when the utility of the agent takes the logarithmic form, i.e. $u(c) = \ln(c)$, a solution to the functional equation (5)–(8) is

$$V(U) = \frac{\gamma T}{1 - \beta} - \frac{\Lambda \exp((1 - \beta)U)}{1 - \beta},$$

(10)
\[
W(U) = \frac{y}{1 - \beta} - \frac{\exp[I] \exp((1 - \beta)U)}{1 - \beta},
\]
(11)

where \( I \) and \( A \) are suitable constants.\(^{11}\) An added bonus is that the implied policies are linear functions in the space of utilities. In particular, for the unemployment state, the policy functions are

\[
u(b(U)) = (1 - \beta)U + \ln A,\]
(12)

\[
U^u(0, U) = U - \frac{\ln A}{\beta},\]
(13)

\[
U^e(0, U) = U + \frac{v/p - \ln A}{\beta}.
\]
(14)

The employment state policy is simply a constant utility \( U^e(1, U) = U \) for each period, guaranteed by a within-period net wage equal to

\[
u(w(U)) = (1 - \beta)U + l.
\]
(15)

By looking at the optimal policies, one can easily verify that—in accordance with Proposition 3—if the worker stays unemployed for a sufficiently long period, his utility may reach arbitrarily low levels. For any two values \( U_0, U \) with \( U_0 > U \), we can indeed define \( A = U_0 - U \) and \( \delta = U - U^u(0, U) \). From (13) we get \( \delta = \ln A/\beta > 0 \), which is independent of the level of the state variable \( U \). As a consequence, \( U \) would be reached with probability \( (1 - p)^{[A/\beta]} > 0 \).

4. Restricted case: imposing utility bounds

A key feature of the present paper is to impose a lower bound \( U_{\text{min}} \) on the expected discounted utility that can be assigned to the agent from any date onward. From the simplicity of the new constraint, it should be easy to see that the problem remains recursive in a very natural way. The Bellman equation for a planner facing an unemployed worker to whom a level \( U \) of utility is promised is now given by

\[
\hat{W}(U) = \sup_{b, U^u, U^e} \left[ b + \beta[p \hat{W}(U^e) + (1 - p)\hat{W}(U^u)] \right]
\]
(16)

\[
s.t. \quad (6), (7) \text{ and } \quad U^e, U^u \geq U_{\text{min}}.
\]
(17)

The problem in the employment state is analogously defined as follows:

\[
\hat{W}(U) = \sup_{w, U^e} \left[ y - w + \beta \hat{W}(U^e) \right]
\]
(18)

\[
s.t. \quad (9) \text{ and } \quad U^e \geq U_{\text{min}}.
\]
(19)

\(^{11}\)In particular, \( \Gamma = \beta p/(1 - (1 - p)\beta) \), and \( A \) solves

\[
\ln[A^{1/\beta} - (1 - p)A] = \ln p + \frac{1 - \beta}{\beta} \frac{v}{p} + l.
\]
Notice that we have simply added the minimum bound constraint to problem (5)–(9). To be relevant, the minimum utility restriction must of course be on the interior of the effective domain of $V$. To this extent, we only consider $U_{\text{min}} > u(0)/(1 - \beta)$.

4.1. Characterization of the optimal compensation scheme

In this section, we study the qualitative characteristics of the optimal contract. Similar to the unrestricted case, one can easily show that the newly defined value functions are strictly decreasing, strictly concave and continuously differentiable. One can also show that the incentive compatibility is binding.\(^{12}\)

We now take advantage of these properties to characterize the optimal contract using first order conditions. The first order conditions for the unemployment state problem are

\begin{align*}
\dot{V}'(U) &= -\frac{1}{u'(b(U))}, \quad (20) \\
\dot{W}'(U'u(0, U)) &= -\frac{1}{u'(b(U))} - \mu, \quad \mu \geq 0, \quad (21) \\
\dot{V}'(U'u(0, U)) &\leq -\frac{1}{u'(b(U))} + \frac{p}{1-p} \mu, \quad (22) \\
\dot{V}'(U) &\geq [p\dot{W}'(U'u(0, U)) + (1-p)\dot{V}'(U'u(0, U))], \quad (23)
\end{align*}

If $U'u(0, U) > U_{\text{min}}$ then both (22) and (23) are satisfied with equality; together with (6), (7), (17). The multiplier $\mu$ is the one associated with the incentive compatibility constraint (6). Notice that condition (21) for $U'u$ is assumed—without loss of generality—to hold as an equality. This is so since incentive compatibility implies that the lower bound constraint can be binding only for $U'u$. Eq. (20) represents the envelope condition. Finally, note that condition (23) is derived from (21) and (22). The optimality conditions for the employment state problem are

\begin{align*}
\dot{W}'(U) &= -\frac{1}{u'(w(U))}, \quad (24) \\
\dot{W}'(U) &\geq \dot{W}'(U'u(1, U)), \quad (25)
\end{align*}

if $U'u(1, U) > U_{\text{min}}$ then (25) is satisfied with equality.

We are now ready to characterize the optimal compensation scheme. Let us first consider the case where the minimum bound constraint is not binding.

**Proposition 5.** For each $U \geq U_{\text{min}}$ such that $U'u(0, U) > U_{\text{min}}$ the policy functions satisfy: (i) for $i = u, e$ and $h \in \{0, 1\}$, $U'h(U)$ is increasing in $U$; (ii) $b(U)$ and $w(U)$ are strictly increasing; (iii) $U'u(1, U) = U$; (iv) $U > U'u(0, U)$ and $U'u(0, U) > U'u(0, U)$.

The monotonicity properties reported in (i) and (ii) are induced by consumption smoothing. Result (iii) implies that the worker is fully insured when employed. This is intuitive since there are no informational restrictions in this state. Finally, for later use note

\(^{12}\)The argument is as follows. If we show that it cannot be that both the incentive compatibility constraint is slack and that (22) is a strict inequality we are allowed to use the same argument as the one used in Lemma 1 also when $U'u = U_{\text{min}}$. However, if the above statement were true then from (20) to (23)—since if $\mu = 0$ we have $\dot{V}'(U) = \dot{W}'(U'u)$—we get $\dot{V}'(U_{\text{min}}) < \dot{V}'(U)$, which from strict concavity implies $U_{\text{min}} > U$ and this is impossible.
that (ii) together with $U > U^a(0, U)$ in (iv) implies that both $b$ and $w$ are (weakly) decreasing in the unemployment spell.

In the next proposition we study the optimal contract when $U^a(0, U) = U_{\text{min}}$.

**Proposition 6.** (i) There is a $U^* > U_{\text{min}}$ such that $U^a(0, U) = U_{\text{min}}$ for every $U$ that belongs to the closed interval $[U_{\text{min}}, U^*]$. (ii) Moreover, if $U \in [U_{\text{min}}, U^*]$ then the within-period utility $u(b(U))$ varies one-to-one with $U$ and the promised expected discounted utility $U^e(0, U)$ is constant and equal to $U_{\text{min}} + v/bp$ for the whole set. (iii) There is a $U^{**} > U_{\text{min}}$ such that $U^*(0, U) > U$ for any $U \in [U_{\text{min}}, U^{**}]$.

Result (i) is important since it implies that $U_{\text{min}}$ will always be reached in finitely many periods of unemployment. The idea of (ii) is that when $U^a(0, U) = U_{\text{min}}$, $u(b(U))$ is used to keep promise (7) and $U^e(0, U)$ is used to satisfy the incentive compatibility constraint (6). Result (iii) is a direct consequence of (i) and (ii).

### 4.2. Description of the optimal contract

We now use graphs to describe the results obtained in Propositions 5 and 6, and to highlight the main features of our optimal contract. To draw Figs. 1 and 2 we use the qualitative features of our scheme and we choose thin lines to draw the unrestricted policies (from the closed form solution) and thick lines for the restricted case.

Fig. 1 reports the results of this section about the unemployment state policy functions. As derived in Proposition 6, Fig. 1(a) shows that the policies for future utilities are flat near $U_{\text{min}}$. Fig. 1(b) depicts the policy $u(b(U))$, which starts at the point $U = U_{\text{min}}$ where the unemployment benefit is equal to $1 - (1 - \beta)U_{\text{min}}$, and it is easy to see that its slope must locally be equal to one.

The problem in the employment state is particularly easy. Full insurance implies $U^e(1, U) = U$ and a net wage starting from the level

$$w(U_{\text{min}}) = u^{-1}((1 - \beta)U_{\text{min}} + l),$$

and increasing with $U$ at a slope equal to $(1 - \beta)$.

Fig. 2 presents a typical example of our scheme. In Fig. 2(a) we consider a worker who starts unemployed in $A$. If he does not find a job during this period, his lifetime utility decreases to $B$ and, eventually, reaches the minimum level $U_{\text{min}}$ in $C$. When $U_{\text{min}}$ is reached, the utility is constant for the whole remaining unemployment spell. Moreover, regardless of the length of the further unemployment duration, once a job is found the expected discounted utility jumps to $D$. Fig. 2(b) describes the policy functions for the within-period utilities $u(b)$ and $u(w)$. From the monotonic relationship between utility level $U$ and unemployment benefit payment $b$, the feature described by the unemployment utility $U^a$ applies to $b$ as well. That is, $u(b)$ is decreasing until the minimum level is reached (in $C$). After that point, the scheme pays a constant transfer $b_{\text{min}}$, and, again, once the

---

13As Proposition 5(i)–(iii) shows, policies are monotone in $U$ but of course they do not have to be linear. Moreover, restricted policies would typically not overlap the unrestricted case policies. Finally, it might be the case that $U^a(0, U)$ crosses the 45° line. However, from Proposition 6(iii) this can only be for sufficiently large utility values: $U > U^{**}$. None of these features—not included into the figures—would affect our discussion below.
A worker finds a job there will be a jump to the level $D$ (on the $u(w)$ line since the state has changed) regardless of the number of periods the worker has remained in $C$.

What happens if a worker finds a job soon and his lifetime utility does not reach the minimal level at the point $C$? From Fig. 2 one can see that actually for very-short-term unemployed workers, the wage tax after reemployment is slightly increasing in the length of the previous unemployment spell. Suppose, for example, that a worker finds a job in the first period of his unemployment (while he is in $A$ in Fig. 2(a)). Then his utility $U_e$ jumps to a point that must lie above $D$ and, from Fig. 2(b) we can see that this implies a net wage which is higher (i.e. with a lower wage tax) than the one received by a long-term unemployed worker (which, as we have just seen, will jump to $D$ when a job is found). However, compared to the unrestricted case, the presence of utility bounds modifies the
characteristics of the wage tax schedule in at least two dimensions. The main (qualitative) difference is that in the restricted model the wage tax increases—slightly—only for a finite number of periods, whereas in the unrestricted model the wage tax is always increasing. Moreover, Fig. 3 shows that also the quantitative aspects are quite important.

Fig. 3 compares the results of a simulation of both our restricted model and the unrestricted one. The base period is 1 month and the unrestricted replacement ratios are computed according to our closed form solution of the unrestricted model of Section 3. Dotted lines represent the unrestricted case and solid lines the restricted one. In both cases, upper lines describe the net wage $w_t$—i.e. the wage (net of taxes) a worker should receive once he finds a new job—as a function of the unemployment period, in percentage terms of the gross wage $y$. Lower lines represent the replacement ratios for unemployment compensation benefits. From the figure one can immediately see how the net wage and the benefit payment in the unrestricted case vary smoothly together during the unemployment spell. In particular, notice that after 4 years the wage tax $(y - w_t)$ is around 20%. In contrast, our restricted contract first presents a much steeper decreasing path for unemployment compensation benefits, which become completely flat roughly after 9 months. Moreover, the net wage slightly decreases for a while, and remains constant afterwards; in particular, the implied wage tax never exceeds 5%.

Let us give an intuitive explanation of the reasons why in Fig. 3 the UI-benefit/net-wage payments behavior for the restricted model is so different from that of the unrestricted case. The

---

Fig. 3. Comparison between the restricted and the unrestricted models. The dotted lines reproduce the simulation results for the benchmark unrestricted model of HN. The solid lines reproduce our simulation results for the restricted model with utility bound. Lower lines represent benefit payments while upper lines represent after-tax wage payments, as a function of the unemployment duration. The horizontal axis represents unemployment duration (in months). The vertical axis represents replacement ratios.

---

14For expositional purposes we anticipate here some quantitative results. The calibration exercise is discussed in detail in the next section.
key forces at work are consumption smoothing and dynamic incentive provision. In the unrestricted case the planner can always fully use intertemporal incentive provision to reduce the within-period difference between unemployment benefit $b_t$ and net wage $w_t$ the worker receives after reemployment. This is the reason why the two dotted lines are so close to each other.

However, this is not the case for the restricted model with utility bounds (represented by solid lines). Consider indeed the case of a long-term unemployed worker whose lifetime utility has reached the minimum level $U_{\min}$. When the lower bound constraint is binding the scope for dynamic incentive possibilities vanishes, and the difference between the unemployment benefit and the net wage must be much larger than in earlier stages of unemployment. This is the reason why in the flat part of the unemployment benefit payments $b_t$ the difference between the two solid lines is so large compared to the one between the two dotted lines.

In the restricted case, dynamic incentive provision is allowed only at the early stages of unemployment, when the minimum bound constraint does not bind yet. Our discussion suggests that an optimal contract for the restricted model requires that the difference between after-tax wage $w_t$ and unemployment benefits $b_t$ must increase during the unemployment spell. There are only two ways of increasing the difference between the two payments: increasing the net wage and decreasing the UI benefits. However, because of consumption smoothing, $w_t$ cannot grow during the unemployment spell. Thus, the net wage can only decrease more slowly (compared with the unrestricted case) for a while, and remain constant thereafter, a feature which leads to a constant reemployment wage tax. On the other hand, the benefits must decrease very rapidly: one reason is the usual incentive provision, as in HN, the second reason is the need to increase the difference between the unemployment benefits and the net wage during the unemployment period we mentioned before. Thus, benefits decrease steeply for a finite number of periods and then become completely flat, because of the utility bound.

In contrast, in the unrestricted model, since dynamic incentives are always feasible both the net wage and the UI-benefit transfer decrease regularly with a much lower slope. Without utility bounds, unemployment benefit payments never stop decreasing and the reemployment wage tax always increases during unemployment. These features, eventually, imply arbitrarily low levels of lifetime utility, as we demonstrated in Proposition 3.

5. Quantitative analysis

We calibrate our model using data on the Spanish economy, and we normalize the gross wage $y = 100$ and set the effort costs to one ($v = l = 1$). We interpret each period as a month so we set the discount factor $\beta = 0.996$ which implies an annual interest rate of 5%.

The main departures from HN are the choice of the risk aversion parameter and the calibration of the minimum level of promised utility $U_{\min}$. HN use CRRA preferences over consumption of the form

$$u(c) = \frac{c^{1-\sigma}}{1-\sigma}$$

and choose $\sigma = 0.5$. In our quantitative exercise we set $\sigma = 1$, i.e. we use the logarithmic utility function. This is mainly done in order to use our closed form

---

15Recall the discussion we made after Proposition 5.
solution. Finally, to estimate the parameter $U_{\text{min}}$ we first set

$$ U_{\text{min}} = \frac{u(qy) - v + \beta p U_{\text{work}}}{1 - \beta(1-p)}, \quad (26) $$

where $q$ expresses the “social salary” as a function of gross wage $y$. Thus, $U_{\text{min}}$ corresponds to the expected discounted utility of receiving the social salary forever and always having the chance of getting a job, and paying no taxes. The parameter $U_{\text{work}}$ is defined as

$$ U_{\text{work}} = \frac{u(y) - l}{(1 - \beta)}, $$

and represents the utility of working forever receiving the gross wage $y$. Thus, $U_{\text{min}}$ is calibrated by setting $q = 0.20$. Note that $u(y) = \ln(100) = 4.6$, so an effort cost of one is between one-fourth and one-fifth of the wage utility, and $U_{\text{work}} = 900$.

In Fig. 3, we report the simulated replacement ratios for our restricted model, and we compare them with the ones obtained using the closed form solution of the unrestricted model. We use $p = 0.15$ and the starting value $U_0 \in (U_{\text{min}}, U_{\text{work}})$ is computed according to the existing Spanish scheme. We have commented in Fig. 3.

The benchmark level used for $p$ is an intermediate one. Bover et al. (2002) emphasize the high degree of heterogeneity between workers’ hazard rates in the Spanish labor market. Thus, in Fig. 4, we investigate how the optimal transfer scheme changes with $p$.

However, Attanasio and Weber (1993) use UK cohort data to estimate the intertemporal elasticity of substitution. In our specification given we assume CRRA preferences—the results of Attanasio and Weber imply a $\sigma$ between 1.3 and 1.5. Moreover, Baily (1977), who studies issues similar to those in this paper, argues for setting $\sigma = 1$, as we do.

The amount of the non-contributive assistance level of transfers is means tested, and varies across different Autonomous Communities between 30,000 to 45,000 pesetas ($150/180), and there is no fixed duration (see Lopez, 1996). For example, in Catalonia, the transfer is called Programa interdepartamental de la renta mínima d’inserció (PIRMI) and, in 1998, the monthly payment was approximately 40,000 pesetas ($160). The Bulletin of Labor Statistics (1999) reports as 300,000 pesetas ($1,200 ) per month the 1998 average wage in non-agricultural activities. We estimated the parameter $q$ following the common assumption that workers subject to severe unemployment risk face a wage that is two-thirds of the average national wage.

Given that we normalized the gross wage $y$ to 100, the current insurance system can be represented by a contract that has no taxes or transfers when unemployed ($w = y$), and pays a first benefit level $b^1$ of 70 for the first six months of unemployment, from the 7th to the 24th month the benefit level $b^2$ is set equal to 60 and from the 25th onward we assume the worker receives an assistance level of benefits $b^3 = 20$. The corresponding expected discounted utility value $U_0$ for an unemployed worker can be calculated backward—partially following HN—as follows. From period $T = 25$ onward the worker’s problem is stationary. The unemployment benefits are at their minimum level, the worker will be searching for a job, which will be found with probability $p$. Given that jobs are permanent, when a worker finds a job his lifetime utility is $U_{\text{work}}$. Thus, the value of his expected discounted utility $U_T$ can be computed as follows:

$$ U_T = \frac{u(b^3) - v + \beta p U_{\text{work}}}{1 - \beta(1-p)} = U_{\text{min}}, $$

where $b^3 = qy = 20$ is the non-contributive assistance level of unemployment benefits. For any $0 \leq t \leq T = 25$ we can now define the value $U_{T-t}$ recursively by

$$ U_{T-t} = u(b_t) - v + \beta p U_{\text{work}} + (1 - p) U_{T-(t-1)}, $$

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

Bover et al. find that the average monthly probabilities of exit from unemployment vary across sector (between 0.11 and 0.17), age (between 0.08 and 0.15) and the business cycle (between 0.12 and 0.18). They also estimate the hazard rate function controlling unobservable heterogeneity.
We consider four possible values for $p$: 0.08, 0.10, 0.15 and 0.20. This figure shows that the first order effect of a change in $p$ is mainly on the steepness of the payments across time. For higher levels of $p$, the benefit payments decrease more rapidly toward a common minimum replacement ratio level of approximately 35%. This simulation exercise has an immediate policy implication. The optimal unemployment compensation scheme should depend on every observable variable (such as age, labor market tightness, geographical region or industrial sector) which implies a different probability of reemployment. Interestingly, however, the wage tax does not seem to be very sensitive to changes in this parameter.

In Fig. 5 we summarize the results of our sensitivity analysis for different minimum bounds $U_{\text{min}}$ as well. They present a similar degree of robustness as those of Fig. 4. It may seem surprising that changes in $U_{\text{min}}$ do not too much affect the long-term replacement ratio (which remains approximately 35%). However, if we look carefully at the figure, we can see that the optimal contract guarantees higher minimum utility levels, mainly through a reduction in the wage tax imposed on long-term unemployed workers. Since jobs are permanent this effect is quantitatively the most important one.

In Fig. 6, we compare our simulations with the Spanish compensation scheme (the solid step-shaped line). The model seems to reproduce most of the qualitative features of the Spanish compensation system, and this confirms the analysis of Section 4.2. From a more quantitative point of view the Spanish scheme appears too generous during intermediate

---

20In 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 2 years. The assistance system pays, for up to 2 years, 75% of the minimum wage to (unemployed) workers with dependents, whose average family income is precisely below that amount. In 1998, the minimum wage was around 70,000 pesetas ($280) (Guia Laboral, 1998 and de Asuntos Sociales, 1998).
periods of unemployment, whereas it provides payments which are too low for both the very-short-term and the long-term unemployed people with respect to the optimal transfer scheme. It seems clear that part of this discrepancy can be accounted for by introducing positive administrative costs in our model. However, we think that a deeper analysis of these quantitative aspects not only requires more accurate measurements but, more

Fig. 5. Sensitivity analysis II. Here we vary the minimum utility bound $U_{\text{min}}$ by changing the value of the parameter $q$ in Eq. (26) between 0.05 and 0.30.

Fig. 6. Comparison with the Spanish compensation scheme. In this figure, the horizontal axis represents unemployment duration (in months). The thick solid step-shaped line reproduces the replacement ratios of the actual unemployment compensation in Spain. All other lines represent the optimal for different levels of $\rho$. 


importantly, probably also requires a larger set of parameters.\textsuperscript{21} This is beyond the scope of the present paper and remains for future research.

6. Conclusions

In this paper, we studied the design of an implementable optimal unemployment compensation scheme. To analyze the trade-off between unemployment insurance and search incentives, we used a dynamic principal–agent relationship between a risk-neutral planner and risk-averse workers, where the planner’s inability to observe workers’ job-search efforts induces moral hazard.

In the first part of the paper, we show that—if $u$ is unbounded below and—if no bound is imposed on the expected discounted utility promised to the agent, optimal unemployment compensation schemes derived from the standard moral hazard model imply that the worker’s lifetime utility falls with positive probability below any arbitrarily low level. This is a weaker form of what is known in the recursive contracts literature as the immiserization result. We thus argue for the importance of restricting the planner’s contract space. In particular, we require that, when designing the optimal program, the planner must respect a lower bound on the expected discounted utility that the agent can have ex post regardless of the previous history. We find that the introduction of utility bounds in the optimal unemployment compensation designing problem has important normative and policy implications.

First, the optimal restricted compensation scheme is qualitatively consistent with the existing ones, meaning that governments are not making such large mistakes as the one highlighted, for example, by HN and Hopenhayn and Nicolini (1997b). We also calibrate and simulate our model for the Spanish economy, and we find only minor quantitative differences between optimal and existing ones.

Second, one should recall that the main policy implication of the optimal program proposed by HN is that an optimizing government should impose a tax on the wage the worker receives when he finds a job, which is increasing in the length of the worker’s previous unemployment spell. We show that the introduction of utility bounds eliminates this characteristic of the optimal unemployment compensation scheme. Since a simple consumption smoothing argument implies that wage tax cannot (strictly) decrease during unemployment, we get a constant wage tax.

References


\textsuperscript{21}In particular, note that we assume permanent jobs and we use a \textit{naive} hazard function. We always allow only for two actions assuming that $p(0) = 0$, and we require time invariance of $p(a)$. See Pavoni (2003) for a more general case.