Appendices

This online appendix contains two main sections. First, in Section A, I provide the proofs for the main results in the paper. I begin by formally proving convexity of the value functions, and then proceed to characterize Proposition 1. I also provide some details on the quantitative analysis of the baseline case in Sections A.1, which discusses calibration, and in Section A.2, which details the naïve planning problem.

Then, in Section B, I present and discuss the alternative specification of preferences, discussed in Sections 3.3 and 6.4 in the main text. I detail which analytical results are possible, prove these results, and discuss why this specification is more difficult analytically. I then provide the details of the quantitative analysis of this model, along with the Hopenhayn and Nicolini version. Both of these alternative specifications are discussed briefly in Section 6.4 in the main text.

A Proofs

This section presents proofs for the Lemmas and Propositions presented in the main text. First, for reference, recall the planning problem from the main body of the paper. For a given choice of effort, $\vec{a}_i$, denote $\pi_j(a_i^k)$ by $\pi_j^{k,i}, k = h, l$. Then, the planner solves:

$$G_j^i(w) = \min_{\tau_j^i} \mu \left\{ \pi_j^{h,i} \left[C(\mathrm{v}_j^h(e)) - y \right] + [1 - \pi_j^{h,i}] C(\mathrm{v}_j^h(u)) + \beta \left[\pi_j^{h,i} G_e(w_j^h(e)) + [1 - \pi_j^{h,i}] G_u(w_j^h(u)) \right] \right\}$$

$$+ (1-\mu) \left\{ \pi_j^{l,i} \left[C(\mathrm{v}_j^l(e)) - y \right] + [1 - \pi_j^{l,i}] C(\mathrm{v}_j^l(u)) + \beta \left[\pi_j^{l,i} G_e(w_j^l(e)) + [1 - \pi_j^{l,i}] G_u(w_j^l(u)) \right] \right\} \tag{26}$$

$$\text{s.t.} \quad w = \mu V_j(\theta_h, \theta_h, a_i^h) + (1-\mu) V_j(\theta_l, \theta_l, a_i^l) \tag{27}$$

$$V_j(\theta_h, \theta_h, a_i^h) \geq V_j(\theta_h, \theta_l, \tilde{a}), \forall \tilde{a} \in A_l^i \tag{28}$$

$$V_j(\theta_h, \theta_h, a_i^h) \geq V_j(\theta_h, \theta_h, \tilde{a}), \forall \tilde{a} \in A_h^i \tag{29}$$

$$V_j(\theta_l, \theta_l, a_i^l) \geq V_j(\theta_l, \theta_h, \tilde{a}), \forall \tilde{a} \in A_h^i \tag{30}$$

$$V_j(\theta_l, \theta_l, a_i^l) \geq V_j(\theta_l, \theta_l, \tilde{a}), \forall \tilde{a} \in A_l^i \tag{31}$$

First I describe the general planning problem with lotteries over the effort choices. Let $\vec{q} = \{q_i\}_{i=1}^3$ denote the choice of lottery, where q_i is the probability of effort choice $\vec{a}_i$. Similarly, let $\vec{w} = \{w_i\}_{i=1}^3$ denote the corresponding choice of promised utility. The value functions $G_j(w), j \in \{u, e\}$ solve

$$G_j(w) = \min_{\vec{q}, \vec{w}} \sum_{i=1}^3 q_i G_j^i(w_i) \tag{32}$$

$$\text{s.t.} \qquad w = \sum_{i=1}^3 q_i w_i \tag{33}$$

$$\sum_{i=1}^3 q_i = 1 \tag{34}$$

With this randomization over $\vec{a}$, the problem is convex, as the following states formally.

Proposition 2 *The value functions $G_j^i(w), i = \{1, 2, 3\}, j = \{u, e\}$ and $G_j(w), j = \{u, e\}$ are convex.*

For lower levels of promised utility, $\vec{a}_2$ dominates, and as w increases, eventually $\vec{a}_1$ dominates, with lotteries optimal for values in between. For even larger values of w, eventually inducing effort from an agent reporting θ_h also becomes prohibitively costly, and $\vec{a}_3$ dominates, again with lotteries possible for intermediate values of w. Intuitively, exerting effort remains costly, as it requires the planner to spread utilities (current or future promises) across employment states. Given a convex cost function, these spreads are costly, and become more costly for higher levels of w; therefore, as w increases, eventually the planner finds it too costly to induce effort from an agent reporting θ_l, and eventually for all reports.

In the analysis of the optimal contract, I assume the lower bound is not prohibitively high, which is summarized in the following assumption:

Assumption 1 *For some w, the optimal contract contains an interior solution for $w_j^k(u), k = \{h, l\}, j = \{e, u\}$.*

Proof of Proposition 2. First, if $G_j(w), j = \{u, e\}$ is convex, then standard results in dynamic programming establish each $G_j^i(w)$ is convex. Now assume that $G_j^i(w), j = \{u, e\}$ is convex, and consider $w^\lambda = \lambda w + (1-\lambda)w'$, $\lambda \in (0,1)$, $w, w' \in W$ (where W is the set of feasible future promised utilities). Further, let $\vec{q}, \vec{w}$ be optimal solutions for w, and $\vec{q}', \vec{w}'$ for w'. Now, consider the following choices for $\vec{q}^\lambda$ and $\vec{w}^\lambda$:

$$\begin{aligned} q_i^\lambda &= \lambda q_i + (1-\lambda)q_i' \\ w_i^\lambda &= \frac{\lambda q_i w_i + (1-\lambda)q_i' w_i'}{q_i^\lambda} \end{aligned}$$

Notice, these represent feasible solutions for the problem in (32)-(34). Since $\sum_i^3 q_{i=1} = 1$ and $\sum_i^3 q_{i=1}' = 1$ by assumption, then

$$\sum_{i=1}^3 q_i^\lambda = \lambda \sum_{i=1}^3 q_i + (1-\lambda) \sum_{i=1}^3 q_i' = 1$$

Further, the promised utility delivered by this allocation is

$$\tilde{w} = \sum_{i=1}^3 q_i^\lambda w_i^\lambda = \sum_{i=1}^3 \lambda q_i w_i + \sum_{i=1}^3 q_i' w_i' = \lambda w + (1-\lambda)w' = w^\lambda$$

The convexity of $G_j^i(w), i = \{1,2,3\}, j = \{u, e\}$ implies that for any η and $(w, w') \in W$,

$$G_j^i(\eta w + (1-\eta)w') \leq \eta G_j^i(w) + (1-\eta)G_j^i(w'). \tag{35}$$

Let $\eta = \frac{\lambda q_i}{q_i^\lambda}$ and given the choice of q_i^λ we can write $1 - \eta = \frac{(1-\lambda)q_i'}{q_i^\lambda}$. Then, equation (35) implies

$$G_j^i(w_i^\lambda) \leq \frac{\lambda q_i G_j^i(w_i) + (1-\lambda)q_i' G_j^i(w_i')}{q_i^\lambda} \tag{36}$$

To complete the proof, the definition of $G_j(w)$ along with equation (36) imply

$$G_j(w^\lambda) \leq \sum_{i=1}^3 q_i^\lambda G_j^i(w_i^\lambda) \leq \lambda \sum_{i=1}^3 q_i G_j^i(w_i) + (1-\lambda) \sum_{i=1}^3 q_i' G_j^i(w_i') = \lambda G_j(w) + (1-\lambda)G_j(w')$$

Therefore, by definition, $G_j(w), j = \{u, e\}$ is convex. ■

Next, I restate the main Proposition from the text:

Proposition 1 *There exists a $\delta > 0$ such that an agent starting an unemployment spell with promised lifetime utility $w_0 \in [\underline{w}, \underline{w} + \delta)$, receives increasing consumption over the length of the unemployment spell with positive probability.*

To prove Proposition 1, several preliminary results are necessary.

Lemma 1 *If there exists interior $w^h_j(i)$ (i.e. $w^h_j(i) > \underline{w}$) for some w, then for any optimal $v^h_j(i)$ and $w^h_j(i) \geq \underline{w}$,*

$$\frac{C'(v^h_j(i))}{\theta_h} \geq G'_i(w^h_j(i))$$

for $i = u, e$ and $j = u, e$.

Proof. The proof begins by showing that for interior $w^h_j(i)$, $\dfrac{C'(\mathrm{v}^h_j(i))}{\theta_h} > G'_i(w^h_j(i))$. Towards this end, suppose instead $\dfrac{C'(\mathrm{v}^h_j(i))}{\theta_h} \leq G'_i(w^h_j(i))$, where $\mathrm{v}^h_j(i), w^h_j(i)$ are optimal solutions, and $w^h_j(i) > \underline{w}$. Then, consider decreasing $w^h_j(i)$ by ε and increasing $\mathrm{v}^h_j(i)$ by $\frac{\beta\varepsilon}{\theta_h}$. This change is possible since by assumption $w^h_j(u)$ is interior. Notice, this leaves $V(\theta_h, \theta_h, a^h)$ unchanged; therefore the constraints in (27)-(29) are unaffected. Further, since $\theta_h > \theta_l$, the constraints in (30)-(31) are still satisfied under this alternative allocation. The R.H.S. of each decreases more than they increase. The alternative allocation, however, decreases costs by $\mu\beta(1-\pi_j)G'_i(w^h_j(i))\varepsilon$, and increases them by $\frac{\mu\beta(1-\pi_j)C'(\mathrm{v}^h_j(i))\varepsilon}{\theta_h}$. Since $\frac{C'(\mathrm{v}^h_j(i))}{\theta_h} \leq G'_i(w^h_j(i))$ by assumption, this makes the planner no worse off, a contradiction to $\mathrm{v}^h_j(i), w^h_j(i)$ as optimal solutions. Then, the continuity of the policy functions, combined with the continuity of $C(\mathrm{v})$ and $G'_i(w)$, establishes $\dfrac{C'(\mathrm{v}^h_j(i))}{\theta_h} \geq G'_i(w^h_j(i))$. ■

Lemma 2 *If there exists interior $w^l_j(i)$ (i.e. $w^l_j(i) > \underline{w}$) for some w, then for any optimal $v^l_j(i)$ and $w^l_j(i) \geq \underline{w}$,*

$$\frac{C'(v^l_j(i))}{\theta_l} \leq G'_i(w^l_j(i))$$

for $i = u, e$ and $j = u, e$.

Proof. Similarly to the proof of Lemma 1, I begin by establishing $\dfrac{C'(\mathrm{v}^l_j(i))}{\theta_l} < G'_i(w^l_j(i))$ for $w^l_j(i) > \underline{w}$. Suppose instead $\frac{C'(\mathrm{v}^l_j(i))}{\theta_l} \geq G'_i(w^l_j(i))$, where $\mathrm{v}^l_j(i), w^l_j(i)$ are optimal solutions. Then, consider increasing $w^l_j(i)$ by ε and decreasing $\mathrm{v}^l_j(i)$ by $\frac{\beta\varepsilon}{\theta_l}$. Notice, this leaves $V(\theta_l, \theta_l, a^l)$ unchanged; therefore the constraints in (27), (30) and (31) are unaffected. Further, since $\theta_h > \theta_l$, the constraints in (28)-(29) are still satisfied under this alternative allocation. The R.H.S. of each decreases more than they increase. The alternative allocation, however, increases costs by $(1-\mu)\beta(1-\pi_j)G'_i(w^l_j(i))\varepsilon$, and decreases them by $\frac{(1-\mu)\beta(1-\pi_j)C'(\mathrm{v}^l_j(i))\varepsilon}{\theta_l}$. Since $\frac{C'(\mathrm{v}^l_j(i))}{\theta_l} \geq G'_i(w^l_j(i))$ by assumption, this makes the planner no worse off, a contradiction to $\mathrm{v}^l_j(i), w^l_j(i)$ as optimal solutions. Then, the continuity of the policy functions, combined with the continuity of $C(\mathrm{v})$ and $G'_i(w)$, establishes $\dfrac{C'(\mathrm{v}^l_j(i))}{\theta_l} \leq G'_i(w^l_j(i))$. ■

Lemma 3 *For the case of $\vec{a}_2 = (1, 1)$, the optimal contract satisfies $v^k_j(e) > v^k_j(u)$, for $k = h, l$ and $j = e, u$.*

Proof. The proof proceeds for $k = h$, as $k = l$ follows analogously. Suppose instead $\mathrm{v}^h_j(u) \geq \mathrm{v}^h_j(e)$. Then, consider decreasing $\mathrm{v}^h_j(u)$ by ϵ, and increasing $\mathrm{v}^h_j(e)$ by $\frac{1-\pi_j}{\pi_j}\epsilon$. Notice, this change will leave the constraints in (27)-(31) intact. Moreover, it will decrease the R.H.S. of (26) by $\mu(1-\pi_j)C'(\mathrm{v}_j^h(u))\epsilon$, and increase it by $\mu(1-\pi_j)C'(\mathrm{v}_j^h(e))\epsilon$. Since $\mathrm{v}_j^h(u) \geq \mathrm{v}_j^h(e)$, and $C(\cdot)$ is strictly convex, $C'(\mathrm{v}_j^h(u)) \geq C'(\mathrm{v}_j^h(e))$; therefore, this change will not increase the R.H.S. of (26), a contradiction. ■

In general, this result holds in any effort allocation where it remains relevant. For the aforementioned case, both reports exert effort and consequently there exist allocations for both when reporting employment and unemployment. For the effort allocation $\vec{a}_1 = (1,0)$, the result only applies to an agent reporting θ_h, since a report of θ_l cannot be accompanied by a report of employment.

Lemma 4 *For the case of $\vec{a}_2 = (1,1)$, the constraint in (31) must bind.*

Proof. For this choice of effort by the planner, two of the incentive constraints are

$$V(\theta_h, \theta_h, 1) \geq V(\theta_h, \theta_l, 0) \tag{37}$$
$$V(\theta_l, \theta_l, 1) \geq V(\theta_l, \theta_l, 0) \tag{38}$$

The first constraint is incentive compatibility for an agent reporting θ_h, who may report θ_l and shirk on effort. The proof begins by noticing that both (37) and (38) cannot simultaneously remain slack. To see this, suppose instead both remain slack. Then, consider decreasing $\mathrm{v}_j^l(e)$ by ϵ, and increasing $\mathrm{v}_j^l(u)$ by $\frac{\pi_j}{1-\pi_j}\epsilon$. Since (37) and (38) remain slack, for ϵ small enough, this still satisfies (27)-(31). Now, costs decrease by $(1-\mu)\pi_j C'(\mathrm{v}_j^l(e))\epsilon$, and they increase by $(1-\mu)\pi_j C'(\mathrm{v}_j^l(u))\epsilon$. From Lemma 3 and the strict convexity of $C(\cdot)$, this change makes the planner strictly better off, a contradiction. Thus, either (37) or (38) must bind. Again using Lemma 3 and $\theta_h > \theta_l$, the following must be true.

$$\theta_h \left[\mathrm{v}_j^l(e) - \mathrm{v}_j^l(u)\right] + \beta \left[w_j^l(e) - w_j^l(u)\right] > \theta_l \left[\mathrm{v}_j^l(e) - \mathrm{v}_j^l(u)\right] + \beta \left[w_j^l(e) - w_j^l(u)\right] \geq \frac{\nu}{\pi_j} \tag{39}$$

Now, suppose (37) binds. Then, using (28) for $\tilde{a} = 1$ and (37),

$$\theta_h \left[\mathrm{v}_j^l(e) - \mathrm{v}_j^l(u)\right] + \beta \left[w_j^l(e) - w_j^l(u)\right] \leq \frac{\nu}{\pi_j}$$

However, this contradicts (39). Thus, (37) remains slack, and from the discussion above, (38) (i.e. (31)) must bind. ■

Lemma 5 *The optimal contract satisfies $w_j^l(u) > w_j^h(u)$.*

Proof. Suppose instead that $w_j^h(u) \geq w_j^l(u)$. Given the convexity of $G_u(w)$, Lemmas 1 and 2 imply that $\dfrac{C'(\mathrm{v}_j^h(u))}{\theta_h} \geq \dfrac{C'(\mathrm{v}_j^l(u))}{\theta_l}$, which given the convexity of $C(\mathrm{v})$ and $\theta_h > \theta_l$ gives $\mathrm{v}_j^h(u) > \mathrm{v}_j^l(u)$. Thus, $V(\theta_l, \theta_h, 0) > V(\theta_l, \theta_l, 0)$; i.e. an agent receiving the shock θ_l and not exerting effort is always better off reporting θ_h. For the cases of $\vec{a}_1$ and $\vec{a}_3$, this clearly violates incentive compatibility since the θ_l report is not asked to exert effort by the planner. In the case of $\vec{a}_2$, where the θ_l report does exert effort, Lemma 4 and incentive compatibility constraints imply $V(\theta_l, \theta_l, 0) = V(\theta_l, \theta_l, 1) \geq V(\theta_l, \theta_h, 0)$, a contradiction. Therefore, it must be the case that $w_j^l(u) > w_j^h(u)$. ■

The intuition for Lemma 5 is straightforward. Given $\theta_h > \theta_l$, the planner allocates $\mathrm{v}_j^h(u) > \mathrm{v}_j^l(u)$. In order to maintain incentives for an agent receiving θ_l to truthfully report, it must be that $w_j^l(u) > w_j^h(u)$.

Lemma 6 *There exists a $\gamma > 0$ such that for all $w \in [\underline{w}, \underline{w} + \gamma]$, $w_j^l(u)(w) > w$.*

Proof. From Lemma 5, $w_j^l(u)(\underline{w}) > w_j^h(u)(\underline{w})$. Since these policy functions are continuous, a standard $\varepsilon - \gamma$ argument produces a $\gamma > 0$ such that for all $w \in [\underline{w}, \underline{w} + \gamma]$, $w_j^l(u)(w) > w$. ■

Lemma 7 *The policy function* $v_u^l(u)(w)$ *is strictly increasing in* w.

Proof. Define the following:

$$\mathcal{V}(\theta_l, \theta_l, a_u^l) \equiv V(\theta_l, \theta_l, a_u^l) - \theta_l \mathrm{v}_u^l(u, w)$$

Using (27) to solve for $\mathrm{v}_u^l(u)$ gives

$$\mathrm{v}_u^l(u, w) = \frac{w - \mu V(\theta_h, \theta_h, a_u^h) - (1-\mu)\mathcal{V}(\theta_l, \theta_l, a_u^l)}{(1-\mu)\theta_l} \tag{40}$$

Plugging this expression into (26) for $j = u$, and differentiating with respect to w gives

$$\left(G_u^i\right)'(w) = \frac{(1 - \pi_u(a_u^l))}{\theta_l} C'(\mathrm{v}_u^l(u, w)) \tag{41}$$

From the strict convexity of $G_j^i(\cdot), i \in \{1, 2, 3\}$ and $C(\cdot)$, as w increases, $\mathrm{v}_u^l(u)(w)$ must also increase. ■.

Now consider the proof of Proposition 1. In the following proof, I focus on a particular allocation for effort; i.e. any given $\vec{a}_i$. Since Lemmas 5 and 6 apply to all allocations of effort $(i = 1, 2, 3)$, it is straightforward to apply the proof to the case where the planner randomizes between different effort allocations, as any convex combination of $\mathrm{v}_j^l(u, i)$ and $w_j^l(u, i)$ still satisfies Lemmas 5 and 6.

Proof of Proposition 1. From Lemma 6, there exists a $\gamma > 0$ such that for all $w \in [\underline{w}, \underline{w} + \gamma]$, $w_u^l(u)(w) > w$. Set $\delta = \gamma$. Now consider an agent unemployed for $n > 1$ periods, and starting the unemployment spell with $w = w_0$, where $w_0 \in [\underline{w}, \underline{w} + \delta)$. Furthermore, consider an agent who receives a sequence of n θ_l shocks, $(\theta_l, \theta_l, ..., \theta_l)$. In period 1 of the unemployment spell, the agent receives consumption $c_1 = C(\mathrm{v}_u^l(u)(w_0))$. Next period, he enters with promised lifetime utility $w_u^l(u)(w_0)$. Since $w_0 \in [\underline{w}, \underline{w} + \gamma)$, from Lemma 6, $w_u^l(u)(w_0) > w_0$. From Lemma 7 and since $C(\cdot)$ is strictly increasing, consumption in period 2, $c_2 = C(\mathrm{v}_u^l(u)(w_u^l(u)(w_0))) > c_1$. This continues until either (i)the unemployment spell ends, or (ii)eventually $w_u^l(u)(w) = w$, at which point consumption remains constant. Thus, consumption increases over the duration of unemployment when the agent receives the history $(\theta_l, \theta_l, ...\theta_l)$. This history occurs with probability $(1-\mu)^n > 0$; therefore, consumption increases over the duration of unemployment with positive probability. ■

A.1 Calibration of $\frac{\theta_h}{\theta_l}$

The smallest ratio of $\frac{\theta_h}{\theta_l}$ that permits the unemployment to non-participation transition can be solved for directly. From (12) and (13) (in the main text), an agent, in the first period of unemployment $(j = u)$ receiving the θ_h shock, remains indifferent between exerting effort or not when

$$\theta_h \pi_u[\mathrm{v}(y) - \mathrm{v}(b)] + \beta \pi_u \left\{ \mu [V_e^{US}(\theta_h) - V_{nb}^{US}(\theta_h)] + (1-\mu)[V_e^{US}(\theta_l) - V_{nb}^{US}(\theta_l)] \right\} = \nu \tag{42}$$

Similarly, the agent in state $j = nb$, receiving the shock θ_l, remains indifferent between exerting effort and not when

$$\theta_l \pi_u[\mathrm{v}(y) - \mathrm{v}(d)] + \beta \pi_u \left\{ \mu [V_e^{US}(\theta_h) - V_{nb}^{US}(\theta_h)] + (1-\mu)[V_e^{US}(\theta_l) - V_{nb}^{US}(\theta_l)] \right\} = \nu \tag{43}$$

Combining (42) and (43) implies

$$\frac{\theta_h}{\theta_l} = \frac{[\mathrm{v}(y-\tau_b) - \mathrm{v}(d)]}{[\mathrm{v}(y-\tau_b) - \mathrm{v}(b)]} \tag{44}$$

Given the values for θ_h and θ_l, I chose ν to satisfy (42).

A.2 Naïve Allocation

The cost of providing the naïve allocation, to an agent in employment state j last period, denoted by $\tilde{G}_j(w)$, solves the following program:

$$\tilde{G}_j(w) = \min \left\{ \pi_j \left[C(\tilde{\mathrm{v}}_j(e)) - y \right] + (1-\pi_j) C(\tilde{\mathrm{v}}_j(u)) + \beta \left[\pi_j \tilde{G}_e(\tilde{w}_j(e)) + (1-\pi_j)\tilde{G}_u(\tilde{w}_j(u)) \right] \right\} \tag{45}$$

$$\text{s.t.} \quad w = \pi_j E(\theta) \left[\tilde{\mathrm{v}}_j(e) + (1-\pi_j)\tilde{\mathrm{v}}_j(u) \right] - \nu + \beta \left[\pi_j \tilde{w}_j(e) + (1-\pi_j)\tilde{w}_j(u) \right] \tag{46}$$

$$\pi_j E(\theta) \left[\tilde{\mathrm{v}}_j(e) + (1-\pi_j)\tilde{\mathrm{v}}_j(u) \right] - \nu + \beta \left[\pi_j \tilde{w}_j(e) + (1-\pi_j)\tilde{w}_j(u) \right] \geq E(\theta)\tilde{\mathrm{v}}_j(u) + \beta \tilde{w}_j(u) \tag{47}$$

Equation (46) represents the promise keeping constraint, while (47) is the incentive compatibility constraint. Notice, the planner does not recognize agents receive a taste shock $\theta \in \{\theta_l, \theta_h\}$, but instead believes preferences are based on $E(\theta)$. The latter fact ensures the allocation satisfies promise keeping, so agents prefer to participate.

B Alternative Specification of Preferences

This section describes in more detail the alternative specification of the model, and describes the analogous analytical results to those in Section 4.1. Preferences are now given by,

$$E_0 \sum_{t=0}^{\infty} \beta^t [\mathrm{v}(c_t) - \theta_t \nu(a_t)]$$

so the preference shock affects the utility cost of effort directly. Under this alternative specification, the timing of the model may be simplified, so that effort only affects next period's employment state, instead of the lottery occurring before consumption is allocated. In addition, I assume that when exerting effort the probability of finding/retaining a job is π_j, and if not exerting effort there exists positive probability of finding/retaining a job, $\underline{\pi}$, where $\pi_j > \underline{\pi}$. This assumption does not affect the results, but makes the analytical results more transparent. In the quantitative analysis of this version I set $\underline{\pi}$ very close to zero.

An agent enters each period with his employment status from last period, j, and his expected lifetime utility promise, w, known. The contract specifies utility for the current period (depending on the θ reported), and the agent's future lifetime utility promise, depending on the reported θ and whether the agent finds employment during the period or not. Thus, the planner offers the following contract to an agent entering the period in employment state j:

$$\tau_j = \left\{ a_j^k, \mathrm{v}_j^k, w_j^k(u), w_j^k(e), \right\}, k = h, l$$

Notice, the agent only receives consumption (v_j^h) depending on the report of θ, while the future promised utilities depend on both the report and next period's employment status. In the version presented in Section 3, both consumption and future utilities depend on the employment draw.

Similarly to before, denote the expected lifetime utility from the contract τ_j, for an agent receiving the shock θ_i, reporting the shock θ_k, and exerting effort a, by $V_j(\theta_i, \theta_k, a)$. This is given by,

$$V_j(\theta_i, \theta_k, 1) = \mathrm{v}_j^k - \theta_i \nu + \beta \left[\pi_j w_j^k(e) + (1-\pi_j) w_j^k(u)\right] \tag{48}$$

$$V_j(\theta_i, \theta_k, 0) = \mathrm{v}_j^k + \beta \left[\underline{\pi} w_j^k(e) + (1-\underline{\pi}) w_j^k(u)\right] \tag{49}$$

This agent exerts effort if $V_j(\theta_i, \theta_k, 1) \geq V_j(\theta_i, \theta_k, 0)$, which occurs when,

$$\beta(\pi_j - \underline{\pi}) \left[w_j^k(e) - w_j^k(u)\right] \geq \theta_i \nu \tag{50}$$

B.1 Planning Problem

The planner's problem is similar to that presented in Section 4, with only the timing and the agent's utility calculation changing. The analogous problem is given by (where χ represents an indicator variable equal to 1 if the agent is employed):

$$\begin{aligned}
G_j^i(w) = \min_{\tau_j^i} \mu &\left\{C(\mathrm{v}_j^h) - \chi y + \beta \left[\pi_j^{h,i} G_e(w_j^h(e)) + [1-\pi_j^{h,i}] G_u(w_j^h(u))\right]\right\} \\
+(1-\mu) &\left\{C(\mathrm{v}_j^l) - \chi y + \beta \left[\pi_j^{l,i} G_e(w_j^l(e)) + [1-\pi_j^{l,i}] G_u(w_j^l(u))\right]\right\} && (51)\\
\text{s.t.} \quad & w = \mu V_j(\theta_h, \theta_h, a_i^h) + (1-\mu) V_j(\theta_l, \theta_l, a_i^l) && (52)\\
& V_j(\theta_h, \theta_h, a_i^h) \geq V_j(\theta_h, \theta_l, \tilde{a}), \forall \tilde{a} \in A_l^i && (53)\\
& V_j(\theta_h, \theta_h, a_i^h) \geq V_j(\theta_h, \theta_h, \tilde{a}), \forall \tilde{a} \in A_h^i && (54)\\
& V_j(\theta_l, \theta_l, a_i^l) \geq V_j(\theta_l, \theta_h, \tilde{a}), \forall \tilde{a} \in A_h^i && (55)\\
& V_j(\theta_l, \theta_l, a_i^l) \geq V_j(\theta_l, \theta_l, \tilde{a}), \forall \tilde{a} \in A_l^i && (56)
\end{aligned}$$

This is the problem for a given effort allocation. Recall, $\vec{a}_i = (a_i^l, a_i^h)$, where $\vec{a}_1 = (0,1)$, $\vec{a}_2 = (1,1)$, $\vec{a}_3 = (0,0)$, and $\vec{a}_4 = (1,0)$. Moreover, unlike the baseline specification, the sets A_k^i are all the same. That is, an agent now always has the choice of exerting effort, since when not exerting effort there exists some probability of finding a job. The full problem also involves a choice of lotteries over the different effort allocations. As in the main body of the paper, I let $\vec{q} = \{q_i\}_{i=1}^3$ denote the choice of lottery, where q_i represents the probability of effort choice $\vec{a}_i$. Note that while the notation is identical to the baseline case presented in the main text, which allocations of effort are relevant is different under this specification of preferences. Now $\vec{a}_1 = (0,1)$ is irrelevant, as the planner never prefers to have the θ_h report exert effort ("bad" shock in this case) while the θ_l report does not.

B.2 Properties of the Optimal Contract

The primary goal of this section is to show that the increasing consumption result from Proposition 1 still obtains in this environment. Under this specification of preferences, fewer analytical results are available, but I can provide some characterization of the optimal contract. First, I only consider the case when the planner asks for effort from the agent reporting θ_l, but not from θ_h; i.e. $\vec{a}_4 = (1,0)$. The case where both agents are required to exert effort is both more difficult to characterize, and potentially less interesting, which I explain below. Also note, for the remainder of this appendix, for notational convenience, I often suppress the subscript for current employment status j, although it is implied everywhere.

The key to the main theoretical result in Proposition 1 is to show that for an interval around the lower bound, w increases under the optimal allocation. Towards this end, consider the following results.

Lemma 8 *The following relationships hold in the optimal contract: (i) $G'_e(w^l(e)) \geq G'_u(w^l(u))$; (ii) $C'(v^l) \geq G'_u(w^l(u))$; (iii) for interior $w^h(e)$, $G'_u(w^h(u)) > G'_e(w^h(e))$; and (iv) $G'_u(w^h(u)) > C'(v^h)$.*

Proof. To prove (i), first note that since $\theta_l \nu > 0$, satisfying (56) requires $w^l(e) > w^l(u)$; therefore, $w^l(e)$ is always interior. Next, I first consider the case where $w^l(u)$ is also interior. Suppose instead that $G'_u(w^l(u)) \geq G'_e(w^l(e))$. Then decrease $w^l(u)$ by ε and increase $w^l(e)$ by $\dfrac{1-\pi}{\pi}\varepsilon$. This maintains incentive compatibility, and given the assumption that $G'_u(w^l(u)) \geq G'_e(w^l(e))$, this change makes the planner better off. Therefore, for interior $w^l(u)$, $G'_e(w^l(e)) > G'_u(w^l(u))$. The result in (i) follows from the continuity of $G'_j(w)$. (ii) is proved in a similar fashion. Begin with interior $w^l(u)$, and assume instead that $G'_u(w^l(u)) \geq C'(\mathrm{v}^l)$. Then increase v^l by ε and decrease $w^l(u)$ by $\dfrac{\varepsilon}{\beta(1-\pi)}$. This maintains incentives and makes the planner better off. Continuity gives (ii).

To show (iii), suppose instead $G'(w^h(u)) \leq G'(w^h(e))$. Then, decrease $w^h(e)$ by ε and increase $w^h(u)$ by $\dfrac{\pi}{(1-\underline{\pi})}\varepsilon$. This maintains incentives and makes the planner better off, a contradiction. Finally, to show (iv), suppose instead that $C'(\mathrm{v}) \geq G'_u(w^h(u))$. Then, decrease v^h by ε and increase $w^h(u)$ by $\dfrac{\varepsilon}{\beta(1-\underline{\pi})}$. This maintains incentives while making the planner better off, a contradiction. ■

The key to the increasing consumption result is to show an interior $w^k(u)$ always obtains (in this case it occurs for the θ_h report). The following Lemma shows this key step for a particular effort allocation.

Lemma 9 *For the case of $\vec{a}_4 = (1,0)$, $w^h_u(u) > w^l_u(u)$.*

Proof. Suppose instead that $w^l(u) \geq w^h(u)$. From Lemma 8:

$$C'(\mathrm{v}^l) \geq G'_u(w^l(u)) \geq G'_u(w^h(u)) > C'(\mathrm{v}^h) \tag{57}$$

which implies $\mathrm{v}^l > \mathrm{v}^h$. Since $w^l(u) \geq w^h(u)$ by assumption, to maintain incentives (prevent θ_h report from lying), it must be that $w^h(e) > w^l(e)$. This implies $w^h(e)$ is interior, so that (iii) of Lemma 8 holds, which gives:

$$G'_e(w^l(e)) \geq G'_u(w^l(u)) \geq G'_u(w^h(u)) > G'_e(w^h(e)) \tag{58}$$

However, this implies $w^l(e) > w^h(e)$, a contradiction. Therefore, $w^h_u(u) > w^l_u(u)$. ■

Lemma 9 applies to the case where the planner prefers to have an agent reporting θ_l exert effort, but requests no effort from an agent reporting θ_h. Intuitively, given the effort allocation, the planner must ask the agent for their preference shock; in order to provide incentives for the agent to truthfully reveal their shock, the planner provides some spreads in their respective allocations. The implication of this result is the analogous result to Proposition 1. Before presenting that result, the following Lemma is used in the proof:

Lemma 10 *At the optimal solution, $\mathrm{v}^h(w)$ is increasing in w.*

Proof. The proof follows the proof of Lemma 7 and the details are omitted here. ■

Proposition 3 *For the case of $\vec{a}_4 = (1,0)$, there exists a $\delta > 0$ such that an agent starting an unemployment spell with promised lifetime utility $w_0 \in [\underline{w}, \underline{w} + \delta)$, receives increasing consumption over the length of the unemployment spell with positive probability.*

Proof. The proof follows closely the proof of Proposition 1 and thus the details are omitted here. Again, the key fact is that $w^h(u) > w^l(u) \geq \underline{w}$; therefore, using the continuity of the policy functions, there exists an interval around $\underline{w}$ where w is increasing for the agent reporting θ_h. Since w is increasing, from Lemma 10, v^h and thus consumption are also increasing for an agent remaining unemployed and receiving a string of consecutive θ_h shocks, which occurs with positive probability. ■

B.3 Technical Issues

In the case where both agents exert effort, it remains more difficult to fully characterize the allocations. Intuitively, since the planner prefers the same effort allocation for both reports, there may not exist any reason for her to offer different allocations to each report; i.e. a "pooling" contract may be optimal. While in many incentive problems a pooling equilibrium represents an interesting feature worth exploring, here it remains less interesting. Essentially, the planner's choices are restricted along the dimension (effort) she most prefers to have flexibility. If the number of effort choices is sufficiently large, the planner always prefers a different choice of effort for each report, and thus separates allocations in a similar manner to Lemma 9. For example, a "pooling" contract is never optimal if effort is a continuous variable.

Few analytical results are possible when the planner allocates effort to both reports. Moreover, the larger issue is the nature of the economy in this case. The shocks affect the utility cost of effort directly, but with only two choices of effort, the planner has limited options for minimizing costs while providing incentives. This represents the reason that I use the version where the shock affects the marginal utility from consumption directly as the baseline case.

Although more appealing analytically, as mentioned in the main body of the paper, the baseline specification of preferences may affect the magnitude of the welfare gains, since even the first best allocation varies consumption by reported θ. To gauge how relevant this is for the welfare gains presented, I now calibrate the alternative specification of preferences and compute the analogous welfare gains.

B.4 Quantitative Analysis

This section describes the calibration of the alternative model, and then offers a comparison of the welfare gains from this model and the baseline specification. This comparison is made as follows. First, I focus only on the partial equilibrium case; i.e. I abstract from the general equilibrium version of the planning problem. The general equilibrium version is very useful for endogenizing the lower bound and linking the theoretical and quantitative results. The focus here, however, is simply on the size of the welfare gains in each specification. What matters is the allocations of consumption and effort across the taste shocks, not the distributional issues involved in the general equilibrium version. Given these factors, and the additional computational cost of computing the general equilibrium version, I only present the partial equilibrium case here.

To calculate the welfare gains, in each specification I find the level of expected lifetime utility in the optimal contract, denoted w_P, such that $G_e(w_P) = 0$. The corresponding value for the naïve planner's allocation is denoted w_N. I then compare w_P (w_N) with w_{US}, the level of expected lifetime utility delivered to an employed agent under the current U.S. system, which remains self-financing.

B.4.1 Calibration

I first describe the calibration of the alternative specification of preferences. The model is similar except in the timing; agents consume before the employment shock is realized. As a result, in this alternative specification, there is an additional non-employed state relative to the baseline specification (an additional period of benefit eligibility).

Given the U.S. system previously described, we can define an agents expected lifetime utility, depending on her current employment state. First, if employed, having received the shock θ, expected lifetime utility is given by:

$$V_e^{US}(\theta) = \mathrm{v}(y-T) - \theta\nu + \beta\left[\pi_e E_{\theta'} V_e^{US}(\theta') + (1-\pi_e) E_{\theta'} V_1^{US}(\theta')\right] \tag{59}$$

Then, for an agent in the first period of non-employment, they are eligible for and collect benefits. Their value functions are given by (depending on search effort):

$$V_1^{US}(\theta, a=1) = \mathrm{v}(b) - \theta\nu + \beta\left[\pi_u E_{\theta'} V_e^{US}(\theta') + (1-\pi_u) E_{\theta'} V_2^{US}(\theta')\right] \tag{60}$$

$$V_1^{US}(\theta, a=0) = \mathrm{v}(b) + \beta\left[\underline{\pi} E_{\theta'} V_e^{US}(\theta') + (1-\underline{\pi}) E_{\theta'} V_2^{US}(\theta')\right] \tag{61}$$

In the second period of non-employment, the agent remains eligible for benefits, but exhausts them if still non-employed moving forward. Thus, expected lifetime utility is given by:

$$V_2^{US}(\theta, a=1) = \mathrm{v}(b) - \theta\nu + \beta\left[\pi_u E_{\theta'} V_e^{US}(\theta') + (1-\pi_u) E_{\theta'} V_3^{US}(\theta')\right] \tag{62}$$

$$V_2^{US}(\theta, a=0) = \mathrm{v}(b) + \beta\left[\underline{\pi} E_{\theta'} V_e^{US}(\theta') + (1-\underline{\pi}) E_{\theta'} V_3^{US}(\theta')\right] \tag{63}$$

Finally, if non-employed for longer than two periods, benefits have expired and expected lifetime utility follows:

$$V_3^{US}(\theta, a=1) = \mathrm{v}(d) - \theta\nu + \beta\left[\pi_u E_{\theta'} V_e^{US}(\theta') + (1-\pi_u) E_{\theta'} V_3^{US}(\theta')\right] \tag{64}$$

$$V_3^{US}(\theta, a=0) = \mathrm{v}(d) + \beta\left[\underline{\pi} E_{\theta'} V_e^{US}(\theta') + (1-\underline{\pi}) E_{\theta'} V_3^{US}(\theta')\right] \tag{65}$$

where T denotes the lump-sum tax that finances the unemployment benefits, and $E_{\theta'}$ is the expectation over values of θ next period.

The calibration follows the baseline case in Section 5.3. The parameters μ, ν, and $\frac{\theta_h}{\theta_l}$ are chosen to ensure that a non-employed agent receiving the shock θ_l always prefers to be unemployed (i.e. exert search effort), while an agent receiving the shocks θ_h prefers to enter non-participation (i.e. not exert search effort). As in the baseline calibration, I find the lowest ratio of $\frac{\theta_h}{\theta_l}$ that delivers this feature. This implies that in the first period of non-employment exerting effort is preferred for an agent with θ_l, and in the third period of non-employment an agent with θ_h prefers not to exert effort. From the value function in (59)-(65) this implies the following two conditions (using the fact that $\mu\theta_h + (1-\mu)\theta_l = 1$):

$$\nu = \mu\beta(\pi_u - \underline{\pi})\left[E_\theta\left(V_e^{US} - V_3^{US}\right)\right] + (1-\mu)\beta(\pi_u - \underline{\pi})\left[E_\theta\left(V_e^{US} - V_2^{US}\right)\right] \tag{66}$$

$$\theta_h = \frac{1}{\mu}\left[1 - \frac{(1-\mu)\beta(\pi_u - \underline{\pi})\left[E_\theta\left(V_e^{US} - V_2^{US}\right)\right]}{\nu}\right] \tag{67}$$

Given the other parameters, these two equations imply $\frac{\theta_h}{\theta_l} = 1.32$. In the computations, I use a higher value, $\frac{\theta_h}{\theta_l} = 3.58$ to ensure that the planner indeed prefers to have only the θ_l report searching (the case I have analyzed analytically), and for comparability with the baseline analysis.

Finally, in this model, there is an additional parameter, $\underline{\pi}$, the probability of transitioning to employment if not exerting search effort. I set this to a low value, $\underline{\pi} = 0.001$ to allow comparisons with the baseline case where $\underline{\pi} = 0$. The parameters for this alternative specification and for the baseline comparison case are given in Table B.1

I also must specify the parameters of the U.S. system. In this comparison, I set the unemployment benefit, $b = 0.66y$ and $d = 0.25y$, which correspond to the values for the baseline case in the main text.

B.4.2 Welfare Comparisons

This section compares the welfare gains from adopting the optimal contract, relative to the current U.S. system described above, and the naïve planner's allocation. To calculate the welfare gains, in each specification I find the level of expected lifetime utility in the optimal contract, denoted w_P, such that $G_e(w_P) = 0$ (w_N for the naïve planner). I then compare w_P (w_N) with w_{US}.

Table B.2 presents the comparison of welfare gains. The second column shows that in both cases the

Table B.1: Parameters

Parameter	Baseline	Alternative
ν	0.98	0.96
μ	0.9790	0.021
$\frac{\theta_h}{\theta_l}$	3.49	3.58
$\underline{w}$	34%	35%

Notes: The first column lists the parameter, and the second and third columns display the value for the Baseline case and the Alternative, respectively. Further note that the value for μ is the same, but in the alternative specification the θ_h shock does not exert effort, while in the baseline case, the θ_l shock does not. In both cases, the value of μ is set to match the unemployment to non-participation transition probability. As in the main text, the value of the lower bound is a per-period consumption equivalent, as a fraction of the per-period wage: $\exp[(1-\beta)\underline{w}]/y$.

Table B.2: Welfare gains (in %), relative to current U.S. system

Specification	Total Gain	Naïve	Extra Gain
Baseline	1.94	1.06	0.87
Alternative	1.59	0.55	1.04

Notes: All welfare gains are in consumption equivalent terms. "Baseline" refers to the preference specification from the main text, while "Alternative" is the case presented in this section. "Total Gain" refers to the gains achieved by the optimal contract, relative to the U.S. system, while "Naïve" represents the gains from the naïve allocation, relative to the U.S. system. The last column, "Extra Gain" displays the additional gain from incorporating adverse selection.

optimal contract provides considerable gains over the current U.S. system, approximately 3% in consumption equivalent terms. The third column shows that the naïve allocation also provides large welfare gains, and the last column shows the additional gains from considering both adverse selection and moral hazard. The main conclusion from Table B.2 is that moving the taste shock to the utility cost of effort does not affect the size of the additional gains. In fact, the additional gains are actually larger in the case of the alternative preferences (with a smaller ratio of $\frac{\theta_h}{\theta_l}$).

In the cases considered here, it is optimal for the planner to recommend no effort from the agent reporting the relatively higher utility cost of effort.

B.4.3 Hopenhayn and Nicolini Version

In the optimal contract (either specification), some of the additional welfare gains arise because the planner achieves a more efficient allocation across taste shocks for *employed* agents. One may argue that such gains apply to a more general social insurance scheme, not unemployment insurance specifically. To analyze this dimension, in this section I consider a similar planning problem to the one in Hopenhayn and Nicolini (1997).

Specifically, I assume that once employed, agents no longer receive taste shocks, and they no longer face a decision to exert job-retention effort or not. They incur a utility cost to working, ν, but they do not have a choice to exert the effort or not. With exogenous probability $1-\pi_e$, they lose the job and transition to non-employment. The non-employment states are identical to the original model. The planner controls consumption of the non-employed as before. If the agent transitions to employment, the planner still sets $w_j^k(e)$; i.e., the planner determines the "employment tax." Once employed, however, the planner does not control w; it remains constant. This is an identical planning problem to Hopenhayn and Nicolini (1997) (except that here employment is stochastic, not permanent).

Table B.3: Welfare gains (in %): Hopenhayn and Nicolini Version

Specification	Total Gain	Naïve	Extra Gain
Baseline	2.45	0.52	1.92
Alternative	1.35	0.55	0.79

Notes: All welfare gains are in consumption equivalent terms. The first row presents the "Hopenhayn and Nicolini" case for the baseline specification of preferences, where the taste shock hits the marginal utility of consumption, and the second row presents the same case but for the alternative specification of preferences, where the taste shock hits the utility cost of effort directly.

Table B.3 presents the welfare gains for this case, for both the baseline specification of preferences, and the alternative. Surprisingly, in the baseline case, the welfare gains increase significantly in this version. This arises, I believe, because the costs of an employed agent decreases significantly. This occurs because the information problems disappear. For employed agents, the planner no longer has to deal with the information problems with respect to the taste shock nor with respect to job-retention effort. She simply offers constant consumption while employed. While indeed the planner loses the gains from smoothing consumption across taste shocks for the employed, these are trumped by the gains from perfect information (among employed agents). In the case of the alternative specification of preferences, the additional welfare gains from the optimal contract decrease slightly in the Hopenhayn and Nicolini case.

The experiments presented in this appendix suggest that the main results presented are robust to several different specifications of preferences and planning problems. Note, Table 6 in Section 6.4 of the main text summarizes the results from these experiments (both Table B.2 and Table B.3).

References

Hopenhayn, H., Nicolini, J., 1997. Optimal unemployment insurance. Journal of Political Economy 105, 412–438.