Unemployment Insurance with Hidden Savings

Mitchell, Matthew and Zhang, Yuzhe

University of Toronto,, University of Iowa

March 2010

Unemployment Insurance with Hidden Savings

Matthew Mitchell*and Yuzhe Zhang†

March 15, 2010

Abstract

This paper studies the design of unemployment insurance when neither the searching effort nor the savings of an unemployed agent can be monitored. If the principal could monitor the savings, the optimal policy would leave the agent savings-constrained. With a constant absolute risk-aversion (CARA) utility function, we obtain a closed form solution of the optimal contract. Under the optimal contract, the agent is neither saving nor borrowing constrained. Counter-intuitively, his consumption declines faster than implied by Hopenhayn and Nicolini [4]. The efficient allocation can be implemented by an increasing benefit during unemployment and a constant tax during employment.

JEL Classification Numbers: D82, D86, J65.

Key Words: hidden savings, hidden wealth, repeated moral hazard, unemployment insurance.

*Rotman School of Management, University of Toronto, 105 St. George St., Toronto, ON M5S 3E6, Canada. Email: matthew.mitchell@rotman.utoronto.ca. Phone: (416)-946-3149.

†Department of Economics, University of Iowa, 20 E Market Street, Iowa City, Iowa, 52246. Email: yuzhe-zhang@uiowa.edu. Phone: (319)-335-2290.

1 Introduction

Unemployment insurance must balance the benefits of insurance against the concern that an over-generous program will discourage search effort. Card, Chetty, and Weber [3] provide evidence that search effort varies according to an agent's financial situation. They find that exogenously richer agents take longer to find a job but do not find higher wages on their next job. In other words, it seems that wealth is one important determinant in the search effort decision.

Since neither search effort nor wealth are readily observable to the planner, it is natural to wonder how these information frictions might affect the optimal unemployment system. We introduce hidden savings into an environment similar to the Hopenhayn and Nicolini [4] version of the model of Shavell and Weiss [7]. We show that the addition of hidden savings leads to faster consumption declines during an unemployment spell than the declines in a model with observable savings. Moreover, with hidden savings, agents with relatively high initial insurance claims have the fastest rate of consumption decline, eventually having *lower* claims than those agents who started out receiving less. We show that the unemployment benefit rises over the course of the spell.

Hidden savings is naturally relevant in repeated moral hazard models like this one. When savings can be monitored, as in Hopenhayn and Nicolini [4] or the repeated moral hazard model of Rogerson [6], the optimal policy leaves the agent savings-constrained: his marginal utility is lower today than tomorrow. By making an agent poor in the future, it encourages the agent to search harder for a job.

Our paper is related to both Werning [9, 10] and Abraham and Pavoni [1, 2], who use the first-order approach to study models with hidden savings and borrowing. Briefly speaking, the first-order approach studies a relaxed problem, which replaces the incentive constraints in the original problem with some first-order conditions of the agent. In these papers, the first-order condition for the type that has never deviated in previous periods and thus has zero hidden wealth is imposed.

Werning [9] acknowledges that imposing first-order conditions may not be sufficient to ensure incentive compatibility. Furthermore, Kocherlakota [5] shows that when the disutility function is linear, the agent's problem is severely non-convex and the first-order condition cannot be sufficient. When the first-order approach is invalid, the number of state variables in a recursive formulation would be infinite, making even numerical computations intractable. We overcome this problem by focusing on a special case of constant absolute risk-aversion (CARA) utility from consumption and linear disutility of effort. In this case we conjecture and verify the countable set of constraints that bind. With this in hand, it is straightforward to explicitly solve for the principal's optimum. It has the interesting feature that the incentive constraints of the searching agent never bind. Instead, the binding incentive constraint in any period is the one for the agent who has always shirked, and meanwhile saved.

The basic intuition for this structure of binding incentive constraints relates to the way in which shirking and saving interact. When an agent shirks, he increases the odds of continuing to be unemployed. The unemployed state involves lower consumption, so he wants to save in preparation for the greater probability of this low outcome. Therefore, saving and shirking are complements. The agent who saves the most is the one who has always done maximum shirking, and who knew he would never become employed. Given that he has saved the most, he is best equipped to do additional shirking, which is, again, complementary with saving. This example shows, in two ways, the sense in which the first-order approach is not appropriate for this kind of problem: first, the complementarity between shirking and saving can make the first-order condition for effort insufficient for optimality, and second, since the binding incentive constraint is not for the agent who always searches, it is not enough to look at the always-searching agent's optimality condition in the first place.

The contract we study always implements, at an optimum, a one-time lottery over always searching as hard as possible, or always searching the least, which we refer to as not searching. The interesting case is when the agent searches. The reason for keeping the agent on the Euler equation in this case relates to the savings-constrained nature of the optimal contract in

Hopenhayn and Nicolini [4]. There, making the agent poor in the future generates incentives to search today. When the agent can freely save, it is no longer possible to keep the marginal utility of consumption high tomorrow, but there is still no reason to make it fall over time. In that case, the same logic as in the case with observable savings would suggest lowering today's marginal utility and raising tomorrow's in order to increase search incentives today and make the contract deliver utility more efficiently. As a result, the agent is on the Euler equation. As a corollary, our solution would also be the solution to a case where the agent had access to both hidden borrowing and hidden savings.

In the state-of-the-art model of optimal unemployment insurance in Shimer and Werning [8], the focus is on the wage-draw aspect of the Shavell and Weiss [7] structure.[1] The fundamental trade-off is between search time and quality of match (in terms of wages). They find that the optimal contract has constant benefits (for CARA utility, and approximately so for constant relative risk-aversion (CRRA) utility), and keeps the agent on the Euler equation. The reason why their agents are not borrowing constrained is different from ours, however. In the CARA model, the agent's reservation wage is independent of wealth; therefore, there is no incentive benefit from distorting the first-order condition, and there is the usual resource cost to the principal from the distortion. In that sense their model has no "wealth effects" of the kind emphasized by Card, Chetty, and Weber [3]. Our model, on the other hand, does have wealth effects; richer agents have less incentive to give effort than poorer agents. Hopenhayn and Nicolini [4] show that there is an incentive benefit in making the agent savings-constrained; in our model such a constraint is impossible, so the optimal contract moves to the Euler equation.

The case where search effort is a monetary cost and utility is CARA is studied in Werning [9]. Monetary search cost implies that poor agents have no greater incentives than rich agents,

[1] It is worth noting that empirical evidence in favor of both the wage-draw view and the search theoretic view of unemployment exists. We have stressed the evidence, such as Card, Chetty, and Weber [3], that emphasizes search. However, recent research suggests that actual time used in search by the unemployed is very small, suggesting that compensating the disutility of search may not be as important a feature of unemployment systems.

since the cost and benefit of search are both proportional to wealth. The main insight of this model is similar to Shimer and Werning [8]: agents are not borrowing constrained and a constant benefit sequence is needed to implement the optimal allocation.

Our results contrast with Kocherlakota [5], who studies a similar model with linear disutility of effort, but *assumes* that the principal aims at implementing interior effort. He finds that agents are borrowing constrained. In contrast, we show that the optimal contract always implements corner solutions for effort. The optimal contract does not leave the agents borrowing constrained, even if doing so is feasible for the principal.

The first-order approach remains a practical and useful method to solve hidden-savings problems when the disutility function is sufficiently convex. In particular, Abraham and Pavoni [1] numerically verify the incentive compatibility of the solution obtained with the first-order approach for a large range of convex disutility functions. Abraham and Pavoni [2] show analytically, in a two-period model, that the first-order approach is valid if the utility function has the non-increasing absolute risk aversion property, and the job-finding probability (as a function of effort) satisfies some concavity condition. Williams [11] provides sufficient conditions for the first-order approach based on the Hamiltonian in a continuous-time model.

The paper proceeds as follows. In the next section we introduce the basic model. In Section 3 we show that the optimal contract either implements high effort forever or no effort forever. In Section 4 we solve for the optimal contract to implement high effort. In Section 5 we show the important characteristics of the optimal contract, and compare them to the case without hidden savings. We then conclude. All proofs are contained in the Appendix.

2 The Problem

In this section we describe an unemployment insurance model in the spirit of Shavell and Weiss [7]. There is a risk-neutral principal and a risk-averse agent. The preferences of the agent are

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

where $c_t \in \mathbb{R}$ and $a_t \in [0,1]$ are consumption and search effort at time t, $u(c) = -\exp(-\gamma c)$ is a constant absolute risk-aversion (CARA) utility function, $\beta < 1$ is the discount factor, and E is the expectations operator. An agent can be employed or unemployed; he begins life unemployed. The choice of a_t affects the probability of becoming employed for an unemployed agent. Specifically, if an agent is unemployed in period t, then the probability of his becoming employed in period $t+1$ is πa_t, and the probability of his staying unemployed is $1-\pi a_t$. We assume $\pi \in (0,1)$, which implies that the agent might not find a job even if he exerts full effort.[2] As in Hopenhayn and Nicolini [4], employment is an absorbing state: if an agent is employed in period t, he earns wage $w > 0$ forever.

The agent's employment status is observable to others, but his choice of a_t is unobservable. The agent can also secretly hold a non-negative amount of assets. The principal can observe neither the consumption nor the savings of the agent. A contract σ in this environment specifies three sequences, $(\{c_t^U\}_{t=0}^{\infty}, \{c_t^E\}_{t=1}^{\infty}, \{a_t\}_{t=0}^{\infty})$. Given such a contract, an agent who is unemployed in period t receives compensation c_t^U from the principal. If an agent became employed for the first time in period t, then his compensation from the principal in period $s \geq t$ is c_t^E. Thus, once an agent is employed, his compensation is constant over time. It is simple to show that, because the principal and agent have the same discount factor, this smooth compensation is efficient in this economy, so, to save on notation, we impose that feature now. The contract will also recommend an effort level a_t, and, given that the principal cannot observe a_t, a_t will be designed to satisfy incentive constraints. To simplify notation, we will let $U_t = -u(c_t^U)$ and $E_t = -u(c_t^E)$. Notice that U_t and E_t are proportional to the marginal utility of consumption, where the constant of proportionality is γ.

Without loss of generality, we require that there be no incentive for the agent to save if

[2]If full effort led to sure employment, the principal could simply recommend full effort and promise grave punishment if a job were not found.

he follows the *recommended* strategy $\{a_t\}_{t=0}^{\infty}$. This is true if and only if all the first-order conditions are satisfied,

$$U_t \geq (a_t\pi)E_{t+1} + (1 - a_t\pi)U_{t+1}, \text{ for all } t \geq 0. \tag{1}$$

We refer to this condition as the Euler equation. Note that once the agent chooses a different search effort $\tilde{a}_t \neq a_t$, Eq. (1) may be violated, and the agent holds positive assets.

In the literature, a deviation in effort combined with hidden savings is called a *joint deviation* by the agent. Hidden savings is an additional friction that makes the contract more difficult to solve than a traditional dynamic contracting problem without hidden savings or borrowing, since it adds a constraint to the problem of the principal. Therefore, we introduce some notations to describe all the incentive constraints.

We must consider a variety of possible deviations. For instance, suppose $\{\tilde{a}_t\}_{t=0}^{\infty}$ is the agent's search strategy. If there is a finite $\bar{t}$, such that $\tilde{a}_t = a_t, \forall t \geq \bar{t}$, then $\{\tilde{a}_t\}_{t=0}^{\infty}$ is called a *finite-deviation* strategy. Otherwise it is called an *infinite-deviation* strategy. Let $V(\{\tilde{a}_t\}_{t=0}^{\infty})$ be an agent's ex ante utility if he uses strategy $\{\tilde{a}_t\}_{t=0}^{\infty}$ and privately saves. The incentive compatibility (I.C.) constraints will be written as

$$V(\{a_t\}_{t=0}^{\infty}) \geq V(\{\tilde{a}_t\}_{t=0}^{\infty}), \text{ for all } \{\tilde{a}_t\}_{t=0}^{\infty}. \tag{2}$$

However, in much of the following discussions, we only examine incentive constraints for finite-deviation strategies. This is without loss of generality, because the payoff of an infinite-deviation strategy can be approximated as closely as possible by those of finite-deviation strategies.

Let D_t denote the expected discounted (to period 0) disutility of efforts $\{a_s\}_{s=t}^{\infty}$ conditional on not finding a job at the beginning of period t,

$$D_t = \sum_{s=t}^{\infty} \beta^s (\Pi_{k=t}^{s-1}(1 - a_k\pi))a_s.$$

This can be generated recursively by $D_t = \beta^t a_t + (1 - a_t\pi)D_{t+1}$; D_0 would be the value of the total discounted disutility for an agent. Since disutility is linear and one unit of labor is compensated with $\beta\pi w/(1-\beta)$ units of wage income, $\beta\pi w D_0/(1-\beta)$ is the expected wage income that the principal can obtain. The expected cost for the principal is equal to the discounted value of consumption goods delivered to the agent, minus the expected wage income,

$$C(\sigma) = \sum_{t=0}^{\infty} \beta^t \left(\Pi_{s=0}^{t-1}(1-a_s\pi)\right) \left[c_t^U + \beta a_t \pi \frac{c_{t+1}^E}{1-\beta}\right] - \frac{\beta\pi w}{1-\beta} D_0.$$

The problem of the principal is to choose a contract σ to minimize $C(\sigma)$, subject to the I.C. constraints (2), the Euler equation (1), and the delivery of an initial promised utility $\bar{V}$.

3 Optimal Effort

In this section we take up the question of what levels of effort the principal will implement. We show that the principal can do best by offering an initial lottery over full effort forever ($a_t = 1, \forall t \geq 0$ is denoted by $\{1\}_{t=0}^{\infty}$), or no effort forever ($a_t = 0, \forall t \geq 0$ is denoted by $\{0\}_{t=0}^{\infty}$). Along the way we show that, at the optimally recommended effort level, the agent is always on the Euler equation; in other words, Eq. (1) holds with equality. Thus, no shirkers will be borrowing constrained. Since shirkers' efforts are below the maximum, they have incentives to hold a positive amount of savings. This result is relevant for two reasons. First, it helps us with the analysis of the principal's problem in the next section. Second, it shows that our results for the model with hidden savings carry over to the case where the agent has access to opportunities for both hidden saving and borrowing, since the ability of the principal to make the agent borrowing constrained is never used.

First we verify that, in any incentive compatible (I.C.) contract in which $a_t > 0$, the agent's consumption in period $t+1$ after an unemployment shock is below that after an employment shock, i.e., $c_{t+1}^U < c_{t+1}^E$. This result is intuitive because punishment is necessary to induce effort. It merits a proof because although the continuation utility after a good shock is $\frac{u(c_{t+1}^E)}{1-\beta}$, the continuation utility after a bad shock is not uniquely pinned down by $u(c_{t+1}^U)$.

Lemma 1 *Let σ be an I.C. contract. If $a_t > 0$, then $c_{t+1}^U < c_{t+1}^E$.*

Second we show that in any period where either full effort or no effort is implemented, the Euler equation holds with equality.

Lemma 2 *Let σ be an optimal contract. If $a_t \in \{0, 1\}$, then*

$$U_t = a_t \pi E_{t+1} + (1 - a_t \pi) U_{t+1}. \tag{3}$$

When $a_t = 0$, the result is simple: there is no reason not to let $U_t = U_{t+1} = E_{t+1}$, since there is no effort to induce. The more interesting case is that of $a_t = 1$. For an agent with no hidden wealth, the benefit of eliminating the borrowing constraint is analogous to the "inverse Euler equation" logic of Hopenhayn and Nicolini [4] and Rogerson [6]: by moving consumption from $t + 1$ to t (and, thereby, raising the marginal utility tomorrow relative to today), the incentive to search rises, because a greater marginal utility tomorrow makes finding a job more beneficial. In the CARA case, for an agent with no hidden savings, the algebra is especially simple: the marginal benefit of effort is proportional to $\beta\pi(-E_{t+1} + U_{t+1})$, so if the principal reduces consumption tomorrow uniformly by δ and raises it today by $\epsilon < \beta\delta$, leaving unchanged the utility for the agent who chooses the recommended effort ($a_t = 1$), the marginal benefit of effort increases by a factor of e^{δ}. Of course, in addition, the principal saves resources for the given level of delivered utility by reducing the intertemporal distortion.

Lemma 2 must also consider whether smoothing consumption for the agent with no hidden wealth would violate incentive constraint for agents with hidden savings. In that case, consider a plan of some level of period t hidden wealth s_{t-1}/β, an action in period t denoted $\tilde{a}_t$, and a level of wealth in period $t + 1$, s_t/β. The proof of the lemma shows that the proposed modification to the contract reduces the payoff to all combinations of s_{t-1}/β, $\tilde{a}_t$, and s_t/β.

To see the logic, consider first an agent with $s_t/\beta > 0$. Since this agent is already on the Euler equation, he must be made worse off by receiving less expected discounted resources. On the other hand, consider an agent with hidden savings at t who saves nothing for $t + 1$. This agent consumes the entire allocation c_t^U in period t, plus his wealth s_{t-1}/β; from period $t + 1$ onward, he is just like the agent who works the recommended amount. As a result, this agent gains less from the proposed change; his marginal utility in period t is lower than the agent who starts with no hidden wealth, and therefore he benefits less from moving consumption forward from $t+1$ to t.

Lemmas 1 and 2 imply that, when $a_t = 1$, the shirker is not borrowing constrained either. Since $U_t = \pi E_{t+1} + (1-\pi)U_{t+1}$ and $E_{t+1} < U_{t+1}$, shirking ($\tilde{a}_t < 1$) implies

$$U_t < \tilde{a}_t \pi E_{t+1} + (1 - \tilde{a}_t \pi) U_{t+1}.$$

Therefore, the shirker always secretly saves and chooses his consumption to restore the Euler equation.

Next, we introduce a result from Kocherlakota [5], that with linear disutility, interior effort can only be implemented when the agent is borrowing constrained:

Lemma 3 *(Kocherlakota [5]) Let σ be an I.C. contract. If $a_t \in (0,1)$, then $U_t > a_t \pi E_{t+1} + (1 - a_t \pi) U_{t+1}$.*

A key feature of the linear disutility model with interior effort is that, for any consumption sequence across states and dates, the agent is indifferent between all effort levels. Intuitively, linear disutility makes the model analogous to the case of effort being either zero or one, and $a_t \in (0,1)$ representing a mixed strategy. With the opportunity to also save, the agent has a further option: shirk and save. This option is attractive unless the contract makes saving unattractive by making the marginal utility high today. With observable savings, the agent's value function is linear in a_t, since the disutility function is linear. However, with hidden savings, when the agent is on the Euler equation, there is a second-order effect. Whenever the agent lowers a_t, he would save to smooth consumption, which makes the value function convex in a_t. This second-order effect cannot be made unprofitable when the agent is on the Euler equation.

A closely related feature in the linear disutility model is that implementing interior effort implies that, for the given consumption sequence, the principal is indifferent between implementing any level of effort a. So, starting from $a_t \in (0,1)$ and fixing consumption across states, the principal (and the agent) should be willing to replace the recommended a_t with $a_t = 1$. Since the agent is borrowing constrained at $a_t < 1$, he is further borrowing constrained at $a_t = 1$ (since his expected marginal utility of consumption tomorrow is lower when he searches harder). The *principal* then has a profitable sort of double deviation: move to $a_t = 1$ *and* adjust the consumption sequence to return the agent to the Euler equation. This must make the principal strictly better off by the line of argument in Lemma 2. We therefore have

Corollary 1 *It is never optimal for the principal to implement interior effort* $a_t \in (0, 1)$.

At this point we know that any optimal effort sequence uses only ones and zeros. However, it turns out that the principal would always prefer to implement a lottery over a sequence of all ones and all zeros rather than implement a sequence that contains both ones and zeros.

Lemma 4 *Any sequence of effort levels* $\{a_t\}_{t=0}^{\infty}$ *where* $a_t = 0$ *for some* t *and* $a_{t'} = 1$ *for some* t' *is dominated by a simple lottery with two outcomes: either* $\{1\}_{t=0}^{\infty}$, *or with the complementary probability,* $\{0\}_{t=0}^{\infty}$.

For instance, consider the sequence $\{a_t\}_{t=0}^{\infty} = (0, 1, 1, 1, ...)$. Alternately, the principal could make the agent search hard immediately, with probability β, and let the agent rest forever with probability $1 - \beta$. If savings were observed, the outcome would be no better or worse for both principal and agent. However, by choosing the lottery, the principal eliminates the possibility of hidden savings before search begins in period $t = 1$, which is possible under the sequence that starts with a period of zero effort. Eliminating this possibility eliminates a profitable deviation for the agent, and is therefore beneficial for the principal. More generally, the proof shows that periods where the agent is not searching are costly because of the hidden-savings possibilities they generate.

We are now ready to state the main result of this section, which gathers the previous results. It states that whenever the principal implements any effort, he implements high effort forever.

Proposition 1 *Suppose the principal has access to a one-time lottery before time* 0*. For any outcome of the lottery where* $a_t > 0$ *for any* t*, the principal implements* $a_t = 1$ *for all* $t \geq 0$*. Moreover,* $U_t = \pi E_{t+1} + (1-\pi)U_{t+1}$ *and shirkers are not borrowing constrained.*

In the next section we take up the issue of how to implement full effort forever. The Euler equation will be an aid in tackling that problem.

4 Implementing Full Effort Forever

We now focus on the problem of implementing full effort forever,

$$\min_{\{c_t^U\}_{t=0}^{\infty},\{c_t^E\}_{t=1}^{\infty}} \quad C(\sigma) \tag{4}$$

$$\begin{aligned} s.t. \quad & U_t = \pi E_{t+1} + (1-\pi)U_{t+1}, \\ & \bar{V} = V(\{1\}_{t=0}^{\infty}), \end{aligned}$$

$$V(\{1\}_{t=0}^{\infty}) \geq V(\{\tilde{a}_t\}_{t=0}^{\infty}), \text{ for all } \{\tilde{a}_t\}_{t=0}^{\infty}. \tag{5}$$

We solve problem (4) in two steps. First, we use the CARA assumption to solve the shirker's problem in a closed form; that is, for any sequence $\{\tilde{a}_t\}$, we compute the expected discounted utility $V(\{\tilde{a}_t\}_{t=0}^{\infty})$ for that plan after optimally undertaking hidden saving. Second, we conjecture that the binding incentive constraint for effort $a_t = 1$ is for agents who shirk up to $t-1$ ($\tilde{a}_s = 0$, for all $0 \leq s \leq t-1$), and search hard thereafter ($\tilde{a}_s = 1$, for all $s \geq t$). These agents are referred to as the *always-shirking-up-to-*$(t-1)$ types. We solve a relaxed problem where we impose only the incentive constraints for those deviations, and show that the solution satisfies Eq. (5) and is therefore a solution to the full problem (4).

4.1 Utility after deviation

Suppose an agent chooses a finite-deviation strategy $\{\tilde{a}_s\}_{s=0}^{\infty}$, and engages in hidden saving. With the CARA utility assumption, we can solve the hidden-savings problem in closed form:

Lemma 5 *Suppose an agent chooses a finite-deviation strategy* $\{\tilde{a}_s\}_{s=0}^{\infty}$, *where* $\tilde{a}_s = 1, \forall s \geq t$ *for some* t. *Then the agent's discounted utility from consumption with hidden savings is*

$$-\frac{U_0^{1-\beta}\left[\tilde{a}_0\pi E_1 + (1-\tilde{a}_0\pi)U_1^{1-\beta}\left[\tilde{a}_1\pi E_2 + (1-\tilde{a}_1\pi)U_2^{1-\beta}\left[...\left[\tilde{a}_{t-1}\pi E_t + (1-\tilde{a}_{t-1}\pi)U_t\right]^{\beta}...\right]^{\beta}\right]^{\beta}\right]^{\beta}}{1-\beta}. \quad (6)$$

Formula (6) is derived by backwards induction, using the agent's Euler equations and CARA preferences repeatedly. Euler equations provide a relationship between marginal utilities across times. However, under the assumption of CARA utility, this is also a relationship between levels of utilities from consumption, which we use to construct (6). Note that in general such an avenue is not available, since statements about marginal utility do not translate directly to level of utility; even for CRRA utility the t-period hidden-savings problem is not tractable.

4.2 Showing which constraints bind: a relaxed problem

We conjecture that the binding constraints are the ones for always-shirking-up-to-$(t-1)$ types. The intuition comes from the fact that shirking and saving are complements, so that shirking at time t is best combined with prior saving. Prior saving, in turn, is most valuable when prior shirking guaranteed that employment would not occur.

To show that these constraints are in fact the relevant ones, consider the relaxed problem

$$\min_{\sigma} \quad C(\sigma) \quad (7)$$

$$s.t \quad U_t = \pi E_{t+1} + (1-\pi)U_{t+1}, \text{ for all } t \geq 0,$$

$$\bar{V} = -\frac{U_0}{1-\beta} - \frac{1}{1-\beta(1-\pi)}, \quad (8)$$

$$\bar{V} \geq -\frac{\left(\Pi_{s=0}^{t-1}U_s^{(1-\beta)\beta^s}\right)U_t^{\beta^t}}{1-\beta} - \frac{\beta^t}{1-\beta(1-\pi)}, \text{ for all } t \geq 1. \quad (9)$$

Note that formula (6) is crucial for the relaxed problem. Without it, we are unable to write the incentive constraints (9) in closed form. In the appendix, we also use (6) to obtain a closed-form solution to (7). Without a closed-form solution to (7), we cannot establish the equivalence between the relaxed problem (7) and the full problem (4), and hence cannot solve the model.

First, we show in the relaxed problem that all the I.C. constraints bind.

Lemma 6 *In the optimal solution to problem (7), all the I.C. constraints bind.*

The relaxed problem has an equal number of choice variables and constraints. However, to understand why all the constraints bind, it is useful to explain the intuition that these constraints are independent. With change of variables $x_t = c^U_{t-1} - c^U_t = \log(U_t/U_{t-1})/\gamma$, the constraints in Eq. (9) can be rewritten as

$$\sum_{s=1}^{t} \beta^s x_s = \frac{1}{\gamma}\log\left(-\bar{V} - \frac{\beta^t}{1-\beta(1-\pi)}\right) - \frac{1}{\gamma}\log\left(-\bar{V} - \frac{1}{1-\beta(1-\pi)}\right).$$

Heuristically, we could pick a period $T = 10$ and divide the constraints into two groups, the "early" incentive constraints with $t \leq T$, and the "late" ones ($t > T$), and then pick one from each group. For example, we pick the incentive constraint for $t = 10$ from the first group and $t = \infty$ from the second group, and consider the constraints:

$$\sum_{s=1}^{10} \beta^s x_s \geq \frac{1}{\gamma}\log\left(-V - \frac{\beta^{10}}{1-\beta(1-\pi)}\right) - \frac{1}{\gamma}\log\left(-V - \frac{1}{1-\beta(1-\pi)}\right), \quad (10)$$

$$\sum_{s=1}^{\infty} \beta^s x_s \geq \frac{1}{\gamma}\log(-V) - \frac{1}{\gamma}\log\left(-V - \frac{1}{1-\beta(1-\pi)}\right). \quad (11)$$

If inequality (10) binds and inequality (11) is slack, then, as x_s ($s > 10$) does not appear in inequality (10), $x_s = 0$ in the solution, which means that consumption is flat in later periods: a contradiction to the fact that the principal wants to implement high effort in all time periods. If (11) binds and (10) is slack, then the principal knows that the shirking type will shirk in all time periods, and thus would reach the unemployment state U_t with probability one, while the equilibrium type would reach U_t with smaller and smaller probabilities as t gets bigger. Therefore, for efficiency reasons, the principal would like the distortion x_t to be increasing, and minimize the punishment in the early periods that the searcher hits relatively often.[3] This implies that

$$\sum_{s=1}^{10} \beta^s x_s < (1-\beta^{10}) \left[\frac{1}{\gamma} \log(-V) - \frac{1}{\gamma} \log\left(-V - \frac{1}{1-\beta(1-\pi)} \right) \right].$$

But the concavity of the function $\log(\cdot)$ implies that

$$\begin{aligned} & \log\left(-V - \frac{\beta^{10}}{1-\beta(1-\pi)} \right) - \log\left(-V - \frac{1}{1-\beta(1-\pi)} \right) \\ > \; & (1-\beta^{10}) \left[\log(-V) - \log\left(-V - \frac{1}{1-\beta(1-\pi)} \right) \right], \end{aligned}$$

contradicting (10). The proof extends this intuition to show that, if any constraint did not bind, the solution would imply an infeasible choice for that t.

Now we verify that the solution to the relaxed problem satisfies all the incentive constraints in the full problem.

Lemma 7 *If all the I.C. constraints bind in the relaxed problem* (7)*, then*

$$V(\{1\}_{s=0}^{\infty}) \geq V(\{\tilde{a}_s\}_{s=0}^{\infty}),$$

for all $\{\tilde{a}_s\}_{s=0}^{\infty}$ *where* $\tilde{a}_s = 1, \forall s \geq t+1$ *for some* t*. Therefore the solution to the relaxed problem* (7) *satisfies all the incentive constraints* (5)*, and is thus the solution to the full problem* (4)*.*

Instead of proving $V(\{1\}_{s=0}^{\infty}) \geq V(\{\tilde{a}_s\}_{s=0}^{\infty})$ directly, we show that the agent is (weakly) worse off if he decreases his effort in period t from 1 to $\tilde{a}_t$, i.e,

$$V(\{\tilde{a}_s\}_{s=0}^{t-1}, 1) \geq V(\{\tilde{a}_s\}_{s=0}^{t-1}, \tilde{a}_t).$$

This is sufficient to prove the lemma because we can then apply backward induction to roll the logic back to $s = 0$. For simplicity, we provide the intuition for the case of $\tilde{a}_t = 0$ only and leave the discussion of $\tilde{a}_t \in (0, 1)$ to the appendix.

[3] In fact, in the proof we show that the sequence of x_s is decreasing when all the constraints bind in the relaxed problem.

Recall that we refer to the agent with $a_s = 0$ for $s \leq t-1$ as the always-shirking-up-to-$(t-1)$ agent. The binding incentive constraints $V(\{0\}_{s=0}^{t-1}, 1) = V(\{0\}_{s=0}^{t-1}, 0)$ says that this agent is indifferent when he decreases his effort from 1 to 0 in period t. Using this fact, the lemma shows that the policy successfully induces effort in period t for any agent who may have worked at some point in the past.

The fundamental economics can be described in terms of wealth levels of different agents. The agent who has always shirked has had the greatest incentive to save, since he knows in every period that his future income is low, and therefore his marginal utility of consumption in the future is the highest. The always-shirking agent therefore has the greatest amassed wealth of workers of any effort history. As a result, the reduction in consumption that comes with a lower future entitlement (c_{t+1}^U instead of c_{t+1}^E) impacts the always-shirking agent the least, since his wealth allows him to smooth consumption. Agents who have worked in the past have less wealth, and therefore less ability to smooth consumption. If the always-shirking agent, who is well-prepared for low entitlements, finds it worth searching, then the "poorer" agents, who have the same disutility of search and a greater disutility from a low future entitlement, find it worthwhile to search as well. An implication of the lemma is that the binding incentive constraint at time t is for an agent who always shirks up to time $t-1$, and then works at t, and not for the agent who gives high search effort in every period.

In the next section we return, then, to the relaxed problem to develop some details of the optimal contract.

5 Analysis of the Optimal Contract

5.1 Decline of consumption over an unemployment spell

We can analyze how consumption changes over an unemployment spell by looking at the constraints in the relaxed problem. We have that

Proposition 2 *U_t/U_{t-1} is decreasing in t, and*

$$\lim_{t\to\infty}\frac{U_t}{U_{t-1}} = \lim_{t\to\infty}\exp(\gamma x_t) = \exp\left(\frac{(1-\beta)}{(-\bar{V})\beta(1-\beta(1-\pi))}\right) > 1. \tag{12}$$

Since U_t is proportional to the negative of utility from consumption, utility from consumption is falling at a decreasing rate, eventually reaching a constant rate of decline. In order to achieve this, consumption must fall by a constant *amount* in the long run.

A useful comparison is to the case of observable savings.

Proposition 3 *If savings are observed, as in Hopenhayn and Nicolini [4], then*

$$\lim_{t\to\infty}(U_t - U_{t-1}) = \frac{(1-\beta)^2}{\beta(1-\beta(1-\pi))}. \tag{13}$$

In the observable savings case, the long-run decline in *utility* from consumption is a constant amount, implying slower and slower declines in *consumption* in the long run. This shows how unobservable savings can make an important difference in the optimal policy. When savings are observed, consumption falls more and more slowly, because, in the long run, high marginal utility makes small consumption declines translate into large utility declines; but when consumption is unobserved, it cannot because the prospective shirker has more and more hidden wealth.

Both our result and the one from Hopenhayn and Nicolini [4] contrast with Kocherlakota [5], who finds the consumption of the unemployed to be constant. The reason is that, by focusing on interior effort levels and linear disutility, Kocherlakota must make agents indifferent between the implemented effort level, full effort, and no effort. This affords no flexibility for time varying consumption, since such a consumption path would always be exploitable by an agent who combined saving with either too much or too little effort. In our model, the principal only has to worry about agents giving too little effort, and therefore can make the consumption path decline in order to motivate the agent.

In order to further describe the impact of hidden savings, we parameterize the model. As in Hopenhayn and Nicolini [4], we view a period as representing a week and set $\beta = 0.999$ and the job-finding probability $\pi = 0.1$. Wage w is normalized to be 100. We pick $\gamma = 0.06$ so that the agent has a constant relative risk aversion of 4 on average. Two simplifications that were immaterial to our analytic results, but which are important in parameterization, need to be addressed. The first is that we have fixed the disutility of effort at one to save notation; we now parameterize that as d. Second, the model we introduced includes no work effort on the job. Clearly agents should consider working to require effort as well. In the interests of parsimony, we take this disutility to also be d. In our benchmark, we set $d = 0.0473$, so that an agent with a job is willing to sacrifice 50% of income to avoid the disutility from working.

At the outset we note that the model is less well suited to calibration than Hopenhayn and Nicolini [4]. The reasons are more than just the fact that we use CARA utility. Here disutility from effort cannot be calibrated in the manner of Hopenhayn and Nicolini [4]. With interior effort levels, there is a connection between the effort level and the job-finding probability, for which we have data. In our model, where effort is at a corner, we can choose the appropriate job-finding probability for the highest effort level, but we then essentially have d as a free parameter. The choice of d is clearly important. For small d, there is almost no costly effort to elicit. As a result, there is almost no required distortion in consumption without hidden savings, and agents are nearly on their Euler equation in that case. This in turn implies that hidden savings is not an important issue quantitatively; agents are not very savings constrained. On the other hand, for sufficiently high d, the distortion can be made very dramatic. We therefore will study other possible values of d below.

When the principal has a balanced budget, Figure 1 shows the unemployed worker's utility flows with and without hidden savings for the case with no possibility of initial wealth. The labels on the right scale are for the corresponding consumption levels c_t^U. Figure 1 confirms the above result that utilities decline faster with hidden savings; however, the two consumption paths are not significantly different until the duration of unemployment exceeds 6 years.

The long period of similar consumption paths described in Figure 1 would suggest that the effects of hidden savings on consumption distortion is not as large as that conjectured

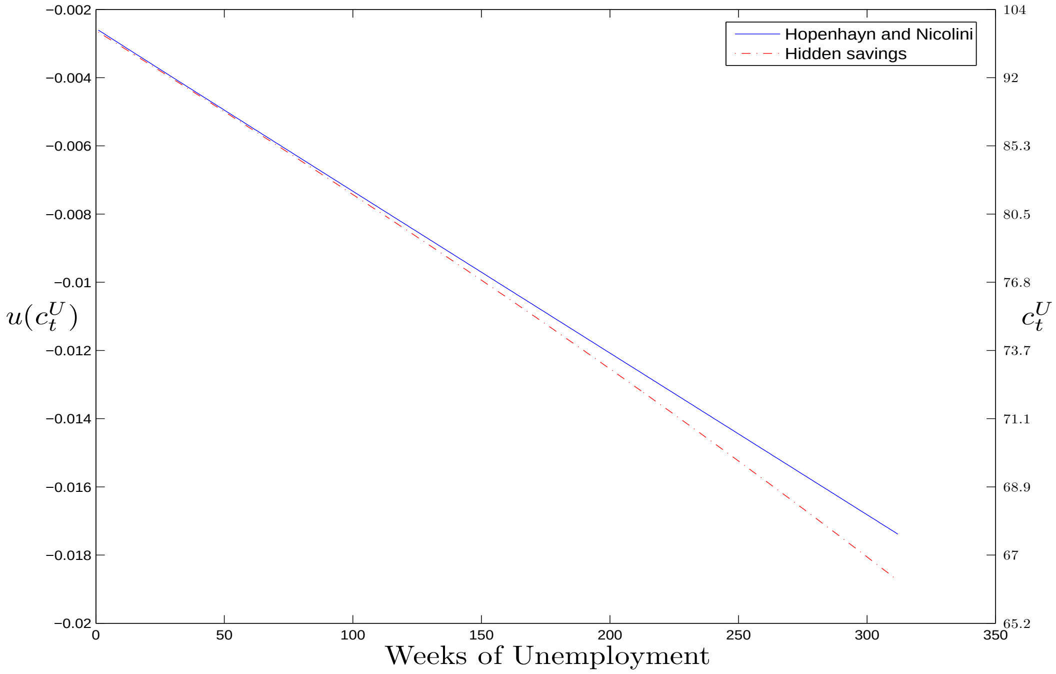

Figure 1: Utility flow of the unemployed.

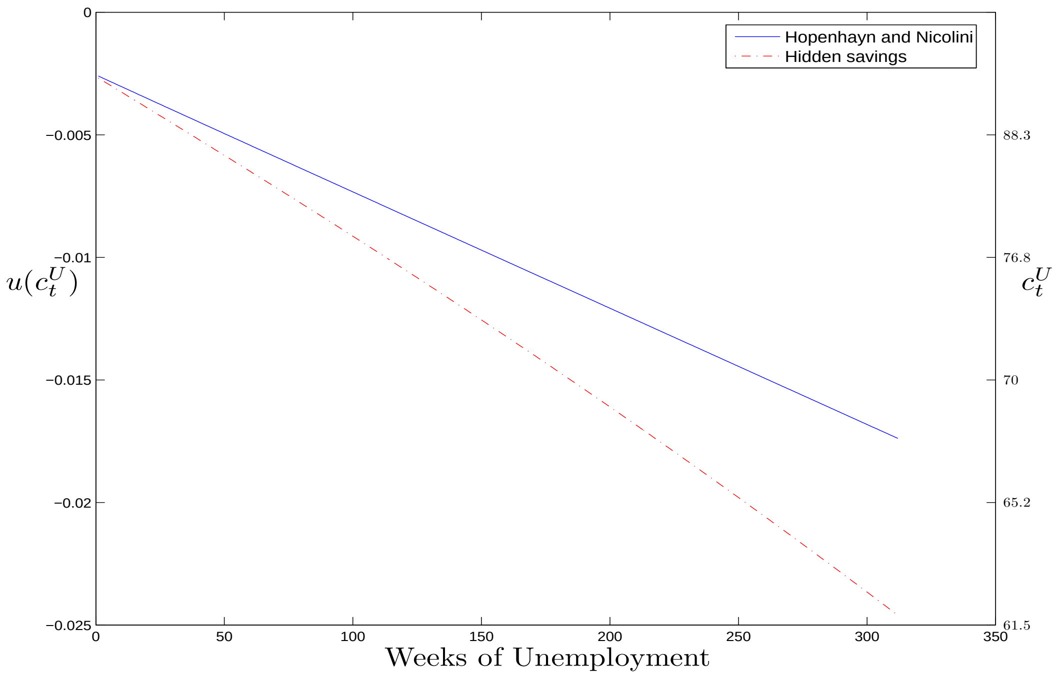

Figure 2: Utility flow of the unemployed with possibly hidden initial wealth.

by previous researchers. However, another reason that the model is not as well suited for calibration as Hopenhayn and Nicolini [4] is the treatment of initial wealth. In Hopenhayn and Nicolini [4], one can simply take initial wealth as immaterial (because it is observed), or rolled into initial promised utility (if one thinks that agents with more wealth are promised superior contracts). Here, however, if we follow the line that initial wealth is hidden (as we assume about subsequent wealth accumulation), then the initial wealth impacts crucially the agents' ability to subsequently shirk and utilize accumulated wealth. Wealth from sources other than saved unemployment benefits, which would include typical financial wealth as well as sources such as assistance from family, is certainly relevant for many people who find themselves unemployed and is hard to monitor.

To get a basic idea about the possible quantitative role of hidden initial wealth, we do the following. First, we assume that agents *may* have (hidden) initial wealth equivalent to 5% of discounted lifetime income. Then the principal provides a contract for each initial wealth level such that it is incentive compatible to report initial wealth truthfully. If we assume that agents can only underreport wealth, then the contract for the rich agent is similar to the contract analyzed before: since a report of positive wealth can only be made truthfully, the issue of initial wealth is moot following such a report and the contract can proceed as above.

The important decision is then how to design the contract for the poor agent. Making the contract unattractive to rich agents, while still eliciting effort, requires distortions. We do not solve the problem of optimally trading off those distortions against the information rents that the contract generates for the rich agent. Instead we compute the contract that elicits effort from the rich agent, should he misreport his wealth as zero. We focus on such a contract for two reasons. First, a contract that elicits effort from the rich agent will elicit effort from the poor agent as well. The agent with initial wealth should be the hardest to motivate, as in our results above, since his amassed wealth makes shirking more attractive. Second, among all contracts that deliver a given level of utility to the poor agent, this contract is the least attractive to the rich agent. Therefore offering such a contract to the poor agent dissuades rich agents as much as possible from claiming to be without initial wealth. In the extreme case where the poor agents are very rare, this would be the optimal policy, since making the contract unattractive to rich agents is more important than the distortions generated for the (rare) poor agents.

Figure 2 shows the consumption path of the unemployed poor agent, compared to the path in Hopenhayn and Nicolini [4]. Notice that the qualitative features of the paths are very similar to the model with no hidden initial wealth; only the quantitative values change as a result of hidden initial wealth. The difference comes from the fact that the path for the poor agent must be sufficiently distorted as to make it unattractive to rich agents.

Next we consider the welfare effects of hidden savings and hidden initial wealth when we vary d in the range of $[0, 0.0473]$. For each value of d, we first compute the promised utility, V_0, when the principal has balanced budget and there is no private information; in other words, when the utility can be delivered with completely smooth consumption. The cost of that policy is zero, so that the costs reported in Figure 3, where we allow for private search efforts and compute the principal's costs of delivering V_0 to a poor agent in environments with and without hidden savings, are measures of distortions in the consumption path. Figure 3 shows that the impact of hidden initial wealth can be dramatic; in particular, throughout the range of d hidden savings has roughly a doubling effect on the costly distortions. Note, as we would expect, the level of the distortion increases in d, since there is more and more costly effort to elicit as d rises.

Our numerical results suggest that studying hidden savings alone, as the literature has done, and as we have followed in the development of our optimal contract here, may be only the first step toward understanding the quantitative role of hidden wealth in the design of the optimal unemployment insurance contract. A full quantitative treatment would require understanding the role of hidden initial wealth in the problem. The qualitative features derived here, however, seem likely to survive the hidden initial wealth case. Certainly the insights from the contract without hidden initial wealth are essential to the study of the problem with hidden

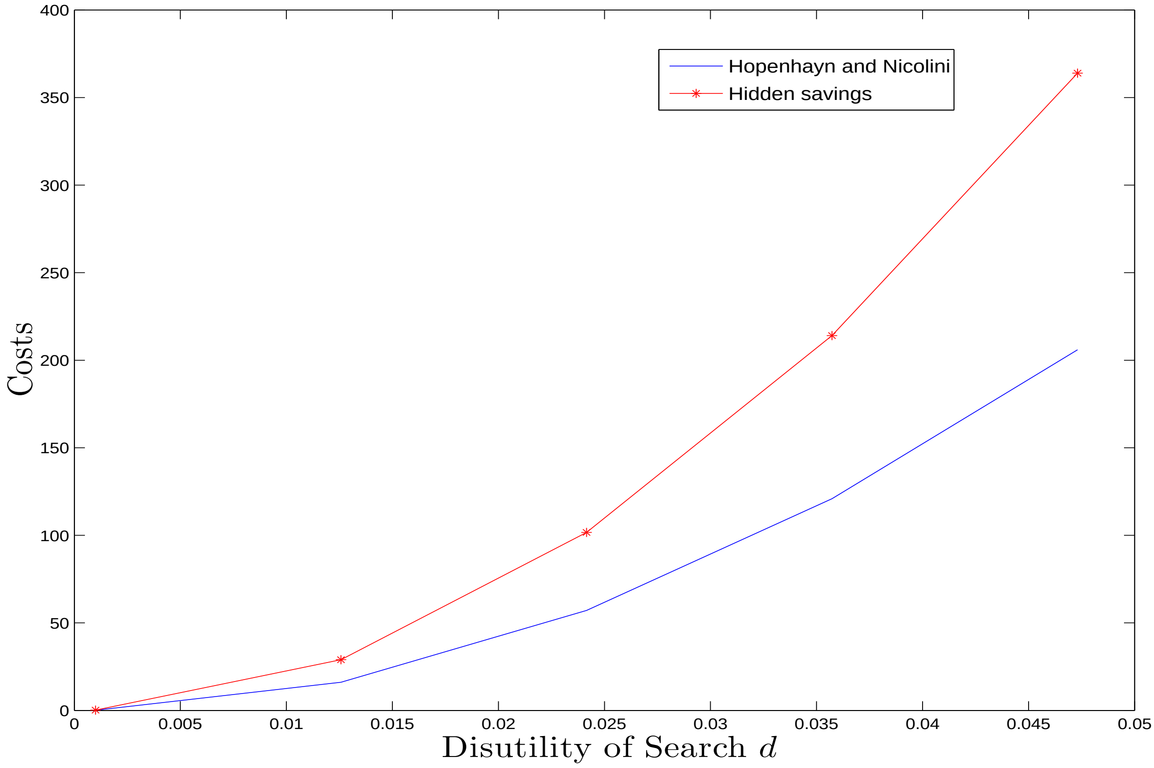

Figure 3: Costs of delivering a fixed level of utility.

initial wealth.

5.2 Rank Reversals in Consumption

An interesting difference between our model and Hopenhayn and Nicolini [4] is that the eventual decline in consumption depends on the initial promised utility when savings are unobserved. This difference gives rise to rank reversals in consumption levels over the spell for agents who start with different promised utilities. It is well known that in the case of observable savings, a larger promised $\bar{V}$ implies higher consumption at every node. That is not the case with unobserved savings. An agent with higher promised utility gets greater consumption at the start of the spell, but that extra consumption makes saving more tempting. In order to undo the temptation to shirk and save, the rate of consumption decline must be faster for the agent who consumes more initially; eventually his consumption is less than an agent with lower initial promised utility.

Proposition 4 *Consider two initial promised utility levels $\bar{V}^a > \bar{V}^b$. Then consumption is initially higher under $\bar{V}^a$ but is eventually lower forever.*

Despite this result, it is worth mentioning that, for large t, although the equilibrium type from $\bar{V}^a$ is poorer than that from $\bar{V}^b$, the always-shirking-up-to-$(t-1)$ type from $\bar{V}^a$ is always richer. In other words, our results are consistent with the monotonicity in Hopenhayn and Nicolini [4]: monotonicity is maintained for the wealth levels that are most likely to consider one deviation, and whose incentive constraints bind. In Hopenhayn and Nicolini [4], there is only one wealth level, that of zero wealth, and the agent is indifferent between shirking and giving effort at t. Here the critical wealth level belongs to the always-shirking-up-to-$(t-1)$ type, who is ex ante indifferent between his strategy and one more time period of shirking. Notice that, since an always-shirking-up-to-$(t-1)$ type from $\bar{V}^a$ always obtains an ex ante utility equivalent to the equilibrium type, he has secretly saved so much wealth so that even if he faces lower future consumption claims, he is able to live better than an always-shirking-up-to-$(t-1)$ type from $\bar{V}^b$.

5.3 Increasing unemployment benefits over the spell

Although the optimal consumption path is uniquely determined, Ricardian equivalence implies that many transfer sequences, combined with asset markets, can achieve the optimum. We pin down the benefit sequence (and the agent's initial asset holdings) by having an employment tax of zero. Then benefits B_t can be obtained for all t from the budget constraint in the optimal savings problem of the agent:

$$c_t^U + \frac{\beta c_{t+1}^E}{1-\beta} - B_t - \frac{\beta w}{1-\beta} = \frac{c_t^E}{1-\beta} - \frac{w}{1-\beta},$$

which implies that

$$B_t = c_t^U + \frac{w - c_t^E}{1-\beta} - \beta \frac{w - c_{t+1}^E}{1-\beta}.$$

Analogous to Shimer and Werning [8], B_t is the insurance for the unemployment risk that an agent faces at period t; $w - B_t$ is the difference in the entitlements after finding a job and not finding a job, and hence can be interpreted as the penalty for shirking. In our model, the benefit B_t is strictly increasing with the duration of unemployment.

Proposition 5 *B_t is strictly increasing in t.*

It is well known that in Shimer and Werning [8], B_t is a constant. In their model, because the effect of $w - B_t$ on the reservation wage is independent of wealth, poor agents and rich agents have the same incentives to accept an offer. Hence, the principal chooses a constant penalty $w - B_t$ regardless of the agent's history or wealth level. In our model, the always-shirking-up-to-$(t-1)$ type is richer than the always-shirking-up-to-t type. In period t, $w - B_t$ motivates the always-shirking-up-to-$(t-1)$ type; while in period $t+1$, $w - B_{t+1}$ motivates the always-shirking-up-to-t type. Since it is more difficult to induce effort from richer agents in our model, $w - B_t$ needs to be larger than $w - B_{t+1}$. In other words, B_t is below B_{t+1}.

Intuitively, the rising benefit sequence is related to the similar one that Shavell and Weiss [7] find for the case where job-finding probabilities are exogenous but borrowing and lending are allowed. There, the single force is consumption smoothing: higher marginal utilities in later periods of the unemployment spell dictate higher shifting benefits to those states. In their model the outcome is a very extreme jump from zero benefits to full benefits (equal to the wage) at some two-period interval. Here, however, due to moral hazard, a wedge between consumption when employed and unemployed must always be maintained.

6 Conclusion

Our paper is a first step toward solving mechanism design problems with hidden savings. We have shown a case under which hidden savings (and, without loss of generality, borrowing) leads to faster utility declines over an unemployment spell than for the contract in Hopenhayn and Nicolini [4]. This result contrasts with the hidden-savings model of Kocherlakota [5], where the rate of decline of utility over a spell of unemployment is zero. Further, our contract has the feature that an agent with an initially higher consumption claim, due to higher promised utility, may eventually have a lower claim than the agent with initially low consumption. This non-monotonicity comes from the nature of the structure of the binding incentive constraints.

Our mode of attack relies on determining the binding incentive constraints. We use the notion that types who shirk also want to save. That motive is central to Kocherlakota's argument about why the first-order approach to these problems might fail. We exploit this feature to prove that the type that shirks the most (and saves the most) will be the hardest agent from whom to extract effort. Therefore, that type has the binding incentive constraint. Whereas the particular results depend on our assumptions on the functional forms, we believe that the results point to what would happen in general when non-implemented types have the binding constraints. The double-deviation incentives make such a structure of binding constraints natural and lead to the results that differ from those in the literature. In particular, the principal must block the possibility of shirking and saving by making a more dramatic decline in the rate of consumption. This problem is exacerbated for agents who get high initial claims and, therefore, would be most prone to saving.

The results here demonstrate that the richest type plays a role in determining the optimal contract, because the richest type's incentive constraint binds. If this type of problems could be studied recursively, an artificial type would need to be added to the state space in addition to the continuation utility of the equilibrium type. Further development of this idea is the task of future research.

7 Appendix

Proof of Lemma 1: Let F_{t+1} denote the discounted (to period $t+1$) utility of consumption conditional on not finding a job at the beginning of $t+1$,

$$F_{t+1} = \sum_{s=t+1}^{\infty} \beta^{s-(t+1)} \left(\Pi_{k=t+1}^{s-1}(1-a_k\pi)\right) \left[-U_s + \beta a_s \pi \frac{-E_{s+1}}{1-\beta}\right].$$

Recall that D_{t+1}/β^{t+1} is the discounted (to period $t+1$) disutility conditional on not finding a job at the beginning of $t+1$. Eq. (1) implies that $F_{t+1} \geq -U_{t+1}/(1-\beta)$. The I.C. condition requires that an agent's benefit of using effort outweighs the cost,

$$\begin{aligned} 1 &\leq \beta\pi\left(\frac{-E_{t+1}}{1-\beta} - \left(F_{t+1} - \frac{D_{t+1}}{\beta^{t+1}}\right)\right) \\ &\leq \beta\pi\left(\frac{-E_{t+1}}{1-\beta} - \left(\frac{-U_{t+1}}{1-\beta} - \frac{D_{t+1}}{\beta^{t+1}}\right)\right) \\ &\leq \beta\pi\frac{U_{t+1}-E_{t+1}}{1-\beta} + \frac{\beta\pi}{1-\beta+\beta\pi} \\ &< \beta\pi\frac{U_{t+1}-E_{t+1}}{1-\beta} + 1, \end{aligned}$$

where the third inequality follows from that $D_{t+1}/\beta^{t+1} \leq 1/(1-\beta+\beta\pi)$. Thus $U_{t+1} > E_{t+1}$. *Q.E.D.*

Proof of Lemma 2: Suppose to the contrary that

$$U_t > \pi E_{t+1} + (1-\pi)U_{t+1}. \tag{14}$$

In the following, we use a variational argument to show that the principal could modify σ to save resources without violating any incentive constraints, which contradicts the optimality of σ. Step (i) describes the variation. Step (ii) shows that the variation reduces the cost for the principal. Steps (iii) and (iv) imply that the contract remains I.C. after the variation. Specifically, step (iii) shows that the equilibrium type is indifferent, while step (iv) shows that all the deviators are (weakly) worse off after the variation.

(i) For $\epsilon > 0$, there is a unique $\delta(\epsilon) > 0$, such that

$$\begin{aligned} & u(c_t^U + \epsilon) + \beta \left[\pi u(c_{t+1}^E - \delta(\epsilon)) + (1-\pi) u(c_{t+1}^U - \delta(\epsilon))\right] \\ = \; & -U_t + \beta \left[\pi(-E_{t+1}) + (1-\pi)(-U_{t+1})\right]. \end{aligned} \tag{15}$$

Thus $\delta'(\epsilon) = u'(c_t^U + \epsilon)/\beta[\pi u'(c_{t+1}^E - \delta(\epsilon)) + (1-\pi)u'(c_{t+1}^U - \delta(\epsilon))]$. Eq. (14) implies that there is a small $\bar{\epsilon} > 0$, such that $\delta'(\epsilon) > 1/\beta$, for all $\epsilon \in [0, \bar{\epsilon}]$. Modify the contract by choosing $\tilde{c}_t^U = c_t^U + \bar{\epsilon}$, $\tilde{c}_{t+1}^E = c_{t+1}^E - \delta(\bar{\epsilon})$, $\tilde{c}_{t+1}^U = c_{t+1}^U - \delta(\bar{\epsilon})$.

(ii) The principal saves resources, because $\bar{\epsilon} < \beta\delta(\bar{\epsilon})$.

(iii) The equilibrium type would not secretly save, and his promised utility is unchanged, given the definition of $\delta(\epsilon)$.

(iv) All deviators are (weakly) worse off. At the beginning of t, consider a deviator who starts with savings $s_t \geq 0$, and exerts effort $\{\tilde{a}_s\}_{s=t}^{\infty}$. Let F_1 and F_2 be the deviator's utility from consumption before and after the modification, respectively.

$$\begin{aligned} F_1 &= \max_{s_t \geq 0} \left\{u(s_t + c_t^U - \beta s_{t+1}) + \beta[\tilde{a}_t \pi v^E(s_{t+1} + c_{t+1}^E) + (1 - \tilde{a}_t\pi) v^U(s_{t+1} + c_{t+1}^U)]\right\}, \\ F_2 &= \max_{s_t \geq 0} \{u(s_{t-1} + c_t^U + \bar{\epsilon} - \beta s_{t+1}) + \beta[\tilde{a}_t \pi v^E(s_{t+1} + c_{t+1}^E - \delta(\bar{\epsilon})) \\ & \quad + (1 - \tilde{a}_t \pi) v^U(s_{t+1} + c_{t+1}^U - \delta(\bar{\epsilon}))]\}, \end{aligned}$$

where $v^E(\cdot), v^U(\cdot)$ are the deviator's value functions starting from period $t+1$. Then

$$\begin{aligned} F_2 = \; & F_1 + \int_0^{\bar{\epsilon}} \{u'(s_t + c_t^U + \epsilon - \beta s_{t+1}(\epsilon)) - \beta[\tilde{a}_t \pi (v^E)'(s_{t+1}(\epsilon) + c_{t+1}^E - \delta(\epsilon)) \\ & + (1 - \tilde{a}_t\pi)(v^U)'(s_{t+1}(\epsilon) + c_{t+1}^U - \delta(\epsilon))]\delta'(\epsilon)\} d\epsilon, \end{aligned} \tag{16}$$

where $s_{t+1}(\epsilon)$ is the optimal savings in the maximization problem. If $s_{t+1}(\epsilon) > 0$, then

$$\begin{aligned}
& u'(s_t + c_t^U + \epsilon - \beta s_{t+1}(\epsilon)) \\
= \ & [\tilde{a}_t \pi (v^E)'(s_{t+1}(\epsilon) + c_{t+1}^E - \delta(\epsilon)) + (1 - \tilde{a}_t \pi)(v^U)'(s_{t+1}(\epsilon) + c_{t+1}^U - \delta(\epsilon))], \text{ thus} \\
& u'(s_t + c_t^U + \epsilon - \beta s_{t+1}(\epsilon)) - \beta[\tilde{a}_t \pi (v^E)'(s_{t+1}(\epsilon) + c_{t+1}^E - \delta(\epsilon)) \\
& +(1 - \tilde{a}_t \pi)(v^U)'(s_{t+1}(\epsilon) + c_{t+1}^U - \delta(\epsilon))]\delta'(\epsilon) \\
= \ & u'(s_t + c_t^U + \epsilon - \beta s_{t+1}(\epsilon))(1 - \beta\delta'(\epsilon)) \\
< \ & 0.
\end{aligned}$$

If $s_{t+1}(\epsilon) = 0$, then

$$\begin{aligned}
& u'(s_t + c_t^U + \epsilon) - \beta[\tilde{a}_t \pi (v^E)'(c_{t+1}^E - \delta(\epsilon)) + (1 - \tilde{a}_t \pi)(v^U)'(c_{t+1}^U - \delta(\epsilon))]\delta'(\epsilon) \\
\leq \ & u'(s_t + c_t^U + \epsilon) - \beta[\tilde{a}_t \pi u'(c_{t+1}^E - \delta(\epsilon)) + (1 - \tilde{a}_t \pi) u'(c_{t+1}^U - \delta(\epsilon))]\delta'(\epsilon) \\
\leq \ & u'(s_t + c_t^U + \epsilon) - \beta[\pi u'(c_{t+1}^E - \delta(\epsilon)) + (1 - \pi) u'(c_{t+1}^U - \delta(\epsilon))]\delta'(\epsilon) \\
\leq \ & u'(c_t^U + \epsilon) - \beta[\pi u'(c_{t+1}^E - \delta(\epsilon)) + (1 - \pi) u'(c_{t+1}^U - \delta(\epsilon))]\delta'(\epsilon) \\
= \ & 0,
\end{aligned}$$

where the first inequality follows from the deviator's non-negative savings (and thus having higher marginal utility of consumption than the equilibrium type) at $t+1$, the second inequality follows from $\tilde{a}_t \leq 1$ and $c_{t+1}^E > c_{t+1}^U$, as shown in lemma 1, and the third inequality follows from $s_t \geq 0$. Therefore, Eq. (16) implies $F_2 \leq F_1$.

Using a similar argument, we can show that if $a_t = 0$, then $U_t = U_{t+1}$. *Q.E.D.*

Proof of Corollary 1: Suppose $\sigma = (\{c_s^U\}_{s=0}^\infty, \{c_s^E\}_{s=1}^\infty, \{a_s\}_{s=0}^\infty)$ is an I.C. contract in which $a_t \in (0,1)$ for some t. Consider a family of contracts $\sigma(x) = (\{c_s^U\}_{s=0}^\infty, \{c_s^E\}_{s=1}^\infty, \{a_0, ..., a_{t-1}, x, a_{t+1}, ...\})$ where the effort recommendation a_t in period t is replaced by $x \in [0,1]$. Since $\sigma = \sigma(a_t)$ is I.C. and the agent's utility is linear in x, the agent must be indifferent between any $x \in [0,1]$ (i.e., $V(\{a_0, ..., a_{t-1}, x, a_{t+1}, ...\}) = V(\{a_0, ..., a_{t-1}, a_t, a_{t+1}, ...\})$ for any x). Therefore $\sigma(x)$ is I.C. for any $x \in [0,1]$.

Because $\sigma(x)$ is I.C. for any $x \in [0,1]$ and the principal's cost function is also linear in x, the principal must be indifferent between implementing $\sigma(a_t)$ and $\sigma(1)$. Because $U_t \geq U_{t+1} > E_{t+1}$, it follows from Lemma 2 and $U_t > \pi E_{t+1} + (1-\pi)U_{t+1}$ that $\sigma(1)$ is not optimal. Thus $\sigma = \sigma(a_t)$ cannot be optimal either. *Q.E.D.*

Before we prove **Lemma 4**, we first describe a relaxed problem (19) in the following. We characterize the solution to this problem in **Lemmas A.1-A.5**. These results will be used later.

Let $\{a_t\}_{t=0}^{\infty}$, $a_t \in \{0,1\}$, be an effort sequence that the principal wants to implement. Let n_t ($t \geq 1$) denote the total number of time periods when high effort is recommended before t, i.e., $n_t = \#\{s : a_s = 1, 0 \leq s \leq t-1\}$. Define $x_t = c_{t-1}^U - c_t^U = \log(U_t/U_{t-1})/\gamma$. Then $c_t^U = c_0^U - \sum_{s=1}^t x_s$. If $a_t = 1$, then Eq. (3) yields

$$\begin{aligned} c_{t+1}^E &= c_t^U - \frac{1}{\gamma}\log\left(\frac{1-(1-\pi)\exp(\gamma x_{t+1})}{\pi}\right) \\ &= c_0^U - \sum_{s=1}^t x_s - \frac{1}{\gamma}\log\left(\frac{1-(1-\pi)\exp(\gamma x_{t+1})}{\pi}\right); \end{aligned}$$

while if $a_t = 0$, Eq. (3) implies that $x_{t+1} = 0$, and c_{t+1}^E does not need to be specified. The cost function $C(\sigma)$ can be decomposed as

$$\begin{aligned} C(\sigma) &= \sum_{t=0}^{\infty} \beta^t \left(\Pi_{s=0}^{t-1}(1-a_s\pi)\right)\left[c_t^U + \beta a_t \pi \frac{c_{t+1}^E}{1-\beta}\right] - \frac{\beta\pi w}{1-\beta}D_0 \\ &= \sum_{t=0}^{\infty} \beta^t \left(\Pi_{s=0}^{t-1}(1-a_s\pi)\right)\left[c_0^U - \sum_{s=1}^t x_s + \beta a_t \pi \frac{c_0^U - \sum_{s=1}^t x_s - \frac{1}{\gamma}\log\left(\frac{1-(1-\pi)\exp(\gamma x_{t+1})}{\pi}\right)}{1-\beta}\right] \\ &\quad - \frac{\beta\pi w}{1-\beta}D_0 \\ &= \frac{c_0^U}{1-\beta} - \sum_{t=1}^{\infty} \frac{\beta^t}{1-\beta}(1-\pi)^{n_t}\left[(1-a_{t-1}\pi)x_t + a_{t-1}\pi\frac{1}{\gamma}\log\left(\frac{1-(1-\pi)\exp(\gamma x_t)}{\pi}\right)\right] \\ &\quad - \frac{\beta\pi w}{1-\beta}D_0 \\ &= \frac{c_0^U}{1-\beta} - \sum_{t=1}^{\infty} \frac{\beta^t}{1-\beta}(1-\pi)^{n_t}\chi_{\{a_{t-1}=1\}}\left[(1-\pi)x_t + \pi\frac{1}{\gamma}\log\left(\frac{1-(1-\pi)\exp(\gamma x_t)}{\pi}\right)\right] \end{aligned}$$

$$-\frac{\beta\pi w}{1-\beta}D_0,$$

where χ is the indicator function. Define a function $f(x) = -(1-\pi)x - \pi\frac{1}{\gamma}\log\left(\frac{1-(1-\pi)\exp(\gamma x)}{\pi}\right)$. Since $f(0) = 0$, $f'(x) = -1 + \frac{\pi}{1-(1-\pi)\exp(\gamma x)}$, $f'(0) = 0$, $f''(x) = \pi(1-\pi)\gamma(1-(1-\pi)\exp(\gamma x))^{-2}\exp(\gamma x) > 0$, we see that f is strictly convex and has a unique minimizer at $x = 0$. Thus

$$C(\sigma) = \frac{c_0^U}{1-\beta} + \sum_{t=1}^{\infty}\frac{\beta^t}{1-\beta}(1-\pi)^{n_t}\chi_{\{a_{t-1}=1\}}f(x_t) - \frac{\beta\pi w}{1-\beta}D_0. \quad (17)$$

Next consider a shirker who shirks in all periods up to a given t ($\tilde{a}_s = 0$, for all $0 \le s \le t-1$), and follows recommended efforts from then on ($\tilde{a}_s = a_s$, for all $s \ge t$). Using the same proof as in Lemma 5, we obtain

$$V(\{\tilde{a}_s\}_{s=0}^{\infty}) = -\frac{\left(\Pi_{s=0}^{t-1}U_s^{(1-\beta)\beta^s}\right)U_t^{\beta^t}}{1-\beta} - D_t.$$

We study a relaxed problem where only shirkers in the above are considered,

$$\begin{aligned}
\min_{\sigma} \quad & C(\sigma) \\
s.t. \quad & U_t = \pi E_{t+1} + (1-\pi)U_{t+1}, \text{ if } a_t = 1, \\
& U_t = U_{t+1}, \text{ if } a_t = 0, \\
& \bar{V} = -\frac{U_0}{1-\beta} - D_0, \\
& \bar{V} \ge -\frac{\left(\Pi_{s=0}^{t-1}U_s^{(1-\beta)\beta^s}\right)U_t^{\beta^t}}{1-\beta} - D_t, \text{ for all } t \ge 1.
\end{aligned}$$

Denote the optimal solution to the above by $\sigma^*(\{a_t\}_{t=0}^{\infty})$. The incentive constraint $-\frac{U_0}{1-\beta} - D_0 \ge -\frac{\left(\Pi_{s=0}^{t-1}U_s^{(1-\beta)\beta^s}\right)U_t^{\beta^t}}{1-\beta} - D_t$ is equivalent to

$$\sum_{s=1}^{t}\beta^s x_s \ge \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + D_0 - D_t\right) - \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}\right). \quad (18)$$

Using Eqs. (17) and (18), the relaxed problem is rewritten as

$$\min_{\{x_t\}_{t=1}^{\infty}} \quad \sum_{t=1}^{\infty} \frac{\beta^t}{1-\beta}(1-\pi)^{n_t}\chi_{\{a_{t-1}=1\}}f(x_t) \tag{19}$$

$$s.t. \quad \sum_{s=1}^{t} \beta^s x_s \geq \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}+D_0-D_t\right) - \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}\right), \text{ for all } t \geq 1, \tag{20}$$

$$x_t = 0, \text{ if } a_{t-1} = 1.$$

Lemma A.1 *For $s < t$, if $a_s = 1$ and $a_t = 1$ ($D_s > D_{s+1} \geq D_t > D_{t+1}$), and $\frac{D_s - D_{s+1}}{\beta^{s+1}} \geq \frac{D_t - D_{t+1}}{\beta^{t+1}}$, then*

$$\frac{\frac{1}{\gamma}\log(\frac{U_0}{1-\beta}+D_0-D_{s+1}) - \frac{1}{\gamma}\log(\frac{U_0}{1-\beta}+D_0-D_s)}{\beta^{s+1}} > \frac{\frac{1}{\gamma}\log(\frac{U_0}{1-\beta}+D_0-D_{t+1}) - \frac{1}{\gamma}\log(\frac{U_0}{1-\beta}+D_0-D_t)}{\beta^{t+1}}.$$

Proof of Lemma A.1: This follows from the strict concavity of function $\log(\cdot)$. *Q.E.D.*

Lemma A.2 *If there is an $s \geq 0$, such that $a_t = 1$, for all $t \geq s$, then $\sigma^*(\{a_t\}_{t=0}^{\infty})$ satisfies*

$$\sum_{i=1}^{s} \beta^i x_i = \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}+D_0-D_s\right) - \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}\right).$$

Proof of Lemma A.2: In the optimal solution, $x_t \geq 0$, for all $t \geq 1$. We also know that

$$\sum_{t=1}^{\infty} \beta^t x_t = \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}+D_0\right) - \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}\right). \tag{21}$$

By contradiction, suppose

$$\sum_{i=1}^{s} \beta^i x_i > \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}+D_0-D_s\right) - \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}\right). \tag{22}$$

Then we will show that

$$\beta^{s+1}x_{s+1} > \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}+D_0-D_{s+1}\right) - \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}+D_0-D_s\right). \tag{23}$$

To see this, let s^* $(1 \leq s^* \leq s)$ be the time index such that

$$\begin{aligned} \sum_{i=1}^{s^*-1} \beta^i x_i &= \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + D_0 - D_{s^*-1}\right) - \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}\right), \\ \sum_{i=1}^{t} \beta^i x_i &> \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + D_0 - D_t\right) - \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}\right), s^* \leq t \leq s. \end{aligned} \tag{24}$$

Therefore, $\beta^{s^*} x_{s^*} > \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + D_0 - D_{s^*}\right) - \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + D_0 - D_{s^*-1}\right)$. For $\epsilon \neq 0$, changing x_{s^*} by ϵ/β^{s^*} and x_{s+1} by $-\epsilon/\beta^{s+1}$ will keep $\beta^{s^*} x_{s^*} + \beta^{s+1} x_{s+1}$ unchanged. Eqs. in (24) imply that no constraints in (20) will be violated after the change if ϵ is sufficiently small. Thus the first-order condition for a minimum at $\epsilon = 0$ is

$$(1-\pi)^{n_{s^*}} f'(x_{s^*}) + (1-\pi)^{n_{s+1}} f'(x_{s+1}) = 0,$$

which, with the convexity of f, implies that $x_{s+1} > x_{s^*}$. Thus

$$\begin{aligned} x_{s+1} &> x_{s^*} \\ &> \frac{\frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + D_0 - D_{s^*}\right) - \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + D_0 - D_{s^*-1}\right)}{\beta^{s^*}} \\ &> \frac{\frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + D_0 - D_{s+1}\right) - \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + D_0 - D_s\right)}{\beta^{s+1}}, \end{aligned}$$

where the last inequality follows from $(D_{s^*-1} - D_{s^*})/\beta^{s^*} = (\beta^{s^*-1} + (1-\pi)D_{s^*} - D_{s^*})/\beta^{s^*} \geq (1/\beta - 1)/(1 - \beta(1-\pi)) = (D_s - D_{s+1})/\beta^{s+1}$, and Lemma A.1. That is, Eq. (23) is proved. Therefore Eqs. (22) and (23) imply

$$\sum_{i=1}^{s+1} \beta^i x_i > \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + D_0 - D_{s+1}\right) - \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}\right).$$

Thus, using the same proof we used for Eq. (23) and by induction, we obtain

$$\beta^{t+1} x_{t+1} > \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + D_0 - D_{t+1}\right) - \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + D_0 - D_t\right), \text{ for all } t \geq s,$$

which contradicts (21). *Q.E.D.*

Lemma A.3 *If there is an $s \geq 0$, such that $a_t = 1$, for all $t \geq s$, then $\sigma^*(\{a_t\}_{t=0}^{\infty})$ satisfies, for all $t \geq s+1$,*

$$\beta^t x_t = \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + D_0 - D_t\right) - \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + D_0 - D_{t-1}\right). \tag{25}$$

Thus x_t is strictly decreasing in t when $t \geq s+1$.

Proof of Lemma A.3: For any $s' \geq s$, the conditions in Lemma A.1 ($a_t = 1$, for all $t \geq s'$) are satisfied, which implies

$$\sum_{i=1}^{s'} \beta^i x_i = \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + D_0 - D_{s'}\right) - \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}\right), \text{ for all } s' \geq s,$$

which implies (25). When $s \geq t+1$, $(D_{t-1} - D_t)/\beta^t = (\beta^{t-1}/(1-\beta(1-\pi)) - \beta^t/(1-\beta(1-\pi)))/\beta^t = (1/\beta - 1)/(1-\beta(1-\pi))$, thus it follows from Eq. (25) and Lemma A.1 that x_t is strictly decreasing in t, when $t \geq s+1$. *Q.E.D.*

Lemma A.4 *If there is an $s \geq 0$, such that $a_t = 1$, for all $t \geq s$, then $\sigma^*(\{a_t\}_{t=0}^{\infty})$ satisfies*

$$\frac{\sum_{i=1}^{s} \beta^i (1-\pi)^{n_i} \chi_{\{a_{i-1}=1\}} f(x_i)}{\sum_{i=1}^{s} \beta^i (1-\pi)^{n_i} \chi_{\{a_{i-1}=1\}}} > f(x_{s+1}), \text{ when } \sum_{i=1}^{s} \beta^i (1-\pi)^{n_i} \chi_{\{a_{i-1}=1\}} > 0.$$

Proof of Lemma A.4: Define $\tilde{x}_t = (\log(U_0/(1-\beta) + D_0 - D_t) - \log(U_0/(1-\beta) + D_0 - D_{t-1}))/(\gamma\beta^t)$, for $t \leq s$. If $a_{t-1} = 0$, then $\tilde{x}_t = 0$; otherwise, Lemma A.1 implies $\tilde{x}_t > x_{s+1}$, for $t \leq s$. Eq. (20) implies

$$\sum_{i=1}^{t} \beta^i x_i \geq \sum_{i=1}^{t} \beta^i \tilde{x}_i, \text{ for all } t \leq s,$$

which yields

$$\sum_{i=1}^{s} \beta^i (1-\pi)^{n_i} \chi_{\{a_{i-1}=1\}} x_i \geq \sum_{i=1}^{s} \beta^i (1-\pi)^{n_i} \chi_{\{a_{i-1}=1\}} \tilde{x}_i > \left(\sum_{i=1}^{s} \beta^i (1-\pi)^{n_i} \chi_{\{a_{i-1}=1\}}\right) x_{s+1}.$$

Therefore, the convexity of f and $f'(x)$ being positive (when $x > 0$) imply

$$\frac{\sum_{i=1}^{s} \beta^i (1-\pi)^{n_i} \chi_{\{a_{i-1}=1\}} f(x_i)}{\sum_{i=1}^{s} \beta^i (1-\pi)^{n_i} \chi_{\{a_{i-1}=1\}}} \geq f\left(\frac{\sum_{i=1}^{s} \beta^i (1-\pi)^{n_i} \chi_{\{a_{i-1}=1\}} x_i}{\sum_{i=1}^{s} \beta^i (1-\pi)^{n_i} \chi_{\{a_{i-1}=1\}}}\right) > f(x_{s+1}).$$

Q.E.D.

Lemma A.5 *If there is an $s \geq 1$, such that $a_{s-1} = 0$ and $a_t = 1$, for all $t \geq s$, then $\sigma^*(\{a_t\}_{t=0}^{\infty})$ is dominated by a lottery with two outcomes: either $\sigma^*(\{1\}_{t=0}^{\infty})$, or with the complementary probability, $\sigma^*(\{0\}_{t=0}^{\infty})$. Notice that $\sigma^*(\{0\}_{t=0}^{\infty})$ is trivially I.C. and by Lemma 6 and Lemma 7, $\sigma^*(\{1\}_{t=0}^{\infty})$ satisfies Eq.* (5) *as well.*

Proof of Lemma A.5: Lemma A.2 and A.3 yield that

$$\begin{aligned}\sum_{i=1}^{s-1}\beta^i x_i &= \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}+D_0-D_{s-1}\right)-\frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}\right),\\ x_s &= 0,\\ \beta^t x_t &= \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}+D_0-D_t\right)-\frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}+D_0-D_{t-1}\right),\ \text{for all } t\geq s+1.\end{aligned}$$

Now construct a new contract $\tilde{\sigma}$ to implement effort sequence $\{\tilde{a}_t\}_{t=0}^{\infty}$, where

$$\begin{aligned}\tilde{a}_t &= a_t,\ \text{if } t < s-1,\\ \tilde{a}_t &= 1,\ \text{if } t \geq s-1.\end{aligned}$$

This implies that $\{\tilde{D}_t\}_{t=0}^{\infty}$ is

$$\begin{aligned}\tilde{D}_0 &= D_0 + (1-\pi)^{n_{s-1}}\frac{\beta^{s-1}-\beta^s}{1-\beta(1-\pi)},\\ \tilde{D}_t &= D_t + (1-\pi)^{n_{s-1}-n_t}\frac{\beta^{s-1}-\beta^s}{1-\beta(1-\pi)},\ \text{for all } t < s-1,\\ \tilde{D}_t &= \frac{\beta^t}{1-\beta(1-\pi)},\ \text{for all } t \geq s-1.\end{aligned}$$

Define $\{\tilde{x}_t\}_{t=0}^{\infty}$ by

$$\begin{aligned}\tilde{x}_t &= x_t,\ \text{for all } t \leq s-1,\\ \tilde{x}_t &= x_{t+1},\ \text{for all } t \geq s.\end{aligned}$$

We will verify that $\{\tilde{x}_t\}_{t=0}^{\infty}$ satisfies (20) in the relaxed problem where $\{\tilde{a}_t\}_{t=0}^{\infty}$ is implemented.

For $t \leq s-1$,

$$\begin{aligned}
\sum_{i=1}^{t} \beta^i \tilde{x}_i = \sum_{i=1}^{t} \beta^i x_i \quad &\geq \quad \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + D_0 - D_t\right) - \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}\right) \\
&\geq \quad \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + \tilde{D}_0 - \tilde{D}_t\right) - \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}\right),
\end{aligned}$$

where the last inequality follows from $\tilde{D}_0 - \tilde{D}_t \leq D_0 - D_t$. For $t \geq s$, we need to show

$$\begin{aligned}
\sum_{i=1}^{s-1} \beta^i x_i + \sum_{i=s}^{t} \beta^i \tilde{x}_i \quad &= \quad \sum_{i=1}^{s-1} \beta^i x_i + \sum_{i=s}^{t} \beta^i x_{i+1} \\
&= \quad \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + D_0 - D_{s-1}\right) - \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}\right) + \\
&\quad\ \ \frac{1}{\beta}\left(\frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + D_0 - D_{t+1}\right) - \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + D_0 - D_s\right)\right), \\
&\geq \quad \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta} + \tilde{D}_0 - \tilde{D}_t\right) - \frac{1}{\gamma}\log\left(\frac{U_0}{1-\beta}\right).
\end{aligned}$$

Since $a_{s-1} = 0, D_s = D_{s-1}$, it is equivalent to show that

$$\log\left(\frac{U_0}{1-\beta} + D_0 - D_{t+1}\right) \quad \geq \quad (1-\beta)\log\left(\frac{U_0}{1-\beta} + D_0 - D_s\right) + \beta\log\left(\frac{U_0}{1-\beta} + \tilde{D}_0 - \tilde{D}_t\right),$$

which follows from the concavity of $\log(\cdot)$ and

$$\begin{aligned}
D_0 - D_{t+1} \quad &\geq \quad D_0 - \beta\tilde{D}_t \\
&\geq \quad D_0 - \beta\tilde{D}_t - (1-\beta)D_s + (1-\pi)^{n_{s-1}}\frac{(1-\beta)\beta^s}{1-\beta(1-\pi)} \\
&= \quad (1-\beta)(D_0 - D_s) + \beta\left(D_0 + (1-\pi)^{n_{s-1}}\frac{\beta^{s-1} - \beta^s}{1-\beta(1-\pi)} - \tilde{D}_t\right) \\
&= \quad (1-\beta)(D_0 - D_s) + \beta(\tilde{D}_0 - \tilde{D}_t),
\end{aligned}$$

where the second inequality follows from $D_s = \beta^s/(1-\beta(1-\pi))$. Thus $\tilde{\sigma}$ satisfies the I.C. constraints in the relaxed problem.

Contract $\tilde{\sigma}$ is more efficient than $\sigma^*(\{a_t\}_{t=0}^{\infty})$ in the sense that $\tilde{\sigma}$ lowers the cost of distortion per unit of wage income generated,

$$
\begin{aligned}
& \frac{\sum_{t=1}^{\infty}\beta^t(1-\pi)^{n_t}\chi_{\{a_{t-1}=1\}}f(x_t)}{\sum_{t=1}^{\infty}\beta^t(1-\pi)^{n_t}\chi_{\{a_{t-1}=1\}}} \\
= \;& \frac{\sum_{t=1}^{s-1}\beta^t(1-\pi)^{n_t}\chi_{\{a_{t-1}=1\}}f(x_t)+\sum_{t=s+1}^{\infty}\beta^t(1-\pi)^{n_t}f(x_t)}{\sum_{t=1}^{s-1}\beta^t(1-\pi)^{n_t}\chi_{\{a_{t-1}=1\}}+\sum_{t=s+1}^{\infty}\beta^t(1-\pi)^{n_t}} \\
> \;& \frac{\sum_{t=1}^{s-1}\beta^t(1-\pi)^{n_t}\chi_{\{a_{t-1}=1\}}f(x_t)+\sum_{t=s+1}^{\infty}\beta^{t-1}(1-\pi)^{n_t}f(x_t)}{\sum_{t=1}^{s-1}\beta^t(1-\pi)^{n_t}\chi_{\{a_{t-1}=1\}}+\sum_{t=s+1}^{\infty}\beta^{t-1}(1-\pi)^{n_t}} \\
= \;& \frac{\sum_{t=1}^{s-1}\beta^t(1-\pi)^{n_t}\chi_{\{a_{t-1}=1\}}f(x_t)+\sum_{t=s}^{\infty}\beta^t(1-\pi)^{\tilde{n}_t}f(\tilde{x}_t)}{\sum_{t=1}^{s-1}\beta^t(1-\pi)^{n_t}\chi_{\{a_{t-1}=1\}}+\sum_{t=s}^{\infty}\beta^t(1-\pi)^{\tilde{n}_t}},
\end{aligned}
$$

where the inequality follows from Lemma A.3. Hence $\sigma^*(\{a_t\}_{t=0}^{\infty})$ is dominated by a lottery with two outcomes: with probability $D_0/\tilde{D}_0$, the principal uses contract $\tilde{\sigma}$ to implement efforts $\{\tilde{a}_t\}_{t=0}^{\infty}$ and deliver promised utility $-U_0/(1-\beta)-\tilde{D}_0$; and with probability $(1-D_0/\tilde{D}_0)$, the principal implements $\{0\}_{t=0}^{\infty}$ and delivers utility $-U_0/(1-\beta)$. This lottery provides the same ex ante utility, but lowers the cost function $C(\sigma)$.

If there are still zero efforts remaining in $\{\tilde{a}_t\}_{t=0}^{\infty}$, we could repeat the above procedure and move forward the tail sequence of full efforts one step further. Eventually the first outcome in the lottery will specify $a_t = 1$, for all $t \geq 0$. *Q.E.D.*

Proof of Lemma 4: The proof follows from **Lemma A.1-A.5**. Let $\{a_t\}_{t=0}^{\infty}$, $a_t \in \{0, 1\}$, be an effort sequence that the principal wants to implement. And for $s > 0$, let $(\{a_t\}_{t=0}^{s-1}, \{1\}_{t=s}^{\infty})$ denote a modified sequence where the effort starting from s is always 1. The solution to the relaxed problem (19), $\sigma^*(\{a_t\}_{t=0}^{\infty})$, can be approximated by $\sigma^*(\{a_t\}_{t=0}^{s-1}, \{1\}_{t=s}^{\infty})$ as close as possible, when $s \to \infty$. Lemma A.5 states that for all s, $\sigma^*(\{a_t\}_{t=0}^{s-1}, \{1\}_{t=s}^{\infty})$ is dominated by lotteries with two outcomes, thus $\sigma^*(\{a_t\}_{t=0}^{\infty})$ is also dominated by a lottery with two outcomes. Since the optimal I.C. contract is always dominated by $\sigma^*(\{a_t\}_{t=0}^{\infty})$, we obtain the result. *Q.E.D.*

Proof of Lemma 5: Denote the agent's hidden assets at the beginning of period s by A_s^U. Initially the agent has zero asset, $A_0^U = 0$. In period $s \geq t$, since he will follow strategy

$\{a_s = 1\}_{s=t}^{\infty}$, the optimal consumption is the sum of the goods claimed from the principal and the interest payment $(1-\beta)A_t^U$, i.e., hidden savings is time invariant after period t. To see this, notice that $U_s = \pi E_{s+1} + (1-\pi)U_{s+1}$ and CARA utilities imply, for all $s \geq t$,

$$u'(c_s^U + (1-\beta)A_t^U) = \pi u'(c_{s+1}^E + (1-\beta)A_t^U) + (1-\pi)u'(c_{s+1}^U + (1-\beta)A_t^U).$$

An unemployed agent's period-t discounted utility of consumption taking A_t^U as a state variable is

$$F_t(A_t^U) = -\frac{\exp(-\gamma c_t^U)}{1-\beta}\exp(-\gamma(1-\beta)A_t^U) = -\frac{U_t}{1-\beta}\exp(-\gamma(1-\beta)A_t^U).$$

Now suppose an unemployed agent's discounted utility of consumption at period $s < t$ is $F_s(A_s^U) = -x\exp(-\gamma(1-\beta)A_s^U)/(1-\beta)$, where x is a parameter. We can calculate $F_{s-1}(A_{s-1}^U)$ as follows. With probability $\tilde{a}_{s-1}\pi$, the agent will find a job next period and has a utility $-E_s\exp(-\gamma(1-\beta)A_s^U)/(1-\beta)$, and with probability $(1-\tilde{a}_{s-1}\pi)$, he is still unemployed and has a utility $-x\exp(-\gamma(1-\beta)A_s^U)/(1-\beta)$. The agent's optimal savings problem at period $s-1$ is

$$\begin{aligned} F_{s-1}(A_{s-1}^U) &= \max_{c_{s-1},A_s^U} -\exp(-\gamma c_{s-1}) + \beta[\tilde{a}_{s-1}\pi(-\frac{E_s}{1-\beta}\exp(-\gamma(1-\beta)A_s^U)) \\ &\quad +(1-\tilde{a}_{s-1}\pi)(-\frac{x}{1-\beta}\exp(-\gamma(1-\beta)A_s^U))] \\ &s.t. \quad c_{s-1} + \beta A_s^U = A_{s-1}^U + c_{s-1}^U. \end{aligned} \tag{26}$$

The first-order condition is

$$\exp(-\gamma c_{s-1}) = [\tilde{a}_{s-1}\pi E_s + (1-\tilde{a}_{s-1}\pi)x]\exp(-\gamma(1-\beta)A_s^U). \tag{27}$$

Substituting Eq. (27) into Eq. (26) yields

$$\begin{aligned} c_{s-1} &= (1-\beta)A_{s-1}^U + (1-\beta)c_{s-1}^U + \beta\log(\tilde{a}_{s-1}\pi E_s + (1-\tilde{a}_{s-1}\pi)x)/(-\gamma) \\ A_s^U &= A_{s-1}^U + \log\left(\frac{\tilde{a}_{s-1}\pi E_s + (1-\tilde{a}_{s-1}\pi)x}{U_{s-1}}\right)/\gamma. \end{aligned} \tag{28}$$

Thus

$$\begin{aligned}
F_{s-1}(A_{s-1}^U) &= \frac{-\exp(-\gamma c_{s-1})}{1-\beta} \\
&= -\exp(-\gamma(1-\beta)A_{s-1}^U)(\exp(-\gamma c_{s-1}^U))^{1-\beta}\left[\tilde{a}_{s-1}\pi E_s + (1-\tilde{a}_{s-1}\pi)x\right]^\beta/(1-\beta) \\
&= -\frac{U_{s-1}^{1-\beta}\left[\tilde{a}_{s-1}\pi E_s + (1-\tilde{a}_{s-1}\pi)x\right]^\beta}{1-\beta}\exp(-\gamma(1-\beta)A_{s-1}^U).
\end{aligned}$$

It follows from the above recursive formula and $A_0^U = 0$ that the agent's utility in period $s = 0$ is

$$\frac{-U_0^{1-\beta}[\tilde{a}_0\pi E_1 + (1-\tilde{a}_0\pi)U_1^{1-\beta}[\tilde{a}_1\pi E_2 + (1-\tilde{a}_1\pi)U_2^{1-\beta}[...[\tilde{a}_{t-1}\pi E_t + (1-\tilde{a}_{t-1}\pi)U_t]^\beta...]^\beta]^\beta]^\beta}{1-\beta}.$$

Q.E.D.

Proof of Lemma 6: It is a special case of Lemma A.2. *Q.E.D.*

Proof of Lemma 7: Consider a finite-deviation strategy $\{\tilde{a}_s\}_{s=0}^\infty$, where $\tilde{a}_s = 1, \forall s \geq t+1$ for some t. Fix the efforts $\tilde{a}_s$ $(s \leq t-1)$ and vary $\tilde{a}_t$. Denote the utility in (6) by $F(\{\tilde{a}_s\}_{s=0}^{t-1}, \tilde{a}_t)$. Recall the disutility of $\{\tilde{a}_s\}_{s=0}^\infty$ is

$$\begin{aligned}
&D_0(\{\tilde{a}_s\}_{s=0}^{t-1}, \tilde{a}_t) \\
= &\left[\tilde{a}_0 + \beta(1-\pi\tilde{a}_0)\tilde{a}_1 + ... + \beta^t\left(\Pi_{s=0}^{t-1}(1-\pi\tilde{a}_s)\right)\left(\tilde{a}_t + \beta(1-\pi\tilde{a}_t)\frac{1}{1-\beta(1-\pi)}\right)\right].
\end{aligned}$$

In the following we only prove that setting $\tilde{a}_t = 1$ and keeping other efforts fixed will (weakly) improve the agent's utility, i.e.,

$$F(\{\tilde{a}_s\}_{s=0}^{t-1}, \tilde{a}_t) - D_0(\{\tilde{a}_s\}_{s=0}^{t-1}, \tilde{a}_t) \leq F(\{\tilde{a}_s\}_{s=0}^{t-1}, 1) - D_0(\{\tilde{a}_s\}_{s=0}^{t-1}, 1), \tag{29}$$

because then the lemma can be proved by backward induction.

First we prove (29) for a simple case, namely, $\tilde{a}_s = 0, \forall s \leq t-1$. Note that

$$F(\{0\}_{s=0}^{t-1}, \tilde{a}_t) = -\frac{\left(\Pi_{s=0}^t U_s^{(1-\beta)\beta^s}\right)(\tilde{a}_t\pi E_{t+1} + (1-\tilde{a}_t\pi)U_{t+1})^{\beta^{t+1}}}{1-\beta}$$

is a convex function of $\tilde{a}_t$, and $D_0(\{0\}_{s=0}^{t-1}, \tilde{a}_t)$ is a linear function. It follows from $F(\{0\}_{s=0}^{t-1}, 0) - D_0(\{0\}_{s=0}^{t-1}, 0) = F(\{0\}_{s=0}^{t-1}, 1) - D_0(\{0\}_{s=0}^{t-1}, 1)$ and the convexity of $F(\{0\}_{s=0}^{t-1}, \tilde{a}_t)$ that (29) holds.

Second we prove

$$\begin{aligned}
& (F(\{\tilde{a}_s\}_{s=0}^{t-1},1)-D_0(\{\tilde{a}_s\}_{s=0}^{t-1},1))-(F(\{\tilde{a}_s\}_{s=0}^{t-1},\tilde{a}_t)-D_0(\{\tilde{a}_s\}_{s=0}^{t-1},\tilde{a}_t)) \\
= & \int_{\tilde{a}_t}^{1}\left(\frac{\partial F(\{\tilde{a}_s\}_{s=0}^{t-1},a)}{\partial a}-\frac{\partial D_0(\{\tilde{a}_s\}_{s=0}^{t-1},a)}{\partial a}\right)da \geq 0,
\end{aligned}$$

for the general case. $\frac{\partial D_0(\{\tilde{a}_s\}_{s=0}^{t-1},a)}{\partial a}=\beta^t\left(\Pi_{s=0}^{t-1}(1-\pi\tilde{a}_s)\right)\frac{1-\beta}{1-\beta(1-\pi)}$ is independent of a. It can be derived from (6) that

$$\frac{\partial F(\{\tilde{a}_s\}_{s=0}^{t-1},a)}{\partial a}=\beta^t\left(\Pi_{s=0}^{t-1}(1-\pi\tilde{a}_s)\right)\frac{U_t-E_t}{1-\beta}\exp\left(-\gamma(1-\beta)A_t^U(\{\tilde{a}_s\}_{s=0}^{t-1},a)\right),$$

where $A_t^U(\{\tilde{a}_s\}_{s=0}^{t-1},a)$ is the hidden savings at the beginning of t, conditional on the strategy $(\{\tilde{a}_s\}_{s=0}^{t-1},a)$. It follows from (28) that the always-shirking-up-to-$(t-1)$ type saves the most, i.e.,

$$A_t^U(\{\tilde{a}_s\}_{s=0}^{t-1},a)\leq A_t^U(\{0\}_{s=0}^{t-1},a),$$

which implies that

$$\frac{\partial F(\{\tilde{a}_s\}_{s=0}^{t-1},a)}{\partial a}/\frac{\partial D_0(\{\tilde{a}_s\}_{s=0}^{t-1},a)}{\partial a}\quad\geq\quad\frac{\partial F(\{0\}_{s=0}^{t-1},a)}{\partial a}/\frac{\partial D_0(\{0\}_{s=0}^{t-1},a)}{\partial a},$$

meaning that the always-shirking type has an advantage in shirking in period t.

Hence

$$\begin{aligned}
& \int_{\tilde{a}_t}^{1}\left(\frac{\partial F(\{\tilde{a}_s\}_{s=0}^{t-1},a)}{\partial a}-\frac{\partial D_0(\{\tilde{a}_s\}_{s=0}^{t-1},a)}{\partial a}\right)da \\
= & \frac{\partial D_0(\{\tilde{a}_s\}_{s=0}^{t-1},a)}{\partial a}\int_{\tilde{a}_t}^{1}\left(\frac{\partial F(\{\tilde{a}_s\}_{s=0}^{t-1},a)}{\partial a}/\frac{\partial D_0(\{\tilde{a}_s\}_{s=0}^{t-1},a)}{\partial a}-1\right)da \\
\geq & \frac{\partial D_0(\{\tilde{a}_s\}_{s=0}^{t-1},a)}{\partial a}\int_{\tilde{a}_t}^{1}\left(\frac{\partial F(\{0\}_{s=0}^{t-1},a)}{\partial a}/\frac{\partial D_0(\{0\}_{s=0}^{t-1},a)}{\partial a}-1\right)da \\
\geq & 0,
\end{aligned}$$

where the last inequality follows from $F(\{0\}_{s=0}^{t-1},\tilde{a}_t)-D_0(\{0\}_{s=0}^{t-1},\tilde{a}_t)\leq F(\{0\}_{s=0}^{t-1},1)-D_0(\{0\}_{s=0}^{t-1},1)$. *Q.E.D.*

Proof of Proposition 2: Lemma A.3 states

$$\beta^t x_t = \frac{1}{\gamma}\log\left(-\bar{V} - \frac{\beta^t}{1-\beta(1-\pi)}\right) - \frac{1}{\gamma}\log\left(-\bar{V} - \frac{\beta^{t-1}}{1-\beta(1-\pi)}\right).$$

Therefore $\lim_{t\to\infty} x_t = (1-\beta)/(\gamma(-\bar{V})\beta(1-\beta(1-\pi))) > 0$. Recall that $U_t/U_{t-1} = \exp(\gamma x_t)$. *Q.E.D.*

Proof of Proposition 3: In Hopenhayn and Nicolini [4], if the principal always implements high efforts, then $\{U_t\}_{t=0}^{\infty}$ is monotonically increasing. We will show

$$\lim_{t\to\infty} U_t = \infty. \tag{30}$$

Suppose to the contrary that $\lim_{t\to\infty} U_t = M$ for some $0 < M < \infty$. Inverse Euler equation

$$\frac{1}{U_t} = \frac{\pi}{E_{t+1}} + \frac{1-\pi}{U_{t+1}} \tag{31}$$

implies that $\lim_{t\to\infty} E_t = M$. Let W_t^U be the promised utility for an agent who has not found a job at the beginning of t. It is easily seen that

$$W_t^U \geq -\sum_{s=t}^{\infty} \beta^{s-t} U_s \geq -\frac{M}{1-\beta},$$

which violates the incentive constraint $\beta\pi(-E_t/(1-\beta) - W_t^U) \geq 1$ for exerting effort for large t.

Secondly we show that

$$\limsup_{t\to\infty}(U_{t+1} - U_t) < \infty. \tag{32}$$

Eq. (31) and $U_t < U_{t+1}$ imply that $E_t < U_t < U_{t+1}$. Thus

$$\begin{aligned}\frac{U_{t+1} - U_t}{1-\beta} &< \frac{U_{t+1} - E_{t+1}}{1-\beta} < \sum_{s=t+1}^{\infty} \beta^{s-(t+1)} U_s - \frac{E_{t+1}}{1-\beta} \\ &= W_{t+1}^U - \frac{E_{t+1}}{1-\beta} = \frac{1}{\beta\pi},\end{aligned}$$

where the last equality follows from the binding incentive constraint.

Thirdly, we have

$$W^U_{t+2} - \frac{E_{t+2}}{1-\beta} = \frac{1}{\beta\pi}, \tag{33}$$

$$U_{t+1} + \beta W^U_{t+2} - \frac{E_{t+1}}{1-\beta} = W^U_{t+1} - \frac{E_{t+1}}{1-\beta} = \frac{1}{\beta\pi}. \tag{34}$$

Substituting Eq. (33) into Eq. (34) yields

$$(1-\beta)U_{t+1} + \beta E_{t+2} - E_{t+1} = \frac{(1-\beta)^2}{\beta\pi}. \tag{35}$$

Substituting Eq. (31) into Eq. (35) yields

$$(U_{t+2} - U_{t+1}) = \frac{M_t(U_{t+1} - U_t) - \frac{(1-\beta)^2}{\beta\pi}}{N_t}, \tag{36}$$

where $M_t = (U_t + (U_{t+1} - U_t))/(\pi U_t + (U_{t+1} - U_t))$, $N_t = (\beta(1-\pi)U_{t+1})/(\pi U_{t+1} + (U_{t+2} - U_{t+1}))$. Rewrite Eq. (36) as

$$(U_{t+2} - U_{t+1}) - C = \frac{M_t}{N_t}\left[(U_{t+1} - U_t) - C\right] + \frac{(M_t - N_t)C - \frac{(1-\beta)^2}{\beta\pi}}{N_t},$$

where $C = (1-\beta)^2/(\beta(1-\beta(1-\pi)))$. Eqs. (30) and (32) imply that $\lim_{t\to\infty} M_t = 1/\pi$ and $\lim_{t\to\infty} N_t = \beta(1-\pi)/\pi$. Therefore $\lim_{t\to\infty} M_t/N_t = 1/(\beta(1-\pi)) > 1$ and $\lim_{t\to\infty}((M_t - N_t)C - (1-\beta)^2/(\beta\pi))/N_t = 0$. For any $\epsilon > 0$ ($\epsilon < (1/(\beta(1-\pi)) - 1)/4$), there is a t^*, such that if $t \geq t^*$,

$$\frac{M_t}{N_t} \geq \frac{1}{\beta(1-\pi)} - \epsilon, \text{ and } \left|\frac{(M_t - N_t)C - \frac{(1-\beta)^2}{\beta\pi}}{N_t}\right| \leq \epsilon^2.$$

So if $|U_{t+1} - U_t - C| \geq \epsilon$,

$$\begin{aligned} |U_{t+2} - U_{t+1} - C| &\geq \left(\frac{1}{\beta(1-\pi)} - \epsilon\right)|U_{t+1} - U_t - C| - \epsilon^2 \\ &\geq \left(\frac{1}{\beta(1-\pi)} - 2\epsilon\right)|U_{t+1} - U_t - C| \\ &\geq \frac{\frac{1}{\beta(1-\pi)} + 1}{2}\,|U_{t+1} - U_t - C|. \end{aligned}$$

By induction, $|U_{t+2} - U_{t+1} - C|$ will grow at a geometric rate, contradicting Eq. (32). Therefore when $t \geq t^*$, $|U_{t+2} - U_{t+1} - C| \leq \epsilon$. Thus Eq. (13) holds. *Q.E.D.*

Proof of Proposition 4: If $\bar{V}^a > \bar{V}^b$, then it follows from Eq. (8) that $-U_0^a > -U_0^b$ and from Eq. (12) that $\lim_{t\to\infty} U_t^a/U_{t-1}^a > \lim_{t\to\infty} U_t^b/U_{t-1}^b$. *Q.E.D.*

Proof of Proposition 5: Lemma A.3 states that $x_t > x_{t+1} > 0$, for all $t \geq 1$. We will show that both $c_t^E - c_t^U$ and $c_t^E - c_{t+1}^E$ are strictly decreasing in t.

Eq. (3) implies

$$\exp(-\gamma(c_{t-1}^U - c_t^U)) = \pi \exp(-\gamma(c_t^E - c_t^U)) + (1-\pi),$$

which implies that $c_t^E - c_t^U$ is decreasing. Eq. (3) also implies

$$\begin{aligned}\frac{\exp(-\gamma c_{t+1}^E)}{\exp(-\gamma c_t^E)} &= \frac{\exp(-\gamma c_t^U) - (1-\pi)\exp(-\gamma c_{t+1}^U)}{\exp(-\gamma c_{t-1}^U) - (1-\pi)\exp(-\gamma c_t^U)} \\ &= \exp(\gamma x_t)\left(1 + (1-\pi)\exp(\gamma x_{t+1})\frac{\exp(\gamma(x_t - x_{t+1})) - 1}{1 - (1-\pi)\exp(\gamma x_t)}\right).\end{aligned}$$

Since $\exp(\gamma x_t)$, $(x_t - x_{t+1})$, and $1/(1-(1-\pi)\exp(\gamma x_t))$ are all decreasing in t, we conclude that $c_t^E - c_{t+1}^E$ is decreasing in t. Finally

$$N_t = c_t^U + \frac{w - c_t^E}{1-\beta} - \beta\frac{w - c_{t+1}^E}{1-\beta} - = w - (c_t^E - c_t^U) - \frac{\beta(c_t^E - c_{t+1}^E)}{1-\beta}$$

is strictly increasing in t. *Q.E.D.*

Acknowledgements. We thank Manuel Amador, B. Ravikumar, V.V. Chari, April Franco, Maxim Ivanov, Hyeok Jeong, Larry Jones, Ayca Kaya, Narayana Kocherlakota, Fernando Leiva, Elena Pastorino, Nicola Pavoni, Fabrizio Perri, Chris Phelan, Robert Shimer, Galina Vereschagina, Giovanni Violante, Ivan Werning, and Heng-fu Zou for helpful comments.

References

[1] Abraham, A., and N. Pavoni (2008): "Efficient Allocations with Moral Hazard and Hidden Borrowing and Lending: a Recursive Formulation," *Rev. Econ. Dynam.*, 11, 781–803.

[2] ——— (2009): "Optimal Income Taxation and Hidden Borrowing and Lending: The First-Order Approach in Two Periods," *Carlo Alberto Notebooks, No. 102.*

[3] Card, D., R. Chetty, and A. Weber (2007): "Cash-on-Hand and Competing Models of Intertemporal Behavior: New Evidence from the Labor Market," *Quart. J. Econ.*, 122, 1511–1560.

[4] Hopenhayn, H., and J. P. Nicolini (1997): "Optimal Unemployment Insurance," *J. Polit. Economy*, 105(2), 412–438.

[5] Kocherlakota, N. (2004): "Figuring Out the Impact of Hidden Savings on Optimal Unemployment Insurance," *Rev. Econ. Dynam.*, 7, 541–554.

[6] Rogerson, W. (1985): "Repeated Moral Hazard," *Econometrica*, 53, 69–76.

[7] Shavell, S., and L. Weiss (1979): "The Optimal Payment of Unemployment Insurance Benefits over Time," *J. Polit. Economy*, 87, 1347–1362.

[8] Shimer, R., and I. Werning (2008): "Liquidity and Insurance for the Unemployed," *Amer. Econ. Rev.*, 98(5), 1922–42.

[9] Werning, I. (2001): "Repeated Moral Hazard with Unmonitored Wealth: a Recursive First-Order Approach," *Working paper, MIT, http://econ-www.mit.edu/files/1264.*

[10] ——— (2002): "Optimal Unemployment Insurance with Unobservable Savings," *Working paper, MIT, http://econ-www.mit.edu/files/1267.*

[11] Williams, N. (2009): "On Dynamic Principal-Agent Problems in Continuous Time," *Working paper, University of Wisconsin - Madison, http://www.ssc.wisc.edu/ nwilliam/dynamic-pa2.pdf.*