Efficient Allocations in Dynamic Private Information Economies with Persistent Shocks: A First Order Approach

Marek Kapicka

University of California, Santa Barbara

November 30, 2006

Abstract

I study efficient allocations in a dynamic private information economy with a continuum of individual shocks that are persistent. I formulate the problem recursively and develop a first order approach to simplify it. The main advantage of the first order approach is that it allows for a substantial reduction of the state space of the dynamic program. This makes the problem tractable and permits quantitative implementation of the problem.

I provide both qualitative and quantitative solutions for a taste shock economy where the shocks follow a random walk. I show that insurance against the shocks works very differently than in an otherwise identical economy with i.i.d. shocks. Both current and continuation utility are now positively correlated with the current shock and the social planner will optimally overinsure the agents, rather than underinsure. Also, for most of the population the intertemporal wedges are significantly larger than in an i.i.d. economy.

JEL classification: D80, D82, E61

Keywords: dynamic contracting, persistent shocks, private information

1 Introduction

This paper studies efficient allocations in a dynamic economy with continuum of idiosyncratic private information shocks that follow a Markov process. It is well known that persistent shocks significantly complicate the analysis of the problem. As shown by Fernandes and Phelan [8], the efficient allocations do have a recursive structure, but the dimensionality of the state space is proportional to the number of possible shock values. Quantitative analysis thus is feasible only for a small number of shocks.[1] This paper shows that when the first order approach is used, the state space can be reduced to a manageable dimension of three state variables, even though the shocks take a continuum of values. This greatly increases the tractability of the problem and the ability to investigate the efficient allocations quantitatively.

It is well known from Green [11], Thomas and Worrall [18] or Atkeson and Lucas [5] that when the shocks are i.i.d., the recursive formulation takes a very simple form. The whole reporting history of an agent up to any period t can be summarized by a single statistic, "promised utility", which is the lifetime utility the agent is entitled to receive from period t onwards. Based on the agent's report, the social planner chooses current consumption, current output and continuation utility from tomorrow on, such that the promised utility is delivered and it is in the agents' interest to report their shocks truthfully. The continuation utility chosen by the social planner becomes promised utility at the beginning of the next period, and so on. The crucial feature of this recursive formulation is that preferences over the allocations from the current period onwards are identical for all agents, regardless of whether they reported truthfully in the previous period or not.

This formulation breaks down when the shocks are persistent. When shocks are persistent, the probability distribution and hence preferences over the allocations

---

[1] See also Zhang [21] who studies a continuous time version of a private information economy with Markov shocks and is able to provide additional theoretical characterizations of the efficient allocations.

from the current period onward depend on the shock the agent experienced last period. Thus, the continuation utility chosen by the social planner depends both on the current report and on the current true shock. Since the social planner does not observe an agent's true shock, incentive compatibility requires that the continuation utility must be delivered next period for all possible true shocks in the last period. To satisfy this restriction, the state variable must itself be a *function* of the true shock in the previous period. This function, hereafter called the promised utility function, gives the promised utility of an agent for each possible true shock she might have had last period.

The main advantage of the first order approach is that it restricts attention to local deviations from the optimum. When the first order approach is applied, only the continuation utility of the truthtelling agent and its marginal change matter in determining whether an allocation is incentive compatible. As a consequence, I show that one can reduce the state space by replacing the promised utility function by only two values: promised utility for the truthtelling agent and a marginal change in her promised utility. Since one must still keep track of last period's shock, the state space reduces to three numbers. To apply the first order approach, I first show that the continuation utility function is differentiable in the agent's shock if the transition function is differentiable in the previous shock. Assuming this holds, the first order approach is correctly defined. I then find the necessary and sufficient conditions for the first order approach to hold.[2]

To analyze the efficient allocations, I study a simple economy with private taste shocks that follow a random walk with Pareto distributed innovations. I show that the pattern of insurance is very different from the pattern in an otherwise identical economy with i.i.d. shocks.In an economy with i.i.d. shocks, agents with high cur-

---

[2]The first order approach employed in this paper has also been recently used in related dynamic environments. Werning [20] and Abraham and Pavoni [1] use it to study dynamic hidden savings problem while Jarque [13] uses it to analyze moral hazard environment with effort persistence.

rent taste shocks are rewarded more consumption, but in order to ensure incentive compatibility, their continuation utility is reduced. With permanent shocks, this is no longer optimal: if the current shock is high, future shocks are expected to be high and so consumption in the future is expected to be high as well. Therefore, an agent with a high current taste shock receives *both* high current consumption and high continuation utility.

How are the agents then induced to report truthfully, rather than pretending they have higher shocks? I show that an appropriate choice of the marginal continuation utility is essential to ensure incentive compatibility and to provide insurance. The marginal continuation utility is shown to be positively related to the sensitivity of the agent's lifetime utility from tomorrow on to the next period shock. A lifetime utility profile that is very sensitive to next period's shock is beneficial for high shock agents because they expect to have a high shock tomorrow as well, but it is very costly for low shock agents, because they expect to continue having low shocks. High marginal continuation utility thus deters the low shock agents from reporting high shocks and helps to relax the incentive compatibility constraint. I show that the social planner will then optimally choose the marginal continuation utility to be "too high", compared to the first best allocation. I.e., agents' lifetime utility will be more sensitive to the current shock than it would be in a first best allocation, and the social planer overinsures the agents. This conclusion is very different from an i.i.d. economy, where the social planner typically underinsures the agents.

Recent research on dynamic optimal taxation[3] has focused attention on the analysis of the intertemporal wedges implied by the efficient allocations. The reason is that, in any implementation with taxes, the wedges are closely related to capital taxes. I therefore solve numerically for the efficient allocations and study both the efficient allocations and the intertemporal wedges. I define the ex-ante intertemporal

---

[3]See Werning [19], Golosov, Kocherlakota, Tsyvinski [9], Kocherlakota [16], Albanesi and Sleet [4], Albanesi [3] or Farhi and Werning [7], Kapicka [14] and others.

wedge as a wedge between the marginal utility of consumption today and the expected marginal utility of consumption tomorrow. I find that the wedge is mostly negative for low marginal promised utilities but turns positive for higher ones. For most of the marginal promised utilities and for most of the population, the ex-ante intertemporal wedge is significantly larger than it would have been in an i.i.d. economy. I also define an ex-post intertemporal wedge as a wedge between current marginal utility of consumption and the marginal utility of consumption tomorrow. If the well known inverse Euler equation were to hold, then the expected value of the ex-post intertemporal wedge would be zero. But the inverse Euler equation does not hold in the taste shock economy I study and the expected value of the ex-post intertemporal wedge is in general significantly different from zero. I decompose the ex-ante intertemporal wedge into the expected ex-post intertemporal wedge and the difference between the two, which is strictly positive due to Jensen's inequality. I find that the magnitude of the expected ex-post intertemporal wedge is the major reason why the ex-ante intertemporal wedge is so large. However, even the difference between these two is significantly larger than in an i.i.d. economy.

2 The Model

The world begins at time $t = 1$. The economy is populated by a continuum of agents of measure one. The agents are heterogeneous with respect to their productivity shock $\theta_t \in \Theta = [\underline{\theta}, \overline{\theta}] \subseteq \mathcal{R}_+$. The productivity shock is a continuous random variable that follows a first order Markov process. Let $\Pi : \Theta \times \mathcal{B}_\Theta \to [0, 1]$ be its transition function, where $\mathcal{B}_\Theta$ is the $\sigma-$algebra on Θ. It is assumed that Π is twice differentiable with respect to its first argument. Moreover, it is assumed that the transition function is such that $\left| \frac{\Pi_1(\theta, B)}{\Pi(\theta, B)} \right|$ is bounded for all $B \in \mathcal{B}_\Theta$. For each $t \geq 1$, one can construct a probability measure μ^t on $(\Theta^t, \mathcal{B}_\Theta^t)$ over the shock sequences $\theta^t \in \Theta^t$ that follow an initial given shock θ_0. Similarly, for $1 \leq j < t$ I construct a probability measure μ_j^t on

5

$(\Theta^{1+t-j}, \mathcal{B}_\Theta^{1+t-j})$ over the subsequences of shocks $\theta_j^t \in \Theta^{1+t-j}$, given the initial shock θ_0.

The agents are infinitely lived. Their preferences in each period are defined over consumption $c \in \mathcal{R}_+$, output $y \in \mathcal{R}_+$, and depend on the productivity shock $\theta \in \Theta$. Let $\mathcal{Y} \subseteq \mathcal{R}_+ \times \Theta$ be the set of all feasible output-productivity shock pairs.[4] The period utility function is given by $U : \mathcal{R}_+ \times \mathcal{Y} \to [\underline{u}, \bar{u}]$. It is assumed that the utility function is increasing in consumption and productivity shock, decreasing in output, concave in consumption, convex in output and twice differentiable with respect to all three arguments. Furthermore, the utility function is bounded, i.e. $-\infty < \underline{u} < \bar{u} < +\infty$. For technical reasons, I restrict the utility function to be linear in θ.[5] Agents discount the future by a discount factor $\beta \in (0, 1)$. Let $\mathcal{W} = [\frac{\underline{u}}{1-\beta}, \frac{\bar{u}}{1-\beta}]$ be the set of lifetime utilities the agent may receive.

At the beginning of each period, agents observe their current productivity shock. After that, consumption and output are determined. While both consumption and output are observed by the social planner, the productivity of the agent is her private information. The only exception is the initial shock θ_0, which is observed by the social planner and is assumed to be the same for everyone.

The social planner needs to finance a sequence of expenditures $\{\tilde{e}_t\}_{t=1}^\infty$, where $\tilde{e}_t \in \mathcal{R}$. He has access to a credit market and can freely borrow or save. The sequence of intertemporal prices of consumption is given by $q = \{q_t\}_{t=1}^\infty$ where $q_t > 0$ is a relative price of consumption between periods t and $t+1$. Let $e \in \mathcal{E} \subset \mathcal{R}$ be the present value of the expenditures:

$$e = \sum_{t=1}^\infty Q_t \tilde{e}_t,$$

where $Q_1 = 1$ and $Q_t = \prod_{i=1}^{t-1} q_i$ for $t > 1$ is the relative price of consumption goods

---

[4]This formulation allows for the possibility that not all output levels are feasible for a given productivity shock. For example, in a standard Mirrleesean economy where labor supply $l = \frac{y}{\theta}$ is between 0 and 1, $\mathcal{Y}$ is defined by the inequality $0 \le y \le \theta$.

[5]Two most common examples that satisfy this assumption is a taste shock specification without production ($U = \theta u(c)$) or with additively separable production ($U = u(c) - \theta v(y)$).

in period t in terms of consumption in period 1. To ensure that the present value of expenditures is finite, it is assumed that the prices satisfy a boundedness condition $\sum_{t=1}^{\infty} Q_t \leq +\infty$. In what follows I will take the sequence of prices as a parameter of the social planner's problem and the present value of government expenditures as an exogenous variable.

3   Constrained Pareto Optima

For a given present value of expenditures e and prices q, the social planner selects a pair of consumption assignments $C_t : \mathcal{E} \times \Theta^t \rightarrow \mathcal{R}_+$ and income assignments $Y_t : \mathcal{E} \times \Theta^t \rightarrow \mathcal{R}_+$ for each period $t \geq 1$. The collection of these assignments for all periods $Z = \{C_t, Y_t\}_{t \geq 1}$ is called an *allocation*.

At the beginning of period t, the agents report their current type to the social planner. The *reporting strategy* of an agent can be described by a collection of functions $\sigma = \{\sigma_t\}_{t \geq 1}$ where $\sigma_t : \Theta^t \rightarrow \Theta$ is a reporting strategy in period t. The set of all reporting strategies is denoted by Σ and the history of reports up to period t is denoted by $\hat{\theta}^t \in \Theta^t$.

The agent's continuation preferences over an allocation Z from period t on are given, for any history of reports $\hat{\theta}^{t-1} \in \Theta^{t-1}$ and any last period shock $\theta_{t-1} \in \Theta$, by

$$W_t(Z, \hat{\theta}^{t-1}, \theta_{t-1}) = \sum_{j=0}^{\infty} \int_{\Theta^{j+1}} \beta^j U[C_{t+j}(\hat{\theta}^{t-1}, \theta_t^{t+j}), Y_t(\hat{\theta}^{t-1}, \theta_t^{t+j}), \theta_{t+j}] \mu^j(\theta_{t-1}, d\theta_t^{t+j}).$$

Preferences from the initial period 1 on will be denoted simply as $W(Z, \theta_0) = W_1(Z, \oslash, \theta_0)$.

The allocation must satisfy two constraints. First, since the shocks are private information of an agent, the allocation must be such that the agent always prefers to report her shock truthfully. If the agent chooses reporting strategy σ she receives consumption and output assignment $Z \circ \sigma = \{C_t(\sigma^t(\theta^t)), Y_t(\sigma^t(\theta^t))\}_{t \geq 1}$. Thus, the allocation is *incentive compatible* if

$$W(Z, \theta_0) \geq W(Z \circ \sigma, \theta_0) \qquad \forall \sigma \in \Sigma. \tag{1}$$

7

Second, the allocation must be *feasible*, i.e. must collect enough resources to finance the present value of expenditures e. The budget constraint of the social planner is

$$\sum_{t=1}^{\infty} Q_t \int_{\Theta^t} [C_t(\theta^t) - Y_t(\theta^t)] \mu^t(\theta_0, d\theta^t) = e. \tag{2}$$

Denote the set of all allocations that are both incentive compatible and feasible for a present value of expenditures e and last period shock θ_0 by $\mathfrak{Z}^q(e, \theta_0)$.[6] The social planner's objective is to maximize the expected utility of an agent by choosing an allocation that is incentive compatible and feasible[7] for government expenditures e and last period shock θ_0 :

$$\Psi^q(e, \theta_0) = \max_{Z \in \mathfrak{Z}^q(e, \theta_0)} W(Z, \theta_0). \tag{3}$$

I will refer to this problem as a *sequence utility maximization* problem of the social planner.

The incentive compatibility constraint (1) requires that all possible reporting strategies $\sigma \in \Sigma$ must be weakly dominated by the truthful reporting strategy. It turns out, however, that it is not necessary to check whether all possible reporting strategies satisfy this. It is enough to check one period deviations are weakly dominated by truthtelling. The following Lemma, which extends a similar result of Fernandes and Phelan [8], shows this result formally.[8]

Lemma 1 *An allocation Z is incentive compatible if and only if for all $t \geq 1$, for*

---

[6]Note that the set $\mathfrak{Z}^q(E, \theta_0)$ may be empty for some E, if the present value of the expenditures is high enough so that no incentive compatible allocation can finance it.

[7]In principle, the social planner could assign different Pareto weights to different agents, but since all agents are ex ante identical, I will assume that they all have equal weights. This assumption simplifies the algebra considerably but can be easily relaxed.

[8]The proof of Lemma (1) can be found in the Appendix, as well as all the other proofs in this paper (unless mentioned otherwise).

almost all histories $\theta^{t-1} \in \Theta^{t-1}$ and almost all current period shocks $\theta_t \in \Theta$,

$$U(C_t(\theta^t), Y_t(\theta^t), \theta_t) + \beta W_{t+1}(Z, \theta^t, \theta_t) \geq U(C_t(\theta^{t-1}, \hat{\theta}), Y_t(\theta^{t-1}, \hat{\theta}), \theta_t) + \beta W_{t+1}(Z, (\theta^{t-1}, \hat{\theta}), \theta_t)$$

$$(4)$$

for all reports $\hat{\theta} \in \Theta$.

Lemma (1) shows that, although shocks are persistent in the model, there is no persistence in deviations. The incentive to tell the truth is unaffected by the agent's past behavior, in particular whether the agent reported truthfully in the previous periods. The reason why the temporary incentive compatibility constraint (4) holds is that although θ_{t-1} affects the preferences at the beginning of period t, the report is made after the current shock θ_t is realized. At this point, θ_{t-1} is no longer relevant: future preferences depend only on θ_t. Consequently, θ_{t-1} does not affect the incentive to tell the truth.

4   Dual Recursive Formulation

In this section I will study a dual problem where the social planner minimizes costs of delivering certain promised utility $w_1 \in \mathcal{W}$ and, in addition, I will formulate the dual problem recursively. In each period $t > 1$, the social planner's choice is conditioned on a function $\mathbf{w}(\cdot) : \Theta \to \mathcal{W}$, called a *promised utility function*. This function determines the agent's lifetime utility entitlement for each possible shock the agent may have received in the previous period: if last period's shock was $\hat{\theta}_-$, then the agent is entitled to receive lifetime utility $\mathbf{w}(\hat{\theta}_-)$. Even though the social planner does not observe the agent's true last period shock, he is now able to deliver the promised utility to any such agent, regardless of whether or not the agent has reported her shock truthfully.

I define an *allocation rule* $\tilde{z} = \{\tilde{c}_t, \tilde{y}_t, \tilde{w}_{t+1}\}_{t=1}^{\infty}$ as follows. Let $\mathcal{L}^{\Theta}$ be the space of functions $\mathbf{w} : \Theta \to \mathcal{W}$. Define, for each $t > 1$, a consumption function $\tilde{c}_t : \mathcal{L}^{\Theta} \times \Theta^2 \to$

$\mathcal{R}_+$, an output function $\tilde{y}_t : \mathcal{L}^\Theta \times \Theta^2 \to \mathcal{R}_+$ and a continuation utility function $\tilde{w}_{t+1} : \mathcal{L}^\Theta \times \Theta^3 \to \mathcal{W}$. The continuation utility function $\tilde{w}_{t+1}(\mathbf{w}(\cdot), \theta_-, \hat{\theta}, \theta)$ determines the continuation utility of an agent who is promised $\mathbf{w}(\cdot)$, reported θ_- in the previous period, reports $\hat{\theta}$ in the current period and has current shock θ. The consumption and output functions $\tilde{c}_t(\mathbf{w}(\cdot), \theta_-, \hat{\theta})$ and $\tilde{y}_t(\mathbf{w}(\cdot), \theta_-, \hat{\theta}))$ determine the consumption and output of such an agent, but unlike the continuation utility function, do not depend on the current shock θ. For $t = 1$, the definition is modified to take into account that, at the beginning of period 1, only the promised utility $w_1 \in \mathcal{W}$ must be delivered. Define therefore $\tilde{c}_1 : \mathcal{W} \times \Theta^2 \to \mathcal{R}_+$, $\tilde{y}_1 : \mathcal{W} \times \Theta^2 \to \mathcal{R}_+$ and $w_{t+1} : \mathcal{W} \times \Theta^3 \to \mathcal{W}$.

An allocation rule must satisfy three constraints: it must deliver the promised utility function, it must be incentive compatible and the continuation utility function must be consistent with future choices. The promise keeping constraint states that for all $t > 1$ the allocation rule must deliver lifetime utility $\mathbf{w}(\hat{\theta}_-)$ for an agent whose previous period shock was $\hat{\theta}_-$ and that this has to be true for all $\hat{\theta}_- \in \Theta$:[9]

$$\mathbf{w}(\cdot) = \int_\Theta [U(\tilde{c}_t(\theta), \tilde{y}_t(\theta), \theta) + \beta \tilde{w}_{t+1}(\theta, \theta)] \Pi(\cdot, d\theta). \tag{5}$$

This is the promise keeping constraint in the usual sense only when $\mathbf{w}(\cdot)$ is evaluated at the previous period shock θ_-. For any other values $\hat{\theta}_- \neq \theta_-$ it corresponds to the threat keeping constraint in the sense of Fernandes and Phelan [8] and determines what lifetime utility the agent would have been entitled to, were her previous period shock $\hat{\theta}_-$ but she still reported θ_-. The first period needs to be treated differently since the time zero shock θ_0 is publicly known. The promise keeping constraint only requires that promised utility w_1 must be delivered to the agent:

$$w_1 = \int_\Theta [U(\tilde{c}_1(\theta), \tilde{y}_1(\theta), \theta) + \beta \tilde{w}_2(\theta, \theta)] \Pi(\theta_0, d\theta). \tag{6}$$

The temporary incentive compatibility constraint ensures that it is in the agent's

---

[9]For notational convenience, the state variables $w(\cdot)$ and θ_- will be kept implicit whenever possible.

interest to report, in each period $t \geq 1$, her current type truthfully:

$$U(\tilde{c}_t(\theta), \tilde{y}_t(\theta), \theta) + \beta \tilde{w}_{t+1}(\theta, \theta) \geq U(\tilde{c}_t(\hat{\theta}), \tilde{y}_t(\hat{\theta}), \theta) + \beta \tilde{w}_{t+1}(\hat{\theta}, \theta) \qquad \forall \hat{\theta}, \theta \in \Theta^2. \quad (7)$$

To rule out the possibility that the social planner chooses a continuation utility function that will later be impossible to deliver, one needs to put an additional restriction on the social planner's set of choices: for all $t \geq 1$, it is required that

$$\tilde{w}_{t+1}(\theta, \cdot) \in \tilde{\mathcal{U}}^* \qquad \forall \theta \in \Theta, \quad (8)$$

where $\tilde{\mathcal{U}}^*$ is a fixed point of the following operator:

$$T\tilde{\mathcal{U}} = \{\mathbf{w} \in \mathcal{L}^\Theta : \exists \tilde{c} : \Theta \to \mathcal{R}_+, \tilde{y} : \Theta \to \mathcal{R}_+, \tilde{w}' : \Theta^2 \to \mathcal{R}$$
$$\text{such that (5) and (7) holds and } \tilde{w}'(\theta, \cdot) \in \tilde{\mathcal{U}} \forall \theta \in \Theta\}.$$

Standard arguments of Abreu, Pearce and Stacchetti [2] imply that $\tilde{\mathcal{U}}^*$ is nonempty, convex and compact. Denote the set of all allocation rules that satisfy, for a given w_1 and θ_0 the promise keeping constraints (6) and (5), the incentive compatibility constraint (7) and the constraint (8) by $\tilde{\mathfrak{z}}(w_1, \theta_0)$.

An allocation rule implies a sequence of distributions of the continuation utility functions for $t > 1$. Let D be a subset of $\mathcal{L}^\Theta$. The distribution of the promised utility functions at the beginning of period 2, $\tilde{\varphi}_2$ is given by

$$\tilde{\varphi}_2(D) = \int_{\tilde{M}_1(D)} \Pi(\theta_0, d\theta_1)$$

where $\tilde{M}_1(D) = \{\theta \in \Theta : \tilde{w}_2(w_1, \theta_0, \theta, \cdot) \in D\}$. Suppose now that $\tilde{\varphi}_t$ is the distribution of the promised utility functions at the beginning of period $t > 2$. The distribution next period $\tilde{\varphi}_{t+1}$ satisfies the difference equation

$$\tilde{\varphi}_{t+1}(D) = \int_{\tilde{M}_t(D)} \mu_{t-1}^t(\theta_0, d\theta_{t-1}^t) d\tilde{\varphi}_t \quad (9)$$

where $\tilde{M}_t(D) = \{(\mathbf{w}(\cdot), \theta_-, \theta) \in \mathcal{L}^\Theta \times \Theta^2 : \tilde{w}_{t+1}(\mathbf{w}(\cdot), \theta_-, \theta, \cdot) \in D\}$.[10]

---

[10]Note that, unlike the case of i.i.d. shocks, the evolution of the distribution always depends on the initial shock θ_0 and so cannot be written recursively, as in Atkeson and Lucas [5].

Any allocation rule $\tilde{z}$ implies certain present value of the social planner's resource costs. It is denoted by a function $\tilde{E}^q$ and satisfies

$$\tilde{E}^q(\tilde{z}, w_1, \theta_0) = \int_{\Theta} [\tilde{c}_1(w_1, \theta_0, \theta_1) - \tilde{y}_1(w_1, \theta_0, \theta_1)] \Pi(\theta_0, d\theta_1)$$
$$+ \sum_{t=2}^{\infty} Q_t \int_{\Theta^2 \times \mathcal{L}^{\Theta}} \{\tilde{c}_t(\mathbf{w}(\cdot), \theta_{t-1}, \theta_t) - \tilde{y}_t(\mathbf{w}(\cdot), \theta_{t-1}, \theta_t)\} d\mu_{t-1}^t d\tilde{\varphi}_t.$$

For any allocation rule $\tilde{z}$ define an allocation Z as follows. Let $e = \tilde{E}^q(\tilde{z}, w_1, \theta_0)$ be the implied costs of the allocation rule $\tilde{z}$. Let $\hat{W}_t(e, \hat{\theta}^{t-1}, \theta_{t-1})$ solve the following difference equation: $\hat{W}_1(e, \hat{\theta}_1, \theta_1) = \tilde{w}_2(w_1, \theta_0, \hat{\theta}_1, \theta_1)$ and for $t > 1$ $\hat{W}_{t+1}(e, \hat{\theta}^t, \theta_t) = \tilde{w}_{t+1}[\hat{W}_t(e, \hat{\theta}^{t-1}, \cdot), \hat{\theta}_{t-1}, \hat{\theta}_t, \theta_t]$. Set consumption $C_t(e, \hat{\theta}^t) = \tilde{c}_t[\hat{W}_t(e, \hat{\theta}^{t-1}, \cdot), \hat{\theta}_{t-1}, \hat{\theta}_t]$ and output $Y_t(e, \hat{\theta}^t) = \tilde{y}_t[\hat{W}_t(e, \hat{\theta}^{t-1}, \cdot), \hat{\theta}_{t-1}, \hat{\theta}_t]$. Call Z an allocation generated by the allocation rule $\tilde{z}$.

The relationship between allocations and allocation rules is shown in the next lemma.

Theorem 2 *i) Suppose that $\tilde{z}$ is an allocation rule and that $\tilde{z} \in \tilde{\mathfrak{z}}(w_1, \theta_0)$. Let $e = \tilde{E}^q(\tilde{z}, w_1, \theta_0)$ be its implied resource costs. Then the allocation Z that is generated by the allocation rule satisfies $Z \in \mathfrak{Z}^q(e, \theta_0)$ and delivers lifetime utility $W(Z, \theta_0) = w_1$.*

ii) Suppose that Z is an allocation and that $Z \in \mathfrak{Z}^q(e, \theta_0)$. Let $w_1 = W(Z, \theta_0)$ be the lifetime utility it delivers. Then there exists an allocation rule $\tilde{z} \in \tilde{\mathfrak{z}}(w_1, \theta_0)$ that has implied resource costs $\tilde{E}^q(\tilde{z}, w_1, \theta_0) = e$.

The proof of the theorem can be found in Appendix 2. It follows closely the proof of Lemmas 3.1 and 3.2 in Atkeson and Lucas [5] and can be found in the Appendix..

Theorem (14) implies that, instead of solving the social planner's problem (3), one can as well look at the problem of minimizing the resource costs subject to the constraint that the lifetime utility of the agent must be w_1 and then find w_1 such that the desired present value of expenditures is reached. The cost function $\Omega^q : \mathcal{W} \times \Theta \to \mathcal{R}$ is

$$\Omega^q(w_1, \theta_0) = \min_{\tilde{z} \in \tilde{\mathfrak{z}}(w_1, \theta_0)} \tilde{E}^q(\tilde{z}, w_1, \theta_0). \tag{10}$$

Let $\tilde{z}^*$ be the efficient allocation rule that solves this problem. Theorem (14) implies that the allocation Z^* that is generated by $\tilde{z}^*$ solves the sequence utility maximization problem (3) for government expenditures $e = \Omega^q(w_1, \theta_0)$.

The Bellman principle of optimality implies that the cost minimization problem (10) can be written recursively as follows. For $t \geq 1$ define a cost function $\Omega_t^q :$ $\mathcal{L}^\Theta \times \Theta \to \mathcal{R}$ recursively to be the minimal costs of delivering a promised utility function $\mathbf{w}(\cdot)$ to an agent who has reported a shock θ_- in period $t - 1$:

$$\Omega_t^q(\mathbf{w}(\cdot), \theta_-) = \min_{\tilde{c}, \tilde{y}, \tilde{w}'} \int_\Theta \{\tilde{c}(\theta) - \tilde{y}(\theta) + q_t \Omega_{t+1}^q(\tilde{w}'(\theta, \cdot), \theta)\} \Pi(\theta_-, d\theta), \qquad (11)$$

subject to the temporary incentive compatibility constraint (7), the promise keeping constraint (5) and the constraint (8). The solution to the Bellman equation (11) is related to the solution of the cost minimization problem (10) in the next Lemma, which is proved in the Appendix:

Lemma 3 *The cost function Ω^q satisfies, for all $w_1 \in \mathcal{W}$,*

$$\Omega^q(w_1, \theta_0) = \min_{\mathbf{w}(\cdot) \in \mathcal{L}^\Theta} \Omega_1^q(\mathbf{w}(\cdot), \theta_0) \qquad s.t. \quad \mathbf{w}(\theta_0) = w_1. \qquad (12)$$

5   First Order Approach

While the recursive approach of the previous section works well in theory, it is clear that having a function as a state variable prevents one not only from solving for the efficient allocations numerically, but also from gaining more insights into the efficient mechanism. To deal with this weakness, I now develop a first order approach to the problem. The advantage of the first order approach is that, as long as it is justified, one does not need to specify the continuation utility function $\tilde{w}_{t+1}$ for the off-equilibrium values. I show that this in turn implies that one does not need to specify the off-equilibrium values of the state variable as well. Rather, the recursive structure of the problem will depend only on the equilibrium promised utility $\mathbf{w}(\theta_-)$ of the truthtelling agent and on her marginal promised utility $\frac{\partial}{\partial \theta} \mathbf{w}(\theta_-)$, instead of on the whole promised

utility function $\mathbf{w}(\cdot)$. A corresponding recursive cost minimization problem will then depend only on a handful of simple state variables and can be numerically computed and analyzed.

The application of the first order approach is complicated by the fact that, as follows from (7), lifetime utility of the agent in the recursive cost minimization problem is a sum of two terms: $U(\tilde{c}_t(\hat{\theta}), \tilde{y}_t(\hat{\theta}), \theta)$ and $\tilde{w}_{t+1}(\hat{\theta}, \theta)$. The first term is just a period utility function and its properties of interest, namely differentiability with respect to θ, can be imposed exogenously. But $\tilde{w}_{t+1}(\hat{\theta}, \theta)$ is endogenous to the social planner's problem and its properties are not immediately known. Fortunately, I show that the necessary properties of $\tilde{w}_{t+1}(\hat{\theta}, \theta)$ can be determined quite easily. The reason is that, for a given report $\hat{\theta}$, the function $\tilde{w}_{t+1}(\hat{\theta}, \cdot)$ needs to satisfy (5) at the beginning of period $t + 1$. Hence, if the right-hand side of (5) is differentiable in θ, the left-hand side must be as well. But the right-hand side of (5) depends on θ only through the transition function Π. Hence, if the function $\Pi(\theta, \cdot)$ is differentiable in θ, which it is by assumption, then $\tilde{w}_{t+1}(\hat{\theta}, \theta)$ is also differentiable in θ. Put differently, the set $\tilde{\mathcal{U}}^*$ contains only functions that are differentiable in θ. Lemma (4) shows this result formally and also finds the bounds on $\frac{\partial}{\partial \theta} \tilde{w}'(\hat{\theta}, \theta)$.

Lemma 4 *Suppose that $\tilde{w}'(\hat{\theta}, \cdot) \in \tilde{\mathcal{U}}^*$ for all $\hat{\theta} \in \Theta$. Then $\tilde{w}'(\hat{\theta}, \theta)$ is differentiable with respect to θ for all $\hat{\theta} \in \Theta$. Moreover, its derivative $\frac{\partial}{\partial \theta} \tilde{w}'(\hat{\theta}, \theta)$ is bounded for all $(\hat{\theta}, \theta) \in \Theta^2$.*

The next theorem builds on Lemma (4) and derives necessary and sufficient conditions for an allocation rule to be incentive compatible. The proof is similar in many aspects to the one used in static economies, e.g. in Mirrlees [17].

Theorem 5 *An allocation rule $\tilde{z}$ satisfies the temporary incentive compatibility constraint (7) for $t \geq 1$ if and only if*

$$U(\tilde{c}_t(\theta), \tilde{y}_t(\theta), \theta) + \beta \tilde{w}_{t+1}(\theta, \theta) = \int_{\underline{\theta}}^{\theta} [U_\theta(\tilde{c}_t(\varepsilon), \tilde{y}_t(\varepsilon), \varepsilon) + \beta \frac{\partial}{\partial \theta} \tilde{w}_{t+1}(\varepsilon, \varepsilon)] d\varepsilon + \tilde{w}_0, \quad (13)$$

and $U_\theta(\tilde{c}_t(\hat{\theta}), \tilde{y}_t(\hat{\theta}), \theta) + \beta \frac{\partial}{\partial \theta} \tilde{w}_{t+1}(\hat{\theta}, \theta)$ is increasing in $\hat{\theta}$ for all θ.

14

The envelope condition (13) is obtained by applying the Envelope Theorem to the incentive compatibility constraint (7). It shows how lifetime utility must vary with the current period shock in order to be incentive compatible. The expression under the integral consists of two terms. The first term is the current informational rent: the extra current period utility an agent receives from a marginal increase in current skills. The second term corresponds to the future informational rent: extra lifetime utility from tomorrow on that an agent receives from a marginal shift in the future distribution of shocks. The second term is zero if shocks are i.i.d. because current private information bears no information about future shocks.

The global second order condition requires $U_\theta(\tilde{c}_t(\hat{\theta}), \tilde{y}_t(\hat{\theta}), \theta) + \beta \frac{\partial}{\partial \theta} \tilde{w}_{t+1}(\hat{\theta}, \theta)$ to be increasing in $\hat{\theta}$ for all θ. If the shocks are i.i.d., the second term drops out and this condition is equivalent to a simple requirement that $\tilde{y}_t(\theta)$ is increasing in θ, provided that the utility function satisfies the Spence-Mirrlees condition. When the shocks are persistent, monotonicity of $\tilde{y}_t$ is neither required nor implied. In principle, one can have an incentive compatible allocation rule such that $\tilde{y}_t(\theta)$ is decreasing, if $\frac{\partial}{\partial \theta} \tilde{w}_{t+1}$ increases sufficiently fast in $\hat{\theta}$.

In what follows, I will adopt the usual approach by considering a *relaxed social planner's problem* where the monotonicity condition is ignored. The monotonicity condition will only be checked after the solution to the relaxed social planner's problem is found, with the hope that it will not bind.

The envelope condition (13) shows that everything except for the continuation utility and the marginal continuation utility is irrelevant for determining whether an allocation rule is incentive compatible. This suggests that everything else may be also irrelevant for the recursive formulation itself. With that in mind, I define a *modified allocation rule* $z = \{c_t, y_t, w_{t+1}, g_{t+1}\}_{t=1}^\infty$ as follows. Let $\mathcal{G} \subseteq \mathcal{R}$ be the set of all feasible marginal promised utilities.[11] For each $t > 1$, the modified allocation rule consists of the functions $c_t : \mathcal{W} \times \mathcal{G} \times \Theta^2 \to \mathcal{R}_+$, $y_t : \mathcal{W} \times \mathcal{G} \times \Theta^2 \to \mathcal{R}_+$,

---

[11]It follows from Lemma (5) that one can take $\mathcal{G}$ to be a bounded subset or $\mathcal{R}$.

$w_{t+1} : \mathcal{W} \times \mathcal{G} \times \Theta^2 \to \mathcal{W}$ and $g_{t+1} : \mathcal{W} \times \mathcal{G} \times \Theta^2 \to \mathcal{G}$. For $t = 1$ it is defined by $c_1 : \mathcal{W} \times \Theta^2 \to \mathcal{R}_+$, $y_t : \mathcal{W} \times \Theta^2 \to \mathcal{R}_+$, $w_{t+1} : \mathcal{W} \times \Theta^2 \to \mathcal{W}$ and $g_{t+1} : \mathcal{W} \times \Theta^2 \to \mathcal{G}$. The function w_{t+1} now represents the continuation utility of the truthtelling agent, while g_{t+1} represents the marginal change (with respect to the true shock) of the same agent. Compared to the allocation rule, the domain of the modified allocation rule has been modified for $t > 1$ and the promised utility function $\mathbf{w}(\cdot)$ has been replaced by the promised utility w and the marginal promised utility g.

The modified allocation rule z is required to satisfy four conditions. The first is the envelope condition: For all $t \geq 1$,

$$U(c_t(\theta), y_t(\theta), \theta) + \beta w_{t+1}(\theta) = \int_{\underline{\theta}}^{\theta} [U_\theta(c_t(\varepsilon), y_t(\varepsilon), \varepsilon) + \beta g_{t+1}(\varepsilon)] d\varepsilon + w_0, \qquad (14)$$

which is equation (13) applied to the modified allocation rule. The second requirement is that for all $t \geq 1$, the promise keeping constraint must hold:

$$w = \int_\Theta [U(c_t(\theta), y_t(\theta), \theta) + \beta w_{t+1}(\theta)] \Pi(\theta_-, d\theta) \qquad (15)$$

where, unlike constraint (5), it is required to hold only for the last period shock θ_-.

The third constraint, called the *marginal promise keeping constraint,* is new. For all $t > 1$, the social planner to restricted to increase the marginal continuation value of an agent with last period shock θ_- at the rate g:

$$g = \int_\Theta [U(c_t(\theta), y_t(\theta), \theta) + \beta w_{t+1}(\theta)] \Pi_1(\theta_-, d\theta). \qquad (16)$$

Finally, the last requirement is that for all $t > 1$ the modified allocation rule must be such that next period promised utility $w_{t+1}(\theta)$ and marginal promised utility $g_{t+1}(\theta)$ is delivered:

$$[w_{t+1}(\theta), g_{t+1}(\theta)] \in \mathcal{U}^*(\theta) \qquad \forall \theta \in \Theta, \qquad (17)$$

where $\mathcal{U}^*(\theta) \subseteq \mathcal{W} \times \mathcal{G}$ is the fixed point of the following operator:

$$T\mathcal{U}(\theta_-) = \{(w, g) \in \mathcal{W} \times \mathcal{G} : \exists c : \Theta \to \mathcal{R}_+, y : \Theta \to \mathcal{R}_+, w' : \Theta \to \mathcal{W}, g' : \Theta \to \mathcal{G}$$
$$\text{such that (15), (16) and (14) holds and } [w'(\theta), g'(\theta)] \in \mathcal{U}(\theta)\}.$$

The set $\mathcal{U}^*(\theta)$ can again be shown to be nonempty, compact and convex for all $\theta \in \Theta$. The set of all allocation rules that satisfy, for given w_1 and θ_0 the promise keeping constraint (15), the marginal promise keeping constraint (16), the incentive compatibility constraint (14) and the constraint (17) will be denoted by $\mathfrak{z}(w_1, \theta_0)$.

One may be tempted to prove now that an analogue of Theorem 14 holds and that the set of modified allocation rules is in some sense identical to the set of allocation rules or allocations. Such an effort would be unsuccessful. As long as the first order approach is valid, the set of modified allocation rules $\mathfrak{z}(w_1, \theta_0)$ will be strictly smaller than the set of allocation rules $\tilde{\mathfrak{z}}(w_1, \theta_0)$: for each $z \in \mathfrak{z}(w_1, \theta_0)$ one can construct $\tilde{z} \in \tilde{\mathfrak{z}}(w_1, \theta_0)$ that delivers the same consumption and output assignment after each history of reports. The reverse, however, is not true. The reason is that the first order approach is unable to provide arbitrary restrictions off the equilibrium path. But I will show in theorem (6) that these restrictions can only increase the costs to the social planner and will therefore never be used in the optimum. Hence the recursive formulation relying on the first order approach is sufficient to find the efficient allocation.[12]

For these reasons, I will now focus on the cost minimization problem of the social planner directly. Define $V_t^q : \mathcal{W} \times \mathcal{G} \times \Theta \to \mathcal{R}$ recursively by

$$V_t^q(w, g, \theta_-) = \min_{c,y,w',g'} \int_\Theta [c(\theta) - y(\theta) + q_t V_{t+1}^q(w'(\theta), g'(\theta), \theta)]\Pi(\theta_-, d\theta), \qquad (18)$$

subject to (15), (16), (14) and (17). It can then be shown that the smallest present value of the costs of delivering promised utility w_1 to an agent with last period shock θ_0 is given by the following function $V^q : \mathcal{W} \times \Theta$:

$$V^q(w_1, \theta_0) = \min_{g \in \mathcal{G}} V_1^q(w_1, g, \theta_0) \qquad s.t. \quad (w_1, g) \in \mathcal{U}^*(\theta_0). \qquad (19)$$

---

[12]Thus, there is an asymmetry between a first order approach in static models and a first order approach in this model. In static models there are no off-equilibrium allocations and so the first order approach does not limit the set of available allocations in any way.

The next theorem relates the solution to the social planner's problem (19) to the solution of the problem (12) and justifies the recursive formulation of this section.

Theorem 6 *Suppose that the relaxed social planner's problem is valid. Then $V^q = \Omega^q$.*

The intuition behind the proof is the following. I first define a set $\Upsilon(w, g, \theta_-) \subseteq \tilde{\mathcal{U}}^*$ that contains all the promised utility functions $\mathbf{w}(\cdot)$ that deliver promised utility $\mathbf{w}(\theta_-) = w$ and marginal promised utility $\frac{\partial}{\partial \theta_-}\mathbf{w}(\theta_-) = g$ to an agent with the last period shock θ_-. I first show that the value function $V_t^q(w, g, \theta_-)$ attains the minimum of $\Omega_t^q(\mathbf{w}(\cdot), \theta_-)$ over the set $\Upsilon(w, g, \theta_-)$. That is, as long as the first order approach is valid, constraining the social planner only by w and g does no harm. This in turn implies that $V^q \leq \Omega^q$. On the other hand, the first period choice of the promised utility function is constrained only by a requirement that $\mathbf{w}(\theta_0) = w_1$ and so the social planner can attain the minimum of $\Omega_1^q(\mathbf{w}(\cdot), \theta_-)$ over $\Upsilon(w_1, g, \theta_0)$. Hence $\Omega^q = V^q$.

6   Taste Shock Economy with Random Walk

I will now consider a simple economy with multiplicative shocks, labeled as taste shocks, and logarithmic utility. The period utility is thus given by $U(c, y, \theta) = \theta \ln c$.

Since the utility is logarithmic, the assumption that the utility function is bounded will not apply and $\mathcal{W} = \mathcal{R}$. I will study an economy where the taste shocks follow a random walk and where the innovations are Pareto distributed. Let $\eta > 1$ be the scale parameter. If θ_- is the last period shock then the location parameter of the Pareto distribution is given by $\theta_- \theta_L$. In order to ensure that $E(\theta|\theta_-) = \theta_-$ and the shocks follow a random walk, I set $\theta_L = \frac{\eta-1}{\eta}$. Hence the density function $\pi(\theta_-, \theta)$ is given by $\pi(\theta_-, \theta) = \frac{\eta}{\theta}(\frac{\theta}{\theta_- \theta_L})^{-\eta}$ if $\theta \geq \theta_- \theta_L$ and zero otherwise. The intertemporal price is set to be equal to the discount factor in all periods: $q_t = \beta$ for all $t \geq 0$.[13]

---

[13]Because the price of consumption is constant, the efficient modified allocation rule will be time independent. I will therefore drop the time subscript and denote next period with prime.

Instead of working with the consumption allocation directly, I follow the usual change of variables and define current period utility $u = \ln c$.[14] With this change, the incentive compatibility constraint (14) takes the form

$$\theta u(\theta) + \beta w'(\theta) = \int_{\theta - \theta_L}^{\theta} [u(\varepsilon) + \beta g'(\varepsilon)]d\varepsilon + w_0, \tag{20}$$

the promise keeping constraint becomes

$$w = \int_{\theta - \theta_L}^{\infty} [\theta u(\theta) + \beta w'(\theta)]\pi(\theta_-, \theta)d\theta, \tag{21}$$

and the marginal threat keeping constraint can be written as

$$g = \int_0^{\infty} [\theta u(\theta) + \beta w'(\theta)]\Pi_1(\theta_-, d\theta). \tag{22}$$

One can significantly simplify the marginal promise keeping constraint by using the fact that with the Pareto distribution $\pi_1(\theta_-, \theta) = \frac{\eta}{\theta_-}\pi(\theta_-, \theta)$. Rewriting the right-hand side of (22), one sees that the marginal threat keeping constraint holds if and only if

$$g = \frac{\eta}{\theta_-}(w - w_0). \tag{23}$$

This version of the marginal threat keeping constraint gives an additional interpretation to marginal promised utility g. Marginal promised utility is proportional to the spread between the expected utility of the agent w and the utility of the agent with the lowest possible shock, w_0. High marginal promised utility indicates that the variation in the lifetime utilities of agents (after the shock has been revealed) will be large. Furthermore, since w_0 can be uniquely determined from the knowledge of w, g and θ_-, the social planner's commitment to deliver a given marginal promised utility g can alternatively be seen as a commitment to obey a lower bound on agent utilities w_0. The equation (23) also implies that $\mathcal{G} = \mathcal{R}$ and that $\mathcal{U}^*(\theta) = \mathcal{R}^2$ for all θ and so the requirement (17) is never binding.

---

[14]It is customary to call u the utility, although the period utility is in fact θu.

The efficient modified allocation rule can now be found as a solution to the following Bellman equation:[15]

$$V^q(w, g, \theta_-) = \min_{u, w', g'} \int_{\theta_- \theta_L}^{\infty} [e^{u(\theta)} + \beta V^q(w'(\theta), g'(\theta), \theta)] \pi(\theta_-, \theta) d\theta, \qquad (24)$$

subject to the incentive compatibility constraint (20), the promise keeping constraint (21) and the marginal threat keeping constraint (23). Since all the constraints are homogeneous of degree one in the choice variables and the period return function is strictly convex, the value function $V^q(w, g, \theta_-)$ is strictly convex in w, g.

6.1   Two Normalizations

It turns out that with permanent shocks and log utility, the value function and the optimal policy functions take a very simple functional form in both the last period shock θ_- and promised utility w. One can use these functional forms to simplify the dynamic program (24) and to transform it into a simpler one that involves only one state variable: marginal promised utility g. I will first show the normalization of the last period shock θ_-.

Lemma 7 *The value function satisfies $V^q(w, g, \theta_-) = V^q(\frac{w}{\theta_-}, g, 1)$. The optimal policy functions satisfy $u(w, g, \theta_-, \theta) = u(\frac{w}{\theta_-}, g, 1, \frac{\theta}{\theta_-})$, $w'(w, g, \theta_-, \theta) = \theta_- w'(\frac{w}{\theta_-}, g, 1, \frac{\theta}{\theta_-})$ and $g'(w, g, \theta_-, \theta) = g'(\frac{w}{\theta_-}, g, 1, \frac{\theta}{\theta_-})$.*

The current utility u and marginal continuation utility g' for an agent with last period shock θ_- can be obtained by proportionally stretching (if $\theta_- > 1$) or compressing (if $\theta_- < 1$) the current utility u and marginal continuation utility g' of an agent with the last period shock 1 and a rescaled promised utility $\frac{w}{\theta_-}$. In addition, the continuation utility is proportionally scaled up or down by θ_-. The intuition for why the optimal policy functions take these special functional forms can be obtained

20

by looking at the properties of the shock distribution. Because the shocks follow a random walk with Pareto distributed innovations, the distribution of shocks for an agent with last period shock θ_- look like a stretched or compressed distribution of shocks for an agent with last period shock 1. Since the shocks are permanent, the agent's utility of consumption in all the future periods (and hence the continuation utility) is expected to be scaled up or down by θ_- as well. Lemma (7) shows that the optimal policy functions inherit all these properties.

The second normalization involves promised utility w.

Lemma 8 *The value function satisfies $V^q(w, g, 1) = e^{(1-\beta)w} V^q(0, g - w, 1)$. The optimal policy functions satisfy $u(w, g, 1, \theta) = (1-\beta)w + u(0, g-w, 1, \theta)$, $w'(w, g, 1, \theta) = \theta w + w'(0, g - w, 1, \theta)$ and $g'(w, g, 1, \theta) = w + g'(0, g - w, 1, \theta)$.*

If shocks were i.i.d., one could show that the value function is exponential in w and the optimal policy functions are linear in w. Lemma (8) shows that this result partially carries over to the economy with permanent shocks. There are, however, two differences. First, one needs to adjust the marginal promised utility g by w as well. Second, the additive term in the continuation utility function is θw, rather than only w. An important implication of Lemma (8) is that the value function is not necessarily increasing in the promised utility. The intuition is that a given marginal promised utility g may be easily obtainable for some levels of promised utility, but be very costly to deliver for some other levels of promised utility. A decrease in the promised utility may thus impose additional costs on the social planner and if these costs are large enough, the value function will increase.

Define the normalized value function $v(g) = V^q(0, g, 1)$ and the normalized current utility function $r(g, \theta) = u(0, g, 1, \theta)$, continuation utility function $h'(g, \theta) = \frac{w'(0,g,1,\theta)}{\theta}$ and the marginal continuation utility function $m'(g, \theta) = g'(0, g, 1, \theta) - \frac{w'(0,g,1,\theta)}{\theta}$. Note that h' is normalized in such a way that the direct effect of the current taste shock on h' is zero. Let also $\pi(\theta) \sim \pi(1, \theta)$ be simplified notation for the normalized density

function. Using Lemmas (7) and (8), the dynamic program (24) has been reduced to the following:

$$v(g) = \min_{r,h',m'} \int_{\theta_L}^{\infty} [e^{r(\theta)} + \beta e^{(1-\beta)h'(\theta)} v(m'(\theta))] \pi(\theta) d\theta \tag{25}$$

s.t.

$$0 = \int_{\theta_L}^{\infty} \theta[r(\theta) + \beta h'(\theta)] \pi(\theta) d\theta \tag{26}$$

$$\theta[r(\theta) + \beta h'(\theta)] = \int_{\theta_L}^{\theta} [r(\varepsilon) + \beta h'(\varepsilon) + \beta m'(\varepsilon)] d\varepsilon - \frac{g}{\eta}, \tag{27}$$

where the marginal threat keeping constraint (and a choice of w_0) has been eliminated completely by substituting $-\frac{g}{\eta}$ for w_0 in the incentive compatibility constraint.

I will denote the cost minimizing marginal promised utility by $g^* = \arg\min_{\mathcal{G}} v(g)$. One may conjecture that the minimum will be interior and that the value function v will be U-shaped, decreasing for $g \leq g^*$ and increasing for $g \geq g^*$. The cost minimizing marginal promised utility g^* is significant for two reasons. First, the social planner will choose g^* in the first period when he is not constrained by the marginal promise keeping constraint. Second, I will later show that g^* is also a lower bound on the relevant domain of the social planner's problem: provided that the social planner starts with $g \geq g^*$, he will always choose $m' \geq g^*$.

6.2 Properties of Efficient Allocations

The constraints (26) and (27) differ from their original counterparts (21) and (20) in one important respect: both current and continuation utility are multiplied by the current shock θ. This is clearly a consequence of the way the optimal policy functions were normalized in Lemma (7) and in principle says nothing about the optimal allocations themselves. But it indicates that insurance against taste shocks works in a dramatically different way. In an economy with i.i.d. shocks the social planner provides insurance by giving relatively high current utility to agents with high shocks, in exchange for a lower continuation utility. Such a trade-off preserves

22

incentive compatibility, because low shock agents benefit very little from high current consumption but are hit as hard as anyone else by the lower continuation utility. When the shocks are permanent, this trade-off is considerably weakened. A high current shock means that the agents expect to receive high shocks in the future as well - as if the shock directly affected both the current and continuation utility. It is therefore likely that they will receive relatively high consumption in the future as well. The efficient allocation will in such case give the agents *both* high current utility and high continuation utility.

To ensure incentive compatibility, the social planner now introduces a new trade-off, in the form of varying marginal promised utility m' or, equivalently, in the form of a varying lower bound on next period lifetime utilities. To see how this trade-off works, consider first the following Lemma that simplifies the incentive compatibility constraint (27):

Lemma 9 *The policy functions r, h' and m' satisfy the incentive compatibility constraint if and only if*

$$r(\theta) + \beta h'(\theta) = \beta \int_{\theta_L}^{\theta} m'(\varepsilon) \frac{d\varepsilon}{\varepsilon} - \frac{g}{\eta - 1}. \tag{28}$$

A simple proof, relying on the differentiability of the optimal policy functions in θ can be easily obtained by differentiating (27), cancelling terms and integrating back. A proof that does not assume differentiability is also available, but is omitted here. The equation (28) shows that the marginal continuation utility is proportional to the slope of agent's lifetime utility profile $r + \beta h'$. If the social planner wants to reward higher shock agents with higher lifetime utility, he needs to choose a high marginal continuation utility profile. Indeed, if $m' = 0$ for all agents, then the only incentive compatible lifetime utility profile is independent of the shock, which is the autarchic utility profile. Any other lifetime utility profile will not be incentive compatible and some agents will deviate by reporting higher skills than they truly have.

The reason why a high marginal continuation utility profile allows the social planner to choose a steeper lifetime utility profile and provide more insurance is the following. High marginal continuation utility implies, as can be seen from the marginal promise keeping constraint (23), that next period there will be a large spread between the lower bound on lifetime utility and average utility. As a consequence, the lifetime utility profile next period will be steep as well. Such a lifetime utility profile is very harmful for a low shock agent who considers falsely reporting a high shock today. The reason is that she expects that her shock values next period will be low. Therefore her continuation utility next period is also expected to be low. This in turn implies that reporting high shock is costly and so deters her from doing so.

Since the social planner must simultaneously fulfill both (26) and (28), he is not completely free to choose the marginal continuation utilities. For instance, choosing $m' = 0$ for all agents is feasible only for $g = 0$, otherwise the promise keeping constraint (26) is violated. One can substitute (28) into (26) and integrate to show that the function m' must satisfy

$$g = \beta \int_{\theta_L}^{\infty} \theta m'(\theta)\pi(\theta)d\theta. \tag{29}$$

That is, the expected value of $\theta m'$ is equal to $\frac{g}{\beta} > g$. This implies that there must be some degree of persistence in the choice of marginal continuation utilities: if g is high, the function m' will also have high values. At least some fraction of agents will be assigned $m' > g$ and, unless there is a sufficiently strong negative covariance between θ and m', one can expect an *upward* drift in the marginal continuation utility. A negative covariance between θ and m' will be a likely feature of the solution at least for $g \geq g^*$ since for such values the value function is increasing in g and it is especially costly to increase the marginal continuation utilities further.

The equation (28) also shows that the global second order conditions now take a very simple form. They are equivalent to a requirement that $m' \geq 0$ for all θ, which in turn implies that $r + \beta h'$ is increasing in θ. If the second order condition binds for some range of taste shocks, then $r + \beta h'$ will be constant in that range and no

insurance against the taste shocks will be provided. I will show in the next section that the second order condition never binds in a relevant range of the domain.

6.2.1   First Order Conditions

Let $\lambda(g)$ be the Lagrange Multiplier on the promise keeping constraint and $\gamma(g, \theta)$ be the Lagrange Multiplier on the incentive compatibility constraint. The first order conditions in r and h' yield the following necessary condition for the optimum:

$$r = (1 - \beta)h' + \ln[(1 - \beta)v(m')]. \tag{30}$$

This equation determines the division of resources between the current utility r and the continuation utility h'. The relative costs of delivering current and continuation utility depend on the choice of m'. In particular, if m' moves further away from g^*, the relative costs of delivering continuation utility increase and more emphasis is put on current utility. Equation (30) underlies the intuition that with permanent shocks, the trade-off between delivering r versus h' is not key to understanding the efficient allocations. The trade-off between r and h' depend on the current shock only indirectly through the choice of m', and if m' were constant across types, r would differ from h' only by a constant.

Define now

$$x = \frac{1}{\theta} e^{(1-\beta)h'} v(m'). \tag{31}$$

Assuming that the value function is differentiable, it is easy to show that the first order condition in m' combined with (30), yields the following equation in x and m':

$$\pi \theta^2 \epsilon_v(m')x = \int_\theta^\infty \pi \varepsilon [\lambda - (1 - \beta)x]d\varepsilon, \tag{32}$$

where $\epsilon_v(g) = \frac{v_g(g)}{v(g)}$ is the the semielasticity of the value function. Equations (30) and (32), together with the envelope condition (28) and definition (31) are four equations in four functions u, h', x and m' that need to be solved for the efficient allocations.

25

Put $t = \ln \theta$ and rewrite equations (30), (32) and definition (31) as differential equations in t. By eliminating r and h', one sees that the functions m' and x satisfy

$$\frac{dm'}{dt} = \frac{x(1-\beta) - \lambda - x\epsilon_v(m')[(1-\beta)\beta m' - \eta]}{x[\epsilon_v'(m') + \beta\epsilon_v^2(m')]} \tag{33}$$

$$\frac{dx}{dt} = x[(1-\beta)\beta m' + \beta\epsilon_v(g')\frac{dm'}{dt} - 1]. \tag{34}$$

The next lemma establishes the boundary conditions for the differential equations (33) and (34) and finds the minimum of the value function.

Lemma 10

i) For all $g \in \mathcal{G}$, $\lim_{\theta \to \infty} m'(g, \theta) = g^*$ and $\lim_{\theta \to \infty} x(g, \theta) = \frac{\lambda(g)}{1-\beta}$.

ii) The value function $v(g)$ is minimized at $g^* = \frac{1}{\beta - \beta^2}$.

The first part of Lemma (10) shows that, in the limit, the cost minimizing marginal continuation utility is always chosen, independently of the promised utility g. This is essentially the well known "no distortion at the top" result: for the agent with the highest possible shock (here represented by the limiting agent), there is no distortion in the choice of marginal continuation utility. But there is another implication of the result. If $g \neq g^*$, the solution to the social planner's problem is not time consistent. At the beginning of the period, the social planner would like to default on his promise to deliver the marginal promised utility g. Delivering the marginal promised utility g^* decreases the costs but the promised utility of the truthtelling agent is unchanged. Lemma (10) then shows that there is no time consistency problem for the limiting agent.

I will now show that the social planner will always operate in the increasing part of the value function and, unless he starts with $g < g^*$ (which he never will), he will always choose the marginal continuation utility to be greater or equal to g^*.

Lemma 11 *Suppose that $g \geq g^*$. Then $m'(g, \theta) \geq g^*$ for all $\theta \geq \theta_L$.*

Lemma (11) has two important implications. First, it implies that the social planner will have a tendency to *overinsure* the agents. One can show that the marginal continuation utility chosen in the first best allocation is equal to $\frac{1}{\beta-\beta^2} = g^*$. Thus, the marginal continuation utility is always greater than the first best one and, by (28), the lifetime utility profile $r + \beta h'$ will respond more strongly to the current shock than it would respond in a first best allocation. In contrast, an otherwise identical economy with i.i.d. shocks will typically exhibit underinsurance: the lifetime utility profile will be less responsive to the current shocks than the first best allocation would dictate. The distortions from the optimum thus exhibit a very different pattern. But the intuition for why there is overinsurance is fairly straightforward. As discussed in previous section, making the lifetime utility profile more responsive to current shocks deters lower skilled agents from reporting higher types and therefore weakens the incentive compatibility constraint. A second implication of Lemma (11) is that the second order condition will never bind for $g \geq g^*$ and one can safely ignore it.

It is shown in the course of the proof of Lemma (11) that $m'(g^*, \theta_L) = g^*$. Since g^* is the marginal promised utility chosen in the first period, there is thus no time consistency problem for the agent with the lowest shock as well, as long as the agent always had the lowest shock in the past. Note that time consistency at the bottom is much weaker than time consistency at the top: the efficient allocation is time consistent for any limiting agent, regardless of the past history, but only for the lowest shock agent who has always had the lowest shock in the past.

6.2.2 Wedges

This section studies intertemporal wedges in this economy. I define an *ex-ante intertemporal wedge* $\Delta(g, \theta)$ as a percentage wedge between the expected marginal utility of consumption tomorrow and the marginal utility of consumption today:

$$\frac{\theta}{c(\theta)} = [1 - \Delta(g, \theta)]E(\frac{\theta'}{c'(\theta')}|\theta).$$

where $c'(\theta') = c[w'(\theta), g'(\theta), \theta, \theta']$. It is worth noting that, since the taste shock enters multiplicatively, the inverse Euler equation will not hold. Hence there is no reason to expect that the intertemporal wedge will be always positive.

I will also define an *ex-post intertemporal wedge* $\delta(g, \theta, \theta')$ as a wedge between marginal utility of consumption today and ex-post marginal utility of consumption tomorrow:

$$\frac{\theta}{c(\theta)} = [1 - \delta(g, \theta, \theta')]\frac{\theta'}{c'(\theta')}.$$

The importance of the ex-post intertemporal wedge is that it is likely that, in any decentralization of the efficient allocation with taxes, the ex-post intertemporal wedge would correspond to capital income taxes. Since the inverse Euler equation fails to hold, the expected ex-post intertemporal wedge is in general not zero.

Jensen's inequality implies that $\Delta(g, \theta) > E(\delta(g, \theta, \theta')|\theta)$. The difference between the ex-ante intertemporal wedge and the expected ex-post intertemporal wedge is of interest because it helps to separate two forces that constitute the ex ante intertemporal wedge. The ex-ante intertemporal wedge equals to the ex-post intertemporal wedge $E(\delta(g, \theta, \theta')|\theta)$, which can be positive or negative, plus the difference between these two, $\Delta(g, \theta) - E(\delta(g, \theta, \theta')|\theta)$, which is always positive. If the inverse Euler equation were to hold, the expected ex-post intertemporal wedge would be zero and the ex-ante intertemporal wedge would be solely determined by the second force. In the next section I will quantitatively determine the relative importance of both forces.

7 Numerical simulations

In this section I quantitatively investigate the efficient allocations in the taste shock economy with random walk. I set $\beta = 0.96$ and $\eta = 2$. For the numerical exercise I truncate the upper 0.0001% of the Pareto distribution and discretize the space of shocks with 1000 gridpoints. I solve the differential equation (33) and iterate on the

value function (25) until it converges.[16]

Figure 1 studies the issue of overinsurance. It plots the current utility function $r(g, \cdot)$ for several selected high values of g against the logarithm of the shock.[17] I contrast the results with the current utility function in an otherwise identical economy with i.i.d. shocks. In a first best allocation the slope of all these functions would be one and so if the policy function is over the 45° line there is overinsurance, while if it is under the 45° degree line there is underinsurance. The picture shows that overinsurance becomes significant especially in the middle range of shocks. On the other hand, insurance in the i.i.d. economy is almost perfect for low shock values but agents become more underinsured when their shock increases.

Figure 2 shows the continuation utility function $h'(g, \cdot)$ for selected values of g and contrasts it again with the continuation utility function in an i.i.d. economy. One can see that the continuation utility function goes in exactly the opposite direction: it increases with shocks when the shocks are persistent but decreases when the shocks are i.i.d. Moreover, h' is much more sensitive to the current shock than the continuation utility in an i.i.d. economy.[18]

The ex-ante intertemporal wedge Δ is depicted in figure 3, together with the the intertemporal wedge when shocks are i.i.d. One can see that the magnitude of the wedge depends critically on the marginal promised utility. For low values of marginal promised utility it is mostly negative but becomes positive for higher values of g. The

---

[16]Unfortunately, the nature of the problem prevents me from imposing an upper bound on the the space of marginal promised utilities. As follows from (11), for any g and any feasible contract, there must be some set of agents with $m' > g$. I solve this problem by extrapolating the value function beyond a certain value of $\bar{g}$.

[17]All the functions are normalized to have the same intercept.

[18]Note that h' is continuation utility normalized by θ. Thus, the actual continuation utility is $\theta h'$, which is even more sensitive to θ. But it is more appropriate to compare h' and the continuation utility in an i.i.d. economy, because h' is cleared from any "mechanic" shifts resulting purely from a shift in the distribution of shocks. Variations in h' thus represent only the economic forces in play, as do variations in the continuation utility in an i.i.d. economy.
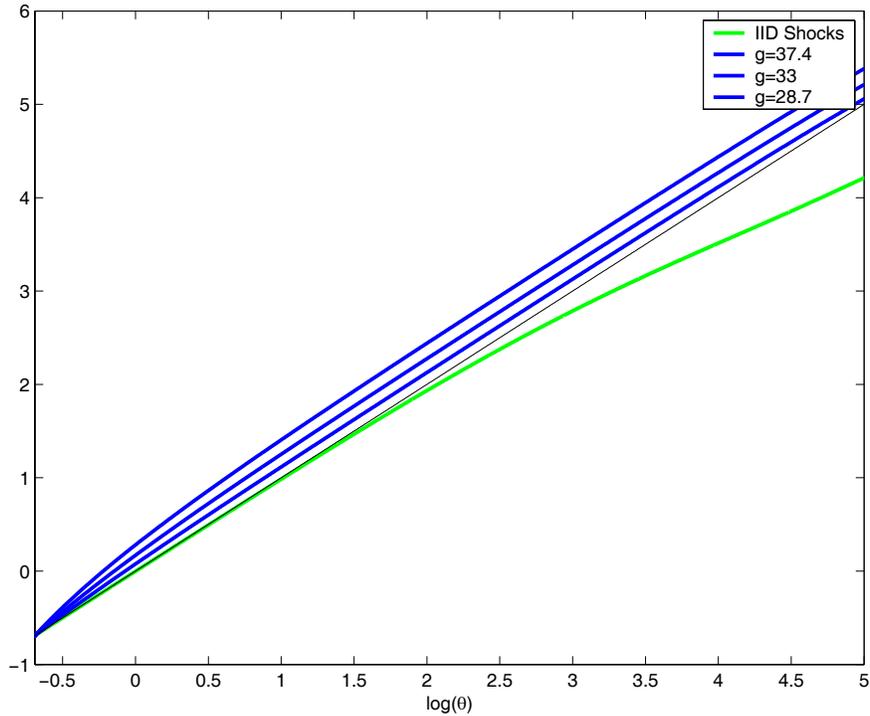
Figure 1: Current Utility

wedge tends to decrease with the shock. This is a very different pattern than one can observe in an i.i.d. economy, where the wedge is very small for low shocks, but becomes significant for very high shocks.

Figure 4 then addresses the question which of the two forces described in the previous section contribute more to the ex-ante intertemporal wedge. It plots both the ex-ante intertemporal wedge and the ex-post intertemporal wedge for $g = g^*$. It shows that the distortions in the inverse Euler equation are mainly responsible for the large magnitudes of the ex-ante intertemporal wedge.[19] However, even the difference between these two is quite sizable, at least when compared to the same difference in an i.i.d. economy where it is almost negligible. Thus, the force stemming from Jensen's inequality is stronger when shocks are permanent. This conclusion is consistent with the findings of Zhang [21], who reports that the ex-ante intertemporal wedge in a

---

[19]Similar result can be found for any other g.
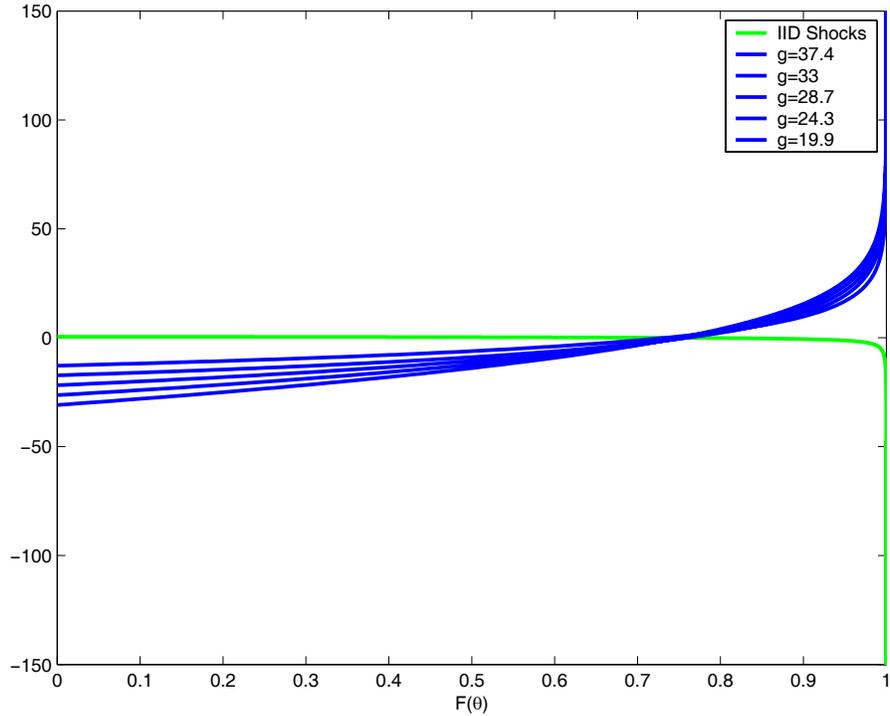
Figure 2: Continuation Utility

Mirrleesean economy where the inverse Euler equation holds becomes more significant once the shocks are permanent.

8 Conclusions

The contribution of this paper is twofold. First, it develops a method of solving dynamic private information models with persistent shocks that is simple enough to be solved numerically, even for a continuum of shocks. Second, I show that the pattern of efficient allocations in a taste shock economy with shocks that follow a random walk is very different from the pattern of efficient allocations in an otherwise identical i.i.d. economy. In particular, I show that the continuation utility increases in the current shock, rather than decreases, and that the efficient allocations involve too much insurance, rather than too little.
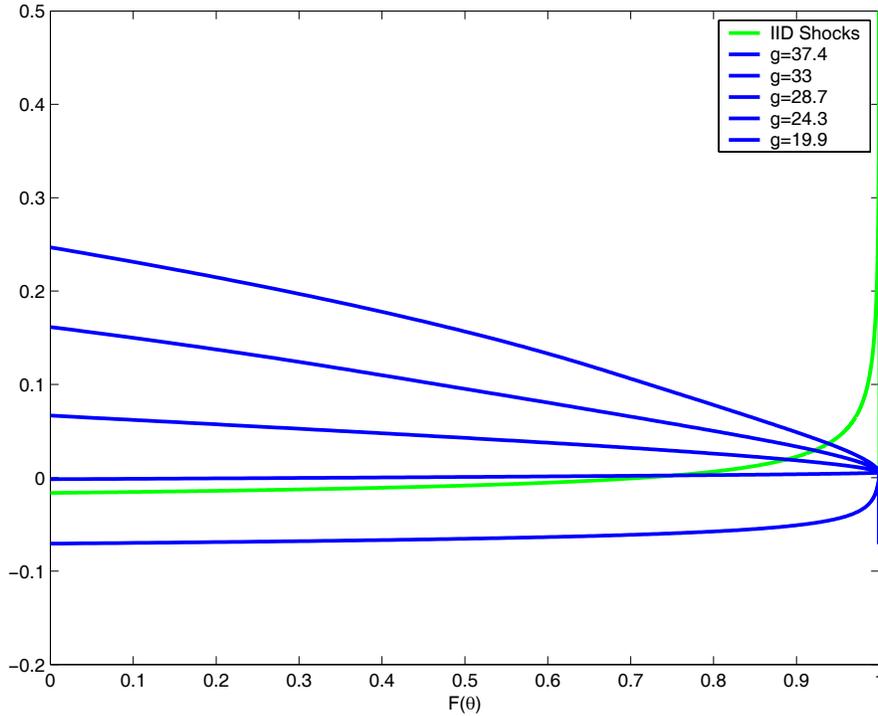
Figure 3: Ex Ante Intertemporal Wedge

While the taste shock economy is a very special one, it is likely that insurance will work in a similar way in more general Mirrleesean economies with persistent productivity shocks and endogenous labor supply. In particular, the continuation utility is still likely to be increasing in the current shock, because the future informational rent will also be positive. If the incentive compatibility can be weakened by increasing the slope of the lifetime utility, then the social planner will still overinsure the agents. This feature may be troubling if one wants to take these models as a positive foundation of market frictions. But, obviously, all these conclusions may be overturned if one assumes less than full persistence of the shocks. More quantitative studies are needed before any of the conclusions of this paper can be generalized in any direction. The first order approach developed in this paper makes such studies tractable.
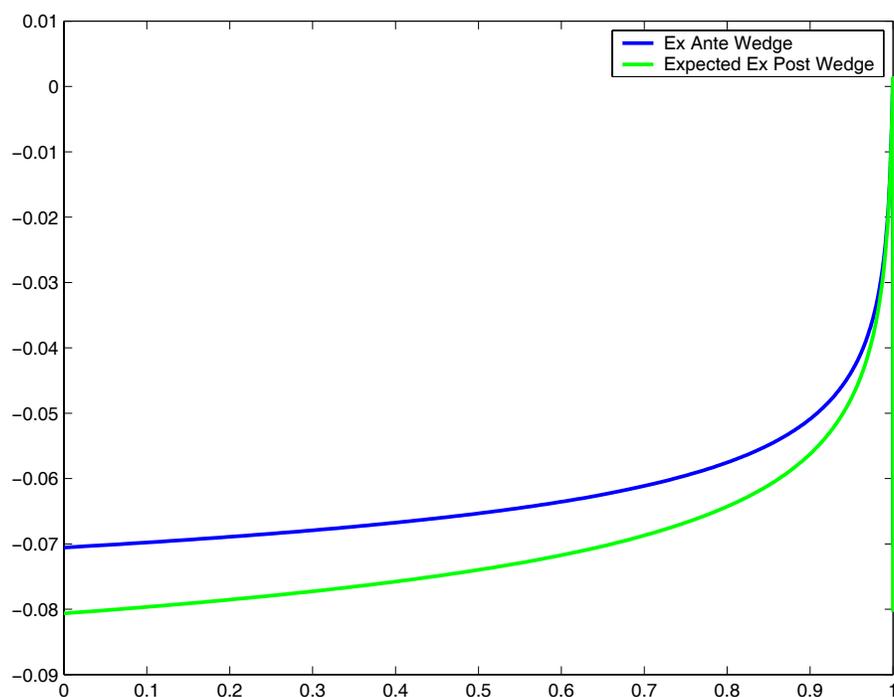
Figure 4: Ex Post and Ex Ante Intertemporal Wedge

References

[1] Abraham, A. and N. Pavoni (2006): "Efficient Allocations with Moral Hazard and Hidden Borrowing and Lending", unpublished manuscript, University of Rochester.

[2] Abreu, D., D. Pearce and E. Stacchetti (1990): " Toward a Theory of Discounted Repeated Games with Imperfect Information", Econometrica **58**, 1041-1063.

[3] Albanesi, S. (2006): "Optimal Taxation of Entrepreneurial Capital with Private Information", unpublished manuscript, Columbia University.

[4] Albanesi, S. and Ch. Sleet (2006): "Dynamic optimal Taxation with Private Information", The Review of Economic Studies **73**, 1-30.

[5] Atkeson, A. and R. E. Lucas, Jr. (1992): "On Efficient Distribution With Private Information", The Review of Economic Studies **59**, 427-453.

[6] Doepke, M. and R. M. Townsend (2006): "Dynamic Mechanism Design with Hidden Income and Hidden Actions", Journal of Economic Theory **126**, 235-286.

[7] Farhi, E. and I. Werning (2006): "Inequality and Social Discounting", unpublished manuscript, MIT.

[8] Fernandes, A. and Ch. Phelan (2000): "A Recursive Formulation for Repeated Agency with History Dependence" Journal of Economic Theory **91**, 223-247.

[9] Golosov, M., N. Kocherlakota and A. Tsyvinski (2003): "Optimal Indirect and Capital Taxation", The Review of Economic Studies **70**, 569-588

[10] Golosov, M. and A. Tsyvinski (2007): "Designing Optimal Disability Insurance", Quarterly Journal of Economics, forthcoming.

[11] Green, E. J. (1987): "Lending and the Smoothing of Uninsurable Income". In Edward C. Prescott and Neil Wallace, editors, Contractual Arrangements for Intertemporal Trade, 3-25. University of Minnesota Press, Minnenapolis.

[12] Heathcote, J., K. Storesletten and G. L. Violante (2004): "The Macroeconomic Implications of Rising Wage Inequality in the United States", unpublished manuscript, NYU.

[13] Jarque, A. (2002): "Repeated Moral Hazard with Effort Persistence", unpublished manuscript, University of Rochester.

[14] Kapicka, M. (2006): "The Dynamics of Optimal Taxation when Human Capital is Endogenous", unpublished manuscript, UC Santa Barbara.

[15] Kapicka, M. (2006): "Efficient Allocations with Persistent Shocks", working paper, UC Santa Barbara.

[16] Kocherlakota, N. (2005): "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation", Econometrica **73,** 1587-1621.

[17] Mirrlees, J. A. (1986): "The Theory of Optimal Taxation", in K.J. Arrow and M.D. Intriligator, eds., Handbook of Mathematical Economics, vol. III, chapter 24.

[18] Thomas, J. and T. Worrall (1990): "Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem". Journal of Economic Theory, 51, 367-90.

[19] Werning, I. (2001): "Dynamic Optimal Taxation", unpublished manuscript, University of Chicago.

[20] Werning, I. (2001): "Moral Hazard with Unobserved Endowments: A Recursive Approach", unpublished manuscript, University of Chicago.

[21] Zhang, Y. (2006): "Dynamic Contracting, Persistent Shocks and Optimal Taxation", unpublished manuscript, University of Minnesota.

9 Appendix

Proof of Lemma (1). Take any history of reports $\theta^{t-1} \in \Theta^{t-1}$ and any last period shock $\theta_{t-1} \in \Theta$. For any reporting strategy σ such that $\sigma^{t-1}(\theta^{t-1}) = \hat{\theta}^{t-1}$ the lifetime utility $W_t(Z \circ \sigma, \hat{\theta}^{t-1}, \theta_{t-1})$ is well defined and one can show that it can be written recursively as follows:

$$
\begin{aligned}
W_t(Z \circ \sigma, \hat{\theta}^{t-1}, \theta_{t-1}) &= \int_\Theta \{ U[C_t[\hat{\theta}^{t-1}, \sigma_t(\theta^t)], Y_t[\hat{\theta}^{t-1}, \sigma_t(\theta^t)], \theta_t] \\
&\quad + \beta W_{t+1}[Z \circ \sigma, (\hat{\theta}^{t-1}, \sigma_t(\theta^t)), \theta_t] \} \Pi(\theta_{t-1}, d\theta_t). \quad (35)
\end{aligned}
$$

I will show necessity and sufficiency of (4).

(a) Necessity of (4). The proof is by contradiction. Suppose there is some period t and sets $D \subseteq \Theta$ and $\hat{D} \subseteq \Theta^{t-1}$, both of nonzero measure such that for all $\theta_t \in D$, all $\theta^{t-1} \in \hat{D}$ the temporary incentive compatibility constraint (4) does not hold. That is, there is some function $\tilde{\sigma}_t : D \times \hat{D} \to \Theta$ such that for all $\theta_t \in D$, all $\theta^{t-1} \in \hat{D}$,

$$
\begin{aligned}
&U[C_t(\theta^t), Y_t(\theta^t), \theta_t] + \beta W_{t+1}(Z, (\theta^t), \theta_t) \\
< \quad &U[C_t(\theta^{t-1}, \tilde{\sigma}_t(\theta^{t-1}, \theta_t)), Y_t(\theta^{t-1}, \tilde{\sigma}_t(\theta^{t-1}, \theta_t)), \theta_t] + \beta W_{t+1}[Z, (\theta^{t-1}, \tilde{\sigma}_t(\theta^{t-1}, \theta_t)), \theta_t].
\end{aligned}
$$

Define a new reporting strategy σ as follows. For $j \neq t$, $j \geq 1$ define $\sigma_j(\theta^j) = \theta_j$ for all $\theta^j \in \Theta^j$. For $j = t$ and $\theta^t \in \hat{D} \times D$ define $\sigma_t(\theta^t) = \tilde{\sigma}_t(\theta^t)$ and for $\theta^t \notin \hat{D} \times D$ define $\sigma_t(\theta^t) = \theta_t$. I will now show that σ dominates truthtelling. For any $\theta^{t-1} \in \hat{D}$,

$$
\begin{aligned}
W_t(Z, \theta^{t-1}, \theta_{t-1}) &= \int_D \{ U[C_t(\theta^t), Y_t(\theta^t), \theta_t] + \beta W_{t+1}[Z, (\theta^t), \theta_t] \} \Pi(\theta_{t-1}, d\theta_t) \\
&\quad + \int_{D^c} \{ U[C_t(\theta^t), Y_t(\theta^t), \theta_t] + \beta W_{t+1}[Z, (\theta^t), \theta_t] \} \Pi(\theta_{t-1}, d\theta_t) \\
&< \int_\Theta \{ U[C_t[\theta^{t-1}, \sigma_t(\theta^{t-1}, \theta_t)], Y_t[\theta^{t-1}, \sigma_t(\theta^{t-1}, \theta_t)], \theta_t] \\
&\quad + \beta W_{t+1}[Z, (\theta^{t-1}, \sigma_t(\theta^{t-1}, \theta_t)), \theta_t] \} \Pi(\theta_{t-1}, d\theta_t) \\
&= W_t(Z \circ \sigma, \theta^{t-1}, \theta_{t-1}). \quad (36)
\end{aligned}
$$

The inequality follows from the fact that the integral over D is strictly smaller on the left hand side by assumption and the integral over D^c is equal since $\sigma_t(\theta^{t-1}, \theta_t) =$

36

θ_t for $\theta_t \in D^c$. The last equality then follows from (35) and from the fact that, by construction, $\sigma^{t-1}(\theta^{t-1}) = \theta^{t-1}$ for all $\theta^{t-1} \in \Theta^{t-1}$. But (36) implies that

$$
\begin{aligned}
W(Z, \theta_0) &= \sum_{j=1}^{t-1} \int_{\Theta^j} \beta^{j-1} U[C_j(\theta^j), Y_j(\theta^j), \theta_j] d\mu^j \\
&\quad + \beta^{t-1} \left\{ \int_{\hat{D}} W_t[Z, \theta^{t-1}, \theta_{t-1}] d\mu^{t-1} + \int_{\hat{D}^c} W_t[Z, \theta^{t-1}, \theta_{t-1}] d\mu^{t-1} \right\} \\
&< \sum_{j=1}^{t-1} \int_{\Theta^j} \beta^{j-1} U[C_j(\theta^j), Y_j(\theta^j), \theta_j] d\mu^j \\
&\quad + \beta^{t-1} \left\{ \int_{\hat{D}} W_t[Z \circ \sigma, \theta^{t-1}, \theta_{t-1}] d\mu^{t-1} + \int_{\hat{D}^c} W_t[Z \circ \sigma, \theta^{t-1}, \theta_{t-1}] d\mu^{t-1} \right\} \\
&= W(Z \circ \sigma, \theta_0),
\end{aligned}
$$

where the inequality follows from the fact that the integral over $\hat{D}$ is strictly smaller on the left hand side by (36) while the integral over $\hat{D}^c$ is identical since $\sigma(\theta^{t-1}) = \theta_{t-1}$ for $\theta^{t-1} \in \hat{D}$. Thus, σ dominates truthtelling, a contradiction.

(b) Sufficiency of (4). The proof is by induction. Consider any function $\tilde{\sigma}_1 : \Theta \to \Theta$. Equations (1) and (35) imply that

$$
\begin{aligned}
W(Z, \theta_0) &= \int_{\Theta} \{ U[C_1(\theta_1), Y_1(\theta_1), \theta_1] + \beta W_2[Z, \theta_1, \theta_1] \} \Pi(\theta_0, d\theta_1) \\
&\geq \int_{\Theta} \{ U[C_1(\tilde{\sigma}_1(\theta_1)), Y_1(\tilde{\sigma}_1(\theta_1)), \theta_1] + \beta W_2[Z, (\tilde{\sigma}_1(\theta_1)), \theta_1] \} \Pi(\theta_0, d\theta_1) \\
&= W(Z \circ \sigma^{(1)}, \theta_0),
\end{aligned}
$$

where σ^1, defined by $\sigma_1^{(1)}(\theta_1) = \tilde{\sigma}_1(\theta_1)$ and $\sigma_j^{(1)}(\theta^j) = \theta_j$ for $j \geq 2$ for all $\theta^j \in \Theta^j$.is a strategy involving only first period deviations.

Assume now that, for any $t \geq 1$, $W(Z, \theta_0) \geq W(Z \circ \sigma^{(t-1)}, \theta_0)$,where $\sigma^{(t-1)}$ is a reporting strategy involving only deviations in periods $1, 2, ..t - 1$. Let $\hat{\theta}^{t-1} = \sigma^{(t-1)(t-1)}(\theta^{t-1})$ by the history of reports in the first $t - 1$ period for some shock history $\theta^{t-1} \in \Theta^{t-1}$. The temporary incentive compatibility constraint (4) applied in

37

period t implies that, for any function $\tilde{\sigma}_t : \Theta \to \Theta$

$$
\begin{aligned}
& W_t(Z, \hat{\theta}^{t-1}, \theta_{t-1}) \\
= & \int_\Theta \{U[C_t(\hat{\theta}^{t-1}, \theta_t), Y_t(\hat{\theta}^{t-1}, \theta_t), \theta_t] + \beta W_{t+1}[Z, (\hat{\theta}^{t-1}, \theta_t), \theta_t]\}\Pi(\theta_{t-1}, d\theta_t)\} \\
\geq & \int_\Theta \{U[C_t(\hat{\theta}^{t-1}, \tilde{\sigma}_t(\theta_t)), Y_t(\hat{\theta}^{t-1}, \tilde{\sigma}_t(\theta_t)), \theta_t] + \beta W_{t+1}[Z, (\hat{\theta}^{t-1}, \tilde{\sigma}_t(\theta_t)), \theta_t]\}\Pi(\theta_{t-1}, d\theta_t)\} \\
= & W_t(Z \circ \sigma^{(t)}, \hat{\theta}^{t-1}, \theta_{t-1}). \tag{37}
\end{aligned}
$$

By expanding the expression for $W(Z \circ \sigma^{(t-1)}, \theta_0)$ one gets

$$
\begin{aligned}
W(Z \circ \sigma^{(t-1)}, \theta_0) = & \sum_{j=1}^{t-1} \int_{\Theta^j} \beta^{j-1} U[C_j(\sigma^{(t-1)j}(\theta^j)), Y_j(\sigma^{(t-1)j}(\theta^j)), \theta_j]d\mu^j \\
& + \beta^{t-1} \int_{\Theta^{t-1}} W_t[Z, \sigma^{(t-1)(t-1)}(\theta^{t-1}), \theta_t]d\mu^{t-1} \\
\geq & \sum_{j=1}^{t-1} \int_{\Theta^j} \beta^{j-1} U[C(\sigma^{(t-1)j}(\theta^j)), Y_j(\sigma^{(t-1)j}(\theta^j)), \theta_j]d\mu^j \\
& + \beta^{t-1} \int_{\Theta^{t-1}} W_t[Z \circ \sigma^{(t)}, (\sigma^{(t)(t-1)}(\theta^{t-1})), \theta_{t-1}]d\mu^j \\
= & W(Z \circ \sigma^{(t)}, \theta_0),
\end{aligned}
$$

where the inequality follows from application of (37). Hence $W(Z, \theta_0) \geq W(Z \circ \sigma^{(t)}, \theta_0)$ and incentive compatibility holds for any reporting strategies involving deviations in the first t periods only. To prove that any strategy $\sigma^{(\infty)} = \sigma$ involving infinite period deviations is also suboptimal, suppose on the contrary there is such strategy $\sigma^{(\infty)}$ that dominates truthtelling. Then there must be a period T such that the gains in periods 1 to T also dominate truthtelling, since the utility is bounded and so the gains after period T can be made arbitrarily small. But this means that a reporting strategy $\sigma^{(T)}$ dominates truthtelling, which is a contradiction. ∎

Proof of Lemma (3). Define $\bar{\Omega}^q(w_1, \theta_0) = \min_{\mathbf{w}(\cdot) \in \mathcal{L}^\Theta} \Omega_1^q[\mathbf{w}(\cdot), \theta_0]$ subject to $w(\theta_0) = w_1$. It will be shown that $\bar{\Omega}^q = \Omega^q$. Define an allocation rule $\bar{z}$ as follows. Let $\{\tilde{c}_1, \tilde{y}_1, \tilde{w}_1'\}$ solve (11) for $t = 1$ and $\bar{w}(w_1, \cdot)$ solves (12). Then define consumption $\bar{c}_1(w_1, \theta_0, \theta_1) = \tilde{c}_1(\bar{w}(w_1, \cdot), \theta_0, \theta_1)$, output $\bar{y}_1(w_1, \theta_0, \theta_1) = \tilde{y}_1(\bar{w}(w_1, \cdot), \theta_0, \theta_1)$ and

continuation utility function $\bar{w}_1'(w_1, \theta_0, \theta_1, \cdot) = \tilde{w}_1'(\bar{w}(w_1, \cdot), \theta_0, \theta_1, \cdot)$. For $t > 1$, let $\bar{z}$ solves the Bellman equation (11).

It is easy to see that $\tilde{E}^q(\bar{z}, w_1, \theta_0) \geq \Omega^q(w_1, \theta_0)$ because $\bar{z}$ satisfies (5) for $t > 1$, (6) for $t = 1$, and (7) and (8) for all time periods. To show that $\tilde{E}^q(\bar{z}, w_1, \theta_0) \leq \tilde{E}^q(\bar{z}^*, w_1, \theta_0) = \Omega^q(w_1, \theta_0)$ note that

$$\Omega_t^q[\mathbf{w}(\cdot), \theta_-] = \int_\Theta \{\bar{c}_t(\theta) - \bar{y}_t(\theta) + q_t \Omega_{t+1}^q[\bar{w}_t'(\theta, \cdot), \theta]\} \Pi(\theta_-, d\theta) \qquad (38)$$
$$\leq \int_\Theta \{\tilde{c}_t^*(\theta) - \tilde{y}_t^*(\theta) + q_t \Omega_{t+1}^q[\tilde{w}_t^{*\prime}(\theta, \cdot), \theta]\} \Pi(\theta_-, d\theta)$$

for any $t > 1$, since $\{\tilde{c}_t^*, \tilde{y}_t^*, \tilde{w}_t^{*\prime}\}$ that satisfy (7), (5) and (8) but not necessarily attain the minimum of (11). Similarly, for $t = 1$,

$$\bar{\Omega}^q(w_1, \theta_0) = \int_\Theta \{\bar{c}_t(\theta) - \bar{y}_t(\theta) + q_1 \Omega_2^q[\bar{w}_1'(\theta, \cdot), \theta]\} \Pi(\theta_-, d\theta)$$
$$\leq \int_\Theta \{\tilde{c}_1^*(\theta) - \tilde{y}_1^*(\theta) + q_1 \Omega_2^q[\tilde{w}_1^{*\prime}(\theta, \cdot), \theta]\} \Pi(\theta_-, d\theta).$$

By recursively substituting Ω_t^q out of the right-hand side using (38), one gets that

$$\tilde{E}^q(\bar{z}, w_1, \theta_0) = \bar{\Omega}^q(w_1, \theta_0) \leq \tilde{E}^q(\bar{z}^*, w_1, \theta_0) = \Omega^q(w_1, \theta_0).$$

Hence $\bar{\Omega}^q(w_1, \theta_0) = \tilde{E}^q(\bar{z}, w_1, \theta_0) = \Omega^q(w_1, \theta_0)$. ∎

Proof of Lemma (4). By definition of the set $\tilde{\mathcal{U}}^*$, for any function $\tilde{w}' \in \tilde{\mathcal{U}}^*$ there must exist some functions $\tilde{c} : \Theta \to \mathcal{R}_+, \tilde{y} : \Theta \to \mathcal{R}_+, \tilde{w}'' : \Theta^2 \to \mathcal{R}$, such that for any $\hat{\theta} \in \Theta$,

$$\tilde{w}'(\hat{\theta}, \theta) = \int_\Theta [U(\tilde{c}(\theta'), \tilde{y}(\theta'), \theta') + \beta \tilde{w}''(\theta', \theta')] \Pi(\theta, d\theta') \qquad (39)$$

and $\tilde{w}''(\theta, \cdot) \in \tilde{\mathcal{U}}^*$ for all $\theta \in \Theta$. The right hand side of (39) depends on θ only through the function $\Pi(\theta, \cdot)$ which it is twice continuously differentiable. Thus, $\tilde{w}'(\hat{\theta}, \theta)$ is also twice continuously differentiable in θ, otherwise (39) cannot be satisfied. Hence if $\tilde{w}'(\theta, .)$ is not twice differentiable then $\tilde{w}'(\theta, .) \notin \tilde{\mathcal{U}}^*$.

To prove the second part of the Lemma, note that

$$\frac{\partial}{\partial \theta}\tilde{w}'(\hat{\theta},\theta) = \int_{\Theta}[U(\tilde{c}(\theta'),\tilde{y}(\theta'),\theta') + \beta\tilde{w}''(\theta',\theta')]\frac{\Pi_1(\theta,d\theta')}{\Pi(\theta,d\theta')}\Pi(\theta,d\theta').$$

By assumption there exists a constant $K < \infty$ such that $\left|\frac{\Pi_1(\theta,B)}{\Pi(\theta,B)}\right| \leq K$ for all $(\theta',B) \in \Theta \times \mathcal{B}_\Theta$. Hence

$$\left|\frac{\partial}{\partial \theta}\tilde{w}'(\hat{\theta},\theta)\right| \leq K\left|\int_{\Theta}[U(\check{c}(\theta'),\check{y}(\theta'),\theta') + \beta\check{w}'(\theta',\theta')]\Pi(\theta,d\theta')\right| \leq K\max(\frac{|\underline{u}|}{1-\beta},\frac{\bar{u}}{1-\beta}),$$

and so $\left|\frac{\partial}{\partial \theta}\tilde{w}'(\hat{\theta},\theta)\right|$ is bounded. $\blacksquare$

Proof of Theorem (5). i) Necessity. Let $X(\hat{\theta},\theta) = U(\tilde{c}(\hat{\theta}),\tilde{y}(\hat{\theta}),\theta) + \beta\tilde{w}'(\hat{\theta},\theta)$. Let $X^*(\theta) = X(\theta,\theta)$. One has, for $\theta > 0$ and for $\hat{\theta} \in \Theta$,

$$\frac{\partial}{\partial \theta}X(\hat{\theta},\theta) = U_\theta(\tilde{c}_t(\hat{\theta}),\tilde{y}_t(\hat{\theta}),\hat{\theta}) + \frac{\partial}{\partial \theta}\tilde{w}'(\hat{\theta},\theta).$$

It follows from Lemma (4) that the derivative is correctly defined. Take $\theta_0 > \underline{\theta}$. For $\theta \geq \theta_0$ the expression is bounded. The first term U_θ is bounded because utility is bounded and linear in θ. The second term is bounded by Lemma (4). By the mean value theorem, $\alpha_\eta(\theta) = \frac{1}{\eta}[X(\theta,\theta+\eta) - X^*(\theta)]$ is also bounded. Because $\lim_{\eta\to 0}\alpha_\eta(\theta) = \frac{\partial}{\partial \theta}X(\hat{\theta},\theta)$, Lebesgue theorem on bounded convergence implies that for $\theta \geq \theta_0$,

$$\lim_{\eta\to 0}\int_{\theta_0}^{\theta}\alpha_\eta(\varepsilon)d\varepsilon = \int_{\theta_0}^{\theta}\lim_{\eta\to 0}\alpha_\eta(\varepsilon)d\varepsilon = \int_{\theta_0}^{\theta}\frac{\partial}{\partial \theta}X(\varepsilon,\varepsilon)d\varepsilon.$$

Hence

$$\eta\int_{\theta_0}^{\theta}\alpha_\eta(\varepsilon)d\varepsilon = \int_{\theta_0}^{\theta}[X(\varepsilon,\varepsilon+\eta) - X^*(\varepsilon)]d\varepsilon$$
$$\leq \int_{\theta_0}^{\theta}[X^*(\varepsilon+\eta) - X^*(\varepsilon)]d\varepsilon = \int_{0}^{\eta}[X^*(\theta+\tilde{\eta}) - X^*(\theta_0+\tilde{\eta})]d\tilde{\eta}.$$

where the inequality follows from the fact that, by incentive compatibility, $X^*(\theta+\eta) \geq X(\theta,\theta+\eta)$ and the last equality follows from the fact that integration over the interval $[\eta,\theta]$ cancels out.

Take now limits for η converging to 0 from the left and for η converging to 0 from the right. The limits must satisfy

$$\lim_{\eta \to 0_+} \frac{1}{\eta} \int_0^\eta [X^*(\theta+\tilde{\eta})-X^*(\theta_0+\tilde{\eta})]d\tilde{\eta} \geq \lim_{\eta \to 0} \int_{\theta_0}^\theta \alpha_\eta(\varepsilon)d\varepsilon \geq \lim_{\eta \to 0_-} \frac{1}{\eta} \int_0^\eta [X^*(\theta+\tilde{\eta})-X^*(\theta_0+\tilde{\eta})]d\tilde{\eta}.$$

The inequalities follow from the fact that η has different signs in the first and third terms. But both these terms converge to $X^*(\theta) - X^*(\theta_0)$ and so

$$X^*(\theta) - X^*(\theta_0) = \int_{\theta_0}^\theta \frac{\partial}{\partial \theta} X(\varepsilon, \varepsilon)d\varepsilon.$$

Letting $\theta_0 \to \underline{\theta}$ concludes the first part of the proof.

ii) Sufficiency. Suppose that $\frac{\partial}{\partial \theta} X(\hat{\theta}, \theta)$ is nondecreasing in $\hat{\theta}$ for all θ. Assume that $\hat{\theta} < \theta$.

$$X^*(\theta) - X^*(\hat{\theta}) = \int_{\hat{\theta}}^\theta \frac{\partial}{\partial \theta} X(\varsigma, \varsigma)d\varsigma \geq \int_{\hat{\theta}}^\theta \frac{\partial}{\partial \theta} X(\hat{\theta}, \varsigma)d\varsigma = X(\hat{\theta}, \theta) - X^*(\hat{\theta})$$

where the first equality follows from the fact that the necessary envelope condition is supposed to hold. Similar arguments apply for $\hat{\theta} > \theta$ and so one gets $X^*(\theta) \geq X(\hat{\theta}, \theta)$ for all $\hat{\theta} \in \Theta$. ∎

Proof of Theorem (6). Define, for each $(w, g, \theta_-) \in \mathcal{W} \times \mathcal{G} \times \Theta$ a set

$$\Upsilon(w, g, \theta_-) = \{\tilde{\mathbf{w}}(\cdot) \in \tilde{\mathcal{U}}^* : \tilde{\mathbf{w}}(\theta_-) = w, \frac{\partial}{\partial \theta} \tilde{\mathbf{w}}(\theta_-) = g\}.$$

The set $\Upsilon(w, g, \theta_-) \subseteq \tilde{\mathcal{U}}^*$ is a set of all the promised utility functions that deliver promised utility w and marginal promised utility g to an agent whose previous period shock was θ_-.

Let

$$\tilde{V}_t^q(w, g, \theta_-) = \min_{\mathbf{w}(\cdot) \in \Upsilon(w,g,\theta_-).} \Omega_t^q[\mathbf{w}(\cdot), \theta_-]$$

be the minimum of the cost function $\Omega_t^q(\cdot, \theta_-)$ over $\Upsilon(w, g, \theta_-)$ in period t and let $\mathbf{w}_{t,w,g,\theta_-}^*(\cdot)$ be the promised utility function that attains the minimum. For each $t \geq 1$ define a set $\hat{\mathcal{U}}_t^*(\theta_-) = \{\tilde{\mathbf{w}}(\cdot) \in \tilde{\mathcal{U}}^* : \exists (w, g) \in \mathcal{W} \times \mathcal{G} : \tilde{\mathbf{w}}(\cdot) = \mathbf{w}_{t,w,g,\theta_-}^*(\cdot)\}$. The set

$\hat{\mathcal{U}}_t^*(\theta_-)$ contains only the promised utility functions that, for some w and g, reach the minimum of $\Omega_t^q[\mathbf{w}(\cdot), \theta_-]$ over $\Upsilon(w, g, \theta_-)$.

I first show that all the promised utility functions that do not belong to $\tilde{\mathcal{U}}_t^*(\theta_-)$ are, in a sense, redundant: they will never be chosen by the social planner because there is some other promised utility function that delivers lower costs and satisfies all the constraints of the social planner's problem. The proof is by contradiction. Suppose that $(\bar{c}_t, \bar{y}_t, \bar{w}_{t+1})$ are the optimal policy functions in period t that attain the minimum of (11) and that $\bar{w}_{t+1}(\theta, \cdot) \notin \tilde{\mathcal{U}}_{t+1}^*(\theta)$ for almost all $\theta \in \Theta$. Define an alternative continuation utility function $\hat{w}_{t+1}(\hat{\theta}, \theta)$ by $\hat{w}_{t+1}(\hat{\theta}, \theta) = \mathbf{w}_{t+1, \bar{w}_t(\hat{\theta}, \hat{\theta}), \frac{\partial}{\partial\theta}\bar{w}_t'(\hat{\theta}, \hat{\theta}), \hat{\theta}}^*(\theta)$. The policy functions $(\bar{c}_t, \bar{y}_t, \hat{w}_{t+1})$ have lower costs than $(\bar{c}_t, \bar{y}_t, \bar{w}_{t+1})$:

$$\Omega_t^q[\mathbf{w}(\cdot), \theta_-] = \int_\Theta \{\bar{c}_t(\theta) - \bar{y}_t(\theta) + q_t\Omega_{t+1}^q[\bar{w}_{t+1}(\theta, \cdot), \theta]\}\Pi(\theta_-, d\theta)$$

$$> \int_\Theta \{\bar{c}_t(\theta) - \bar{y}_t(\theta) + q_t\Omega_{t+1}^q[\hat{w}_{t+1}(\theta, \cdot), \theta]\}\Pi(\theta_-, d\theta).$$

because by definition of the set $\tilde{\mathcal{U}}_{t+1}^*(\theta)$, $\Omega_{t+1}^q[\bar{w}_{t+1}(\theta, \cdot), \theta] > \Omega_{t+1}^q[\hat{h}_{t+1}(\theta, \cdot), \theta]$ for almost all $\theta \in \Theta$. The policy functions $(\bar{c}_t, \bar{y}_t, \hat{w}_{t+1})$ satisfy the promise keeping constraint (5) because, by construction, $\bar{w}_{t+1}(\theta, \theta) = \hat{w}_{t+1}(\theta, \theta)$ and $(\bar{c}_t, \bar{y}_t, \bar{w}_{t+1})$ satisfies (5). They also satisfy the temporary incentive compatibility constraint (7) because $(\bar{c}_t, \bar{y}_t, \bar{w}_{t+1})$ satisfies (7) and hence the envelope condition (13), the first order approach is valid, $\bar{w}_{t+1}(\theta, \theta) = \hat{w}_{t+1}(\theta, \theta)$ and $\frac{\partial}{\partial\theta}\bar{w}_{t+1}(\theta, \theta) = \frac{\partial}{\partial\theta}\hat{w}_{t+1}(\theta, \theta)$. Clearly, $(\bar{c}_t, \bar{y}_t, \hat{w}_{t+1}) \in \tilde{\mathcal{U}}^*$. Therefore, $(\bar{c}_t, \bar{y}_t, \bar{w}_{t+1})$ does not minimize (11), which is a contradiction.

Hence, it must be true that $\bar{w}_{t+1}(\theta, \cdot) \in \tilde{\mathcal{U}}_{t+1}^*(\theta)$ for almost all $\theta \in \Theta$. The value function Ω_t^q then satisfies, by construction of the set $\tilde{\mathcal{U}}_{t+1}^*(\theta)$,

$$\Omega_t^q[\mathbf{w}(\cdot), \theta_-] = \int_\Theta \{\bar{c}_t(\theta) - \bar{y}_t(\theta) + q_t\tilde{V}_{t+1}^q(\bar{w}_{t+1}(\theta, \theta), \frac{\partial}{\partial\theta}\bar{w}_{t+1}(\theta, \theta), \theta)\}\Pi(\theta_-, d\theta).$$

By taking the minimum of both sides over the set $\Upsilon(\mathbf{w}(\theta_-), \frac{\partial}{\partial\theta}\mathbf{w}(\theta_-), \theta_-)$, one gets that the value function $\tilde{V}_t^q$ satisfies

$$\tilde{V}_t^q[\mathbf{w}(\theta_-), \frac{\partial}{\partial\theta}\mathbf{w}(\theta_-), \theta_-] = \int_\Theta \{\hat{c}_t(\theta) - \hat{y}_t(\theta) + q_t\tilde{V}_{t+1}^q[\bar{w}_{t+1}(\theta), \bar{g}_{t+1}(\theta), \theta]\}\Pi(\theta_-, d\theta)$$

$$(40)$$

where

$$\hat{c}_t(\mathbf{w}(\theta_-), \frac{\partial}{\partial \theta}\mathbf{w}(\theta_-), \theta_-, \theta) = \bar{c}_t(\mathbf{w}^*_{t, \mathbf{w}(\theta_-), \frac{\partial}{\partial \theta}\mathbf{w}(\theta_-), \theta_-}(\cdot), \theta_-, \theta)$$

$$\hat{y}_t(\mathbf{w}(\theta_-), \frac{\partial}{\partial \theta}\mathbf{w}(\theta_-), \theta_-, \theta) = \bar{y}_t(\mathbf{w}^*_{t, \mathbf{w}(\theta_-), \frac{\partial}{\partial \theta}\mathbf{w}(\theta_-), \theta_-}(\cdot), \theta_-, \theta)$$

$$\hat{w}_{t+1}(\mathbf{w}(\theta_-), \frac{\partial}{\partial \theta}\mathbf{w}(\theta_-), \theta_-, \theta) = \bar{w}_{t+1}(\mathbf{w}^*_{t, \mathbf{w}(\theta_-), \frac{\partial}{\partial \theta}\mathbf{w}(\theta_-), \theta_-}(\cdot), \theta_-, \theta, \theta)$$

$$\hat{g}_{t+1}(\mathbf{w}(\theta_-), \frac{\partial}{\partial \theta}\mathbf{w}(\theta_-), \theta_-, \theta) = \frac{\partial}{\partial \theta}\bar{w}_{t+1}(\mathbf{w}^*_{t, \mathbf{w}(\theta_-), \frac{\partial}{\partial \theta}\mathbf{w}(\theta_-), \theta_-}(\cdot), \theta_-, \theta, \theta).$$

are the policy functions that attain the minimum. The functions $(\hat{c}_t, \hat{y}_t, \hat{w}_{t+1}, \hat{g}_{t+1})$ satisfy the envelope condition (14) since $(\bar{c}_t, \bar{y}_t, \bar{w}_{t+1})$ satisfy the envelope condition (13). They also satisfy the promise keeping constraint (15) and the marginal promise keeping constraint (16) for $\mathbf{w}(\theta_-)$ and $\frac{\partial}{\partial \theta}\mathbf{w}(\theta_-)$ since $(\bar{c}_t, \bar{y}_t, \bar{w}_{t+1})$ satisfy the promise keeping constraint (5) for $\mathbf{w}^*_{t, \mathbf{w}(\theta_-), \frac{\partial}{\partial \theta}\mathbf{w}(\theta_-), \theta_-}(\cdot)$. Moreover, $(\hat{w}_{t+1}(\theta), \hat{g}_{t+1}(\theta))$ belongs to $\mathcal{U}^*(\theta)$ since $\bar{w}_{t+1} \in \tilde{\mathcal{U}}^*$. The functions $(\hat{c}_t, \hat{y}_t, \hat{w}_{t+1}, \hat{g}_{t+1})$ minimize the right-hand side of (40) subject to (14), (15), (16) and (17): if not, one can find another allocation rule $(\ddot{c}_t, \ddot{y}_t, \ddot{w}_{t+1})$ that delivers lower costs than $(\bar{c}_t, \bar{y}_t, \bar{w}_{t+1})$, a contradiction. Therefore, $V^q_t(w, g, \theta_-) = \tilde{V}^q_t(w, g, \theta_-)$. Consequently,

$$
\begin{aligned}
V^q(w_1, \theta_0) &= \min_{g \in \mathcal{G}:(w_1, g) \in \mathcal{U}^*(\theta_0)} \tilde{V}^q_1(w_1, g, \theta_0) \\
&= \min_{g \in \mathcal{G}:(w_1, g) \in \mathcal{U}^*(\theta_0)} \{ \min_{\mathbf{w}(\cdot) \in \Upsilon(w, g, \theta_0).} \Omega^q_1[\mathbf{w}(\cdot), \theta_0] \} \\
&= \min_{\mathbf{w}(\cdot) \in \mathcal{L}_\Theta : \mathbf{w}(\theta_0) = w_1.} \Omega^q_1[\mathbf{w}(\cdot), \theta_0] \\
&= \Omega^q(w_1, \theta_0).
\end{aligned}
$$

where the second equality follows from the definition of $\tilde{V}^q_1$ and the third equality follows from the fact that by construction of the set Υ, $\min_{g \in \mathcal{G}:(w_1, g) \in \mathcal{U}^*(\theta_0)} \Upsilon(w, g, \theta_-) = \{\tilde{\mathbf{w}}(\cdot) \in \tilde{\mathcal{U}}^* : \tilde{\mathbf{w}}(\theta_-) = w\}$. ∎

Proof of Lemma 7. It will be first verified that the value function satisfies the guess. It will then be shown that all the constraints satisfy the guess as well. Using

the guess for the right hand side of the Bellman equation, one gets

$$
\begin{aligned}
V(w, g, \theta_-) &= \int_{\theta - \theta_L}^{\infty} \{ e^{u(w, g, \theta_-, \theta)} + \beta V[w'(w, g, \theta_-, \theta), g'(w, g, \theta_-, \theta), \theta] \} \pi(\theta_-, d\theta) \\
&= \int_{\theta - \theta_L}^{\infty} \{ e^{u(\frac{w}{\theta_-}, g, 1, \frac{\theta}{\theta_-})} + \beta V[\theta_- w'(\frac{w}{\theta_-}, g, 1, \frac{\theta}{\theta_-}), g'(\frac{w}{\theta_-}, g, 1, \frac{\theta}{\theta_-}), \theta] \} \pi(1, \frac{d\theta}{\theta_-}) \\
&= \int_{\theta - \theta_L}^{\infty} \{ e^{u(\frac{w}{\theta_-}, g, 1, \frac{\theta}{\theta_-})} + \beta V[\frac{\theta_-}{\theta} w'(\frac{w}{\theta_-}, g, 1, \frac{\theta}{\theta_-}), g'(\frac{w}{\theta_-}, g, 1, \frac{\theta}{\theta_-}), 1] \} \pi(1, \frac{d\theta}{\theta_-}) \\
&= \int_{\theta_L}^{\infty} [e^{u(\frac{w}{\theta_-}, g, 1, \varepsilon)} + \beta V(\frac{1}{\varepsilon} w'(\frac{w}{\theta_-}, g, 1, \varepsilon), g'(\frac{w}{\theta_-}, g, 1, \varepsilon), 1)] \pi(1, d\varepsilon) \\
&= V(\frac{w}{\theta_-} g, 1)
\end{aligned}
$$

where the first equality uses the fact that u, w' and g' attain the minimum of the Bellman equation, the second equality uses the guess about the optimal policy functions and the properties of the density function, the third equality uses the guess about the value function, the fourth equality substitutes ε for $\frac{\theta}{\theta_-}$ and the last equality uses again the fact that u, w' and g' are optimal policy functions.

The promise keeping constraint (21) can be, using the guess about the optimal policy functions, written as

$$
\begin{aligned}
w &= \int_{\theta - \theta_L}^{\infty} [\theta u(w, g, \theta_-, \theta) + \beta w'(w, g, \theta_-, \theta)] \pi(\theta_-, d\theta) \\
&= \theta_- \int_{\theta - \theta_L}^{\infty} [\frac{\theta}{\theta_-} u(\frac{w}{\theta_-}, g, 1, \frac{\theta}{\theta_-}) + \beta w'(\frac{w}{\theta_-} g, 1, \frac{\theta}{\theta_-})] \pi(1, \frac{d\theta}{\theta_-}) \\
&= \theta_- \int_{\theta_L}^{\infty} [\varepsilon u(\frac{w}{\theta_-}, g, 1, \varepsilon) + \beta w'(\frac{w}{\theta_-}, g, 1, \varepsilon)] \pi(1, d\varepsilon)
\end{aligned}
$$

where again the substitution $\varepsilon = \frac{\theta}{\theta_-}$ was used.

To show that the marginal threat keeping constraint satisfies the normalization as well, note that $w_0(w, g, \theta_-) = \theta_- w_0(\frac{w}{\theta_-}, g, 1)$. Hence (23) can be written as

$$
g = \frac{\eta}{\theta_-} [w - w_0(w, g, \theta_-)] = \eta [\frac{w}{\theta_-} - w_0(\frac{w}{\theta_-}, g, 1)]
$$

and so the guess satisfies the marginal threat keeping constraint as well.

44

Finally, for the incentive compatibility constraint (20) one gets that

$$\frac{\theta}{\theta_-}u(\frac{w}{\theta_-},g,1,\frac{\theta}{\theta_-}) + \beta w'(\frac{w}{\theta_-},g,1,\frac{\theta}{\theta_-})$$

$$= \frac{1}{\theta_-}[\theta u(w,g,\theta_-,\theta) + \beta w'(w,g,\theta_-,\theta)]$$

$$= \frac{1}{\theta_-}\{\int_{\theta_-\theta_L}^{\theta}[u(w,g,\theta_-,\varepsilon) + \beta g'(w,g,\theta_-,\varepsilon)]d\varepsilon + w_0(w,g,\theta_-)\}$$

$$= \int_{\theta_-\theta_L}^{\theta}[u(\frac{w}{\theta_-},g,\frac{\varepsilon}{\theta_-},1) + \beta g'(\frac{w}{\theta_-}g,\frac{\varepsilon}{\theta_-},1)]\frac{d\varepsilon}{\theta_-} + w_0(\frac{w}{\theta_-},g,1)$$

$$= \int_{\theta_L}^{\frac{\theta}{\theta_-}}[u(\frac{w}{\theta_-},g,\tilde{\varepsilon},1) + \beta g'(\frac{w}{\theta_-},g,\tilde{\varepsilon},1)]d\tilde{\varepsilon} + w_0(\frac{w}{\theta_-},g,1),$$

where the first equality uses the guessed functional forms, the second equality uses the incentive compatibility constraint, the third one uses the guesses again and the last one makes a substitution $\tilde{\varepsilon} = \frac{\varepsilon}{\theta_-}$ in the integration. It follows that the incentive compatibility constraint is satisfied for the guess. ∎

Proof of Lemma 8. The proof follows similar structure as the proof of Lemma (7). To simplify notation, define the normalized variables as follows. Let $\tilde{v}(w,g) = V(w,g,1)$ be the normalized value function. Define the the current utility function by $\tilde{r}(w,g,\theta) = u(w,g,1,\theta)$, the normalized continuation utility $\tilde{h}'(w,g,\theta) = \frac{w'(w,g,1,\theta)}{\theta}$ and the marginal continuation utility by $\tilde{m}'(w,g,\theta) = g'(w,g,1,\theta)$.

The value function satisfies

$$\tilde{v}(w,g) = \int_{\theta_L}^{\infty}\{e^{\tilde{r}(w,g,\theta)} + \beta\tilde{v}[\tilde{h}'(w,g,\theta), w + \tilde{m}'(w,g,\theta)]\}\pi d\theta$$

$$= \int_{\theta_L}^{\infty}\{e^{(1-\beta)w+\tilde{r}(0,g-w,\theta)} + \beta\tilde{v}[w + \tilde{h}'(0,g-w,\theta), w + \tilde{m}'(0,g-w,\theta)]\}\pi d\theta$$

$$= e^{(1-\beta)w}\int_{\theta_L}^{\infty}\{e^{\tilde{r}(0,g-w,\theta)} + \beta e^{(1-\beta)W'(0,g-w,\theta)}\tilde{v}[\tilde{m}'(0,g-w,\theta) - \tilde{h}'(0,g-w,\theta)]\}\pi d\theta$$

$$= e^{(1-\beta)w}\tilde{v}(0,g-w).$$

where the first and last equality uses the fact that $\tilde{r}$, $\tilde{h}'$ and $\tilde{m}'$ attain the maximum of the Bellman equation, the second equality uses the guess about the functional

forms of the optimal policy functions and the third equality uses the guess about the functional form of the value function.

The promise keeping constraint satisfies, using the guess,

$$
\begin{aligned}
w &= \int_{\theta_L}^{\infty} \{\theta \tilde{r}(w, g, \theta) + \beta \theta \tilde{h}'(w, g, \theta)]\} \pi d\theta \\
&= \int_{\theta_L}^{\infty} \{\theta[(1-\beta)w + \tilde{r}(0, g-w, \theta)] + \beta \theta[w + \tilde{h}'(0, g-w, \theta)]\} \pi d\theta \\
&= w + \int_{\theta_L}^{\infty} [\theta \tilde{r}(0, g-w, \theta) + \beta \theta \tilde{h}'(0, g-w, \theta)] \pi d\theta
\end{aligned}
$$

and the promised utility w cancels out.

For the marginal threat keeping constraint note that $\tilde{h}_0(w, g) = \theta_L w + \tilde{h}_0(0, g-w)$. The marginal threat keeping constraint can then be written as

$$
\begin{aligned}
g &= \eta[w - \tilde{h}_0(w, g)] \\
&= \eta[w - \theta_L w - \tilde{h}_0(0, g-w)] \\
&= w - \mu \tilde{h}_0(0, g-w)
\end{aligned}
$$

where the last equality uses the fact that $\eta(1 - \theta_L) = 1$.

Finally, the incentive compatibility constraint can be rewritten as follows:

$$
\begin{aligned}
\theta[\tilde{r}(0, g-w, \theta) &+ \beta \tilde{h}'(0, g-w, \theta)] \\
&= \theta[\tilde{r}(w, g, \theta) + \beta \tilde{h}'(w, g, \theta)] - \theta w \\
&= \int_{\theta_L}^{\theta} [\tilde{r}(w, g, \theta) + \beta \tilde{z}'(w, g, \theta)] d\varepsilon + \tilde{h}_0(w, g) - \theta w \\
&= \int_{\theta_L}^{\theta} [\tilde{r}(0, g-w, \theta) + \beta \tilde{z}'(0, g-w, \theta) + w] d\varepsilon + \tilde{h}_0(0, g-w) + (\theta_L - \theta)w \\
&= \int_{\theta_L}^{\theta} [\tilde{r}(0, g-w, \theta) + \beta \tilde{z}'(0, g-w, \theta)] d\varepsilon + \tilde{h}_0(0, g-w)
\end{aligned}
$$

where the guess about the optimal policy functions is used in the first and third equality, while the second equality uses the incentive compatibility constraint and the fourth equality cancels all the term involving the promised utility w. ∎

Proof of Lemma (10). Let $m'_\infty = \lim_{\theta\to\infty} m'(\theta)$ and $x_\infty = \lim_{\theta\to\infty} x(\theta)$ Consider a truncated problem, where the distribution of shocks is truncated at $\bar\theta > \theta_L$. Let $m'_{\bar\theta}$ and $x_{\bar\theta}$ be the optimal policy functions for such a problem. Equation (32) implies that $m'_{\bar\theta}(\bar\theta) = g^*_{\bar\theta}$ where $g^*_{\bar\theta}$ minimizes the value function in the truncated problem. Since $\epsilon_v(m'_{\bar\theta}(\bar\theta)) = 0$, equation (33) implies that $\frac{dm'_{\bar\theta}(\bar\theta)}{d\ln\theta} \sim \lambda - (1-\beta)x_{\bar\theta}(\bar\theta)$.

Now let $\bar\theta \to \infty$. Then $x_\infty = \lim_{\bar\theta\to\infty} x_{\bar\theta}(\bar\theta)$ and $m'_\infty = \lim_{\bar\theta\to\infty} m'_{\bar\theta}(\bar\theta) = \lim_{\bar\theta\to\infty} g^*_{\bar\theta} = g^*$. But, since $m'_{\bar\theta}(\bar\theta)$ tends to g^*, $\frac{dm'_{\bar\theta}(\bar\theta)}{d\ln\theta}$ tends to zero and so $\lambda = (1-\beta)x_\infty$. Thus, $x_\infty = \frac{\lambda}{1-\beta}$, is finite and by equation (34), $m'_\infty = \frac{1}{\beta-\beta^2}$. Thus, $g^* = \frac{1}{\beta-\beta^2}$. ∎

Proof of Lemma (11). I will show that the Lemma holds for $g = g^*$ and only note that the arguments hold a fortiori for $g > g^*$.

If $g = g^*$ the envelope theorem and the fact that $v_g(g^*) = 0$ implies that $\int f\varepsilon[\lambda - (1-\beta)x]d\varepsilon = 0$. Hence, by equation (32), $m'(g^*,\theta_L) = g^*$. It is also easy to show that $(1-\beta)x(\theta_L) = \frac{1}{\theta_L}e^{-\frac{1}{\beta(\eta-1)}}[(1-\beta)v(g^*)]^\beta$, a fact that will become useful later. The proof proceeds in two steps.

i) I will first show that $m'(g^*,\theta)$ is increasing in the neighborhood of θ_L. The proof is by contradiction. Suppose that $\frac{dm'(g^*,\theta_L)}{dt} \leq 0$. It follows from (33)

$$0 \geq \frac{dm'(g^*,\theta_L)}{dt} = \frac{x(\theta_L)(1-\beta) - \lambda}{x(\theta_L)\epsilon'_v(g^*)}.$$

Since the value function is convex, $\epsilon'_v(g^*) > 0$ and so $(1-\beta)x(\theta_L) < \lambda$. Now assume that there is some $\bar\theta$, $\theta_L < \bar\theta < \infty$ such that $m'(g^*,\bar\theta) = g^*$. Such $\bar\theta$ must exists, otherwise (29) is violated. If there is more than one such $\bar\theta$, take the smallest one. Then, $m'(\theta) < g^*$ for all $\theta \in (\theta_L, \bar\theta)$ by construction and m' must be increasing at $\bar\theta$. The incentive compatibility constraint implies

$$r(\bar\theta) + \beta h'(\bar\theta) = \beta \int_{\theta_L}^{\bar\theta} m'(\varepsilon)\frac{d\varepsilon}{\varepsilon} - \frac{g^*}{\eta-1} < \frac{1}{1-\beta}[\ln(\frac{\bar\theta}{\theta_L}) - \frac{1}{\beta(\eta-1)}].$$

where the inequality follows from the fact that $m'(\theta) < g^* = \frac{1}{\beta-\beta^2}$ for all $\theta \in (\theta_L, \bar\theta)$.

By definition of x and by (30),

$$(1 - \beta)x(\bar{\theta}) = \frac{1}{\theta}e^{(1-\beta)[r(\bar{\theta})+\beta h'(\bar{\theta})]}[(1-\beta)v(g^*)]^\beta < \frac{1}{\theta_L}e^{-\frac{1}{\beta(\eta-1)}}[(1-\beta)v(g^*)]^\beta$$

$$= (1 - \beta)x(\theta_L).$$

Hence $(1 - \beta)x(\bar{\theta}) < \lambda$ and, by (33), $\frac{dm'(g^*,\bar{\theta})}{dt} < 0$. The marginal continuation utility is decreasing at $\bar{\theta}$, which is a contradiction. Therefore $\frac{dm'(g^*,\theta_L)}{dt} > 0$ and the marginal continuation utility is locally increasing at θ_L.

ii) I will now show by contradiction that m' will be always greater than g^*. Assume that there is some $\bar{\theta}$, $\theta_L < \bar{\theta} < \infty$ such that $m'(g^*,\bar{\theta}) = g^*$ If there is more than one such $\bar{\theta}$, take the smallest one. By construction, $m'(\theta) \geq g^*$ for all $\theta \in (\theta_L, \bar{\theta})$ and m' must be decreasing at $\bar{\theta}$. Similarly to part i), one can show that $(1 - \beta)x(\bar{\theta}) \geq (1 - \beta)x(\theta_L)$ and so $\frac{dm'(g^*,\bar{\theta})}{dt} > 0$, which is a contradiction. Thus, there is no such $\bar{\theta}$ and $m'(g^*,\theta) > g^*$ for all $\theta > \theta_L$. ∎

10   Appendix 2: Proof of Theorem 14

Several preliminary results and definitions are needed before one can prove Theorem 14. Define a sequence cost minimization problem as a dual formulation of the sequence utility maximization problem. For a given promised utility $w_1 \in \mathcal{W}$, the social planner selects a pair of consumption and output functions $C_t^d : \mathcal{W} \times \Theta^t \to \mathcal{R}_+$ and $Y_t^d : \mathcal{W} \times \Theta^t \to \mathcal{R}_+$ for each period. Call the collection of these assignments for all period $Z^d = \{C_t^d, Y_t^d\}_{t=1}^\infty$ an allocation in a sequence cost minimization problem. Define the set of allocations that satisfy the incentive compatibility constraint (1) and a promise keeping constraint

$$W(Z^d, \theta_0) = w_1 \tag{41}$$

by $\mathfrak{Z}^d(w_1, \theta_0)$. For a given w_1 and θ_0, the present value of the resource costs implied by Z^d is

$$e^d = \sum_{t=1}^\infty Q_t \int_{\Theta^t} [C_t^d(\theta^t) - Y_t^d(\theta^t)]\mu^t(\theta_0, d\theta^t).$$

48

I will now show that for any allocation in the sequence cost minimization problem one can construct another one that satisfies an *equal treatment property*, defined as follows: All the agents that arrive in period t with identical promised utility will receive an identical allocation in the current period and an identical continuation utility from tomorrow on, regardless of their past history. The Lemma closely parallels Lemma A.2 of Atkeson and Lucas [5].

Lemma 12 *Suppose $Z \in \mathfrak{Z}^d(w_1, \theta_0)$ is an allocation in the sequence cost minimization problem that is incentive compatible and satisfies the promise keeping constraint (41). Then there is another allocation $\bar{Z} \in \mathfrak{Z}^d(w_1, \theta_0)$, that is incentive compatible, satisfies the promise keeping constraint (41), delivers weakly lower costs each period and exhibits the following equal treatment property:*

$$W_t(\bar{Z}, \hat{\theta}^{t-1}, \cdot) = W_t(\bar{Z}, \tilde{\theta}^{t-1}, \cdot)$$

for all $(\hat{\theta}^{t-1}, \tilde{\theta}^{t-1}) \in \Theta^{2t}$ implies

$$\bar{Y}_t(w_1, \hat{\theta}^{t-1}, \theta_t) = \bar{Y}_t(w_1, \tilde{\theta}^{t-1}, \theta_t)$$
$$\bar{C}_t(w_1, \hat{\theta}^{t-1}, \theta_t) = \bar{C}_t(w_1, \tilde{\theta}^{t-1}, \theta_t)$$
$$W_{t+1}(\bar{Z}, (\hat{\theta}^{t-1}, \theta_t), \cdot) = W_{t+1}(\bar{Z}, (\tilde{\theta}^{t-1}, \theta_t), \cdot)$$

for all $\theta_t \in \Theta$.

Proof. Fix w_1. Let t be the earliest date when Z does not satisfy the equal treatment property. Suppose there is a set of histories $D \subseteq \Theta^{t-1}$ such that $W_t(Z, \hat{\theta}^{t-1}, \cdot) = W_t(Z, \tilde{\theta}^{t-1}, \cdot)$ but the equal treatment property fails on D. For $j \geq t$ define $\bar{U}_{t+j}[\hat{\theta}^{t-1}, \theta_t^{t+j}), \theta_{t+j}]$ to be the average over the period utilities: For any $\hat{\theta}^{t-1} \in D$,

$$\bar{U}_{t+j}[(\hat{\theta}^{t-1}, \theta_t^{t+j}), \theta_{t+j}] = \int\limits_{\Theta^{t-1}} U[C_{t+j}(\hat{\theta}^{t-1}, \theta_t^{t+j}), Y_{t+j}(\hat{\theta}^{t-1}, \theta_t^{t+j}), \theta_{t+j}] \tilde{\mu}^{t-1}(d\hat{\theta}^{t-1})$$

where $\tilde{\mu}^{t-1}(B)$ is defined for any Borel measurable $B \subseteq D$ by $\tilde{\mu}^{t-1}(B) = \frac{\mu^{t-1}(\theta_0, B)}{\mu^{t-1}(\theta_0, D)}$.

49

Since the utility is linear in θ, it can take two possible forms. The first is that $U(C,Y,\theta) = \theta U^1(C,Y) + U^2(Y)$, the second is that $U(C,Y,\theta) = \theta U^1(C,Y) + U^2(C)$. I will show the result for the first case; the second one can be treated analogously.

Define a new allocation $\bar{Z}$ as follows. For any $\hat{\theta}^{t-1} \in D$, let

$$\bar{Y}_{t+j}(\hat{\theta}^{t-1}, \theta_t^{t+j}) = \int_{\Theta^{t-1}} \bar{Y}_{t+j}(\hat{\theta}^{t-1}, \theta_t^{t+j}) \tilde{\mu}^{t-1}(d\hat{\theta}^{t-1}) \tag{42}$$

$$\bar{C}_{t+j}(\hat{\theta}^{t-1}, \theta_t^{t+j}) = U^{1-1}\{\bar{U}_{t+j}[(\hat{\theta}^{t-1}, \theta_t^{t+j}), \theta_{t+j}], \bar{Y}_{t+j}(\tilde{\theta}^{t-1}, \theta_t^{t+j})\}. \tag{43}$$

Let $\bar{Z} = Z$ after all other histories. By construction, the allocation $\bar{Z}$ satisfies for any $\hat{\theta}^{t-1} \in D$

$$W_{t+1}[\bar{Z}, (\hat{\theta}^{t-1}, \theta_t), \cdot]$$
$$= \sum_{j=0}^{\infty} \int_{\Theta^{j+1}} \beta^j \bar{U}_{t+j}[(\hat{\theta}^{t-1}, \theta_t^{t+j}), \theta_{t+j}] \mu_{t+1}^{t+1+j}(\cdot, d\theta_{t+1}^{t+1+j})$$
$$= \sum_{j=0}^{\infty} \int_{\Theta^{j+1}} \beta^j \int_{\Theta^{t-1}} U[C_{t+j}(\hat{\theta}^{t-1}, \theta_t^{t+j}), Y_{t+j}(\hat{\theta}^{t-1}, \theta_t^{t+j}), \theta_{t+j}] \tilde{\mu}^{t-1}(d\hat{\theta}^{t-1}) \mu_{t+1}^{t+1+j}(\cdot, d\theta_{t+1}^{t+1+j})$$
$$= \int_{\Theta^{t-1}} \left\{ \sum_{j=0}^{\infty} \int_{\Theta^{j+1}} \beta^j U[C_{t+j}(\hat{\theta}^{t-1}, \theta_t^{t+j}), Y_{t+j}(\hat{\theta}^{t-1}, \theta_t^{t+j}), \theta_{t+j}] \mu_{t+1}^{t+1+j}(\cdot, d\theta_{t+1}^{t+1+j}) \right\} \tilde{\mu}^{t-1}(d\hat{\theta}^{t-1})$$
$$= \int_{\Theta^{t-1}} W_{t+1}[Z, (\hat{\theta}^{t-1}, \theta_t), \cdot] \tilde{\mu}^{t-1}(d\hat{\theta}^{t-1}). \tag{44}$$

It also satisfies the temporary incentive compatibility constraint since for all $\hat{\theta}_t \in \Theta$,

$$\bar{U}_t[(\hat{\theta}^{t-1}, \theta_t), \theta_t] + \beta W_{t+1}[\bar{Z}, (\hat{\theta}^{t-1}, \theta_t), \theta_t]$$
$$= \int \{U[C_t(\hat{\theta}^{t-1}, \theta_t), Y_t(\hat{\theta}^{t-1}, \theta_t), \theta_t] + \beta W_{t+1}[Z, (\hat{\theta}^{t-1}, \theta_t), \theta_t]\} \tilde{\mu}^{t-1}(d\hat{\theta}^{t-1})$$
$$\geq \int \{U[C_t(\hat{\theta}^{t-1}, \hat{\theta}_t), Y_t(\hat{\theta}^{t-1}, \theta_t), \hat{\theta}_t] + \beta W_{t+1}[Z, (\hat{\theta}^{t-1}, \theta_t), \hat{\theta}_t]\} \tilde{\mu}^{t-1}(d\hat{\theta}^{t-1})$$
$$= \bar{U}_t[(\hat{\theta}^{t-1}, \hat{\theta}_t), \theta_t] + \beta W_{t+1}[\bar{Z}, (\hat{\theta}^{t-1}, \hat{\theta}_t), \theta_t]$$

where both equalities follow from (44) and the inequality follows from the fact that Z satisfies the temporary incentive compatibility constraint. The same result applies

after any history $(\hat{\theta}^{t-1}, \theta_t^{t+j}) \in \Theta^j$ for all $j \geq 0$. After all other histories, $\bar{Z}$ is identical to Z and so satisfies the temporary incentive compatibility constraint as well. Hence $\bar{Z}$ is incentive compatible.

The allocation $\bar{Z}$ delivers the same utility as Z. It also requires weakly less resources than Z. This follows from the fact that expected revenues are the same as before because of (42), but expected costs are weakly lower because of (43) and because $U^1(C, Y)$ is concave in C. The argument applies for any number of instances when the equal treatment property fails. This completes the proof. ∎

The next Lemma relates the allocations in the sequence cost minimization problem to the recursive allocations.

Lemma 13 *i) Suppose* $\tilde{z} \in \tilde{\mathfrak{z}}(w_1, \theta_0)$ *is an allocation rule. Let* $e = \tilde{E}^q(\tilde{z}, w_1, \theta_0)$ *be the implied resource costs. Then there exists an allocation in the sequence cost minimization problem* $Z^d \in \mathfrak{Z}^d(w_1, \theta_0)$ *that has the implied resource costs* $e^d = e$.

ii) Suppose $Z^d \in \mathfrak{Z}^d(w_1, \theta_0)$ *is an allocation in the sequence cost minimization problem. Let* e^d *be the implied resource costs it delivers. Then there exists an allocation rule* $\tilde{z} \in \tilde{\mathfrak{z}}(w_1, \theta_0)$ *that has the implied resource costs* $\tilde{E}^q(\tilde{z}, w_1, \theta_0) = e^d$.

Proof. (i) For a given allocation rule $\tilde{z}$ define an allocation in the sequence cost minimization problem as follows. Let $\hat{W}_t^d(w_1, \hat{\theta}^{t-1}, \theta_{t-1})$ solve a difference equation

$$\hat{W}_1^d(w_1, \hat{\theta}_1, \theta_1) = \tilde{w}_2(w_1, \theta_0, \hat{\theta}_1, \theta_1), \tag{45}$$
$$\hat{W}_{t+1}^d(w_1, \hat{\theta}^t, \theta_t) = \tilde{w}_{t+1}[\hat{W}_t^d(w_1, \hat{\theta}^{t-1}, \cdot), \hat{\theta}_{t-1}, \hat{\theta}_t, \theta_t] \quad t > 1.$$

Set $C_t^d(w_1, \hat{\theta}^t) = \tilde{c}_t[\hat{W}_t^d(w_1, \hat{\theta}^{t-1}, \cdot), \hat{\theta}_{t-1}, \hat{\theta}_t]$ and $Y_t^d(w_1, \hat{\theta}^t) = \tilde{y}_t[\hat{W}_t^d(w_1, \hat{\theta}^{t-1}, \cdot), \hat{\theta}_{t-1}, \hat{\theta}_t]$. Call Z^d an allocation generated by the allocation rule $\tilde{z}$.

I will show that Z^d delivers the expected utility w_1 and that it is incentive compatible.

(ia) Promise keeping. Let Z^d be generated from $\tilde{z}$ and let $\{\hat{W}^d_{t+1}(w_1, \hat{\theta}^t, \theta_t)\}^\infty_{t=0}$ solve (45). Fix $w_1, \hat{\theta}^{t-1}$ and θ_{t-1}. The promise keeping constraint (5) implies that

$$\hat{W}^d_t(w_1, \hat{\theta}^{t-1}, \theta_{t-1}) \tag{46}$$
$$= \int_\Theta \{U[\tilde{c}_t[\hat{W}^d_t(w_1, \hat{\theta}^{t-1}, \cdot), \hat{\theta}_{t-1}, \theta_t], \tilde{y}_t[\hat{W}^d_t(w_1, \hat{\theta}^{t-1}, \cdot), \hat{\theta}_{t-1}, \theta_t], \theta_t]$$
$$+ \beta\tilde{w}_{t+1}[\hat{W}^d_t(w_1, \hat{\theta}^{t-1}, \cdot), \hat{\theta}_{t-1}, \theta_t, \theta_t]\}\Pi(\theta_{t-1}, d\theta_t)$$
$$= \int_\Theta \{U[C^d_t(w_1, \hat{\theta}^{t-1}, \theta_t), Y^d_t(w_1, \hat{\theta}^{t-1}, \theta_t), \theta_t] + \beta\hat{W}^d_{t+1}[w_1, (\hat{\theta}^{t-1}, \theta_t), \theta_t]\}\Pi(\theta_{t-1}, d\theta_t).$$

where the last equality follows from the definition of $\hat{W}^d_t$, C^d_t and Y^d_t. Using (5) repeatedly $T-1$ times to substitute for $\hat{W}^d_{t+j}$, $j = 1..T-1$ one gets

$$\hat{W}^d_t(w_1, \hat{\theta}^{t-1}, \theta_{t-1}) = \sum_{j=0}^T \int_{\Theta^{j+1}} \beta^j U[C^d_{t+j}(w_1, (\hat{\theta}^{t-1}, \theta^{t+j}_t), Y^d_{t+j}(w_1, (\hat{\theta}^{t-1}, \theta^{t+j}_t), \theta_{t+j})\mu^{t+j}_t(\theta_{t-1}, d\theta^{t+j}_t)$$
$$+ \beta^T \int_{\Theta_T} \hat{W}^d_{t+T}[w_1, (\hat{\theta}^{t-1}, \theta^{t+T}_t), \theta_T]\}\Pi(\theta_{T-1}, d\theta_T).$$

Using the fact that utility function is bounded, one can take the limit as $T \to \infty$ and obtains

$$\hat{W}^d_t(w_1, \hat{\theta}^{t-1}, \theta_{t-1}) = \sum_{j=0}^\infty \int_{\Theta^j} \beta^j U[C^d_{t+j}(w_1, (\hat{\theta}^{t-1}, \theta^{t+j}_t), Y^d_{t+j}(w_1, (\hat{\theta}^{t-1}, \theta^{t+j}_t), \theta_{t+j})\mu^{t+j}_t(\theta_{t-1}, d\theta^{t+j}_t).$$
$$\tag{47}$$

Hence

$$\hat{W}^d_t(w_1, \hat{\theta}^{t-1}, \theta_{t-1}) = W_t(Z^d, \hat{\theta}^{t-1}, \theta_{t-1}). \tag{48}$$

Setting $t = 1$ and $\theta_{t-1} = \theta_0$ proves (41) because the initial condition of the difference equation (45) implies that the left hand side of (48) is equal to $\hat{W}^d_1(w_1, \varnothing, \theta_0) = w_1$.

(ib) Incentive compatibility. By Lemma (1), it suffices to show that Z^d satisfies (4) to show that Z^d is incentive compatible. Take any (w_1, θ^{t-1}) pair. Applying (7),

one gets that

$$U[C_t^d(w_1, \hat{\theta}^{t-1}, \theta_t), Y_t^d(w_1, \hat{\theta}^{t-1}, \theta_t), \theta_t] + \beta \hat{W}_{t+1}^d[w_1, (\hat{\theta}^{t-1}, \theta_t), \theta_t]$$

$$= U[\tilde{c}_t[\hat{W}_t^d(w_1, \hat{\theta}^{t-1}, \cdot), \hat{\theta}_{t-1}, \theta_t], \tilde{y}_t[\hat{W}_t^d(w_1, \hat{\theta}^{t-1}, \cdot), \hat{\theta}_{t-1}, \theta_t], \theta_t]$$

$$+ \beta \tilde{w}_{t+1}'[\hat{W}_t^d(w_1, \hat{\theta}^{t-1}, \cdot), \hat{\theta}_{t-1}, \theta_t, \theta_t]$$

$$\geq U[\tilde{c}_t[\hat{W}_t^d(w_1, \hat{\theta}^{t-1}, \cdot), \hat{\theta}_{t-1}, \hat{\theta}_t], \tilde{y}_t[\hat{W}_t^d(w_1, \hat{\theta}^{t-1}, \cdot), \hat{\theta}_{t-1}, \hat{\theta}_t], \theta_t]$$

$$+ \beta \tilde{w}_{t+1}'[\hat{W}_t^d(w_1, \hat{\theta}^{t-1}, \cdot), \hat{\theta}_{t-1}, \hat{\theta}_t, \theta_t]$$

$$= U[C_t^d(w_1, \hat{\theta}^{t-1}, \hat{\theta}_t), Y_t^d(w_1, \hat{\theta}^{t-1}, \hat{\theta}_t), \theta_t] + \beta \hat{W}_{t+1}^d[w_1, (\hat{\theta}^{t-1}, \hat{\theta}_t), \theta_t],$$

where the equalities follow from the definition of $\hat{W}_t^d$, C_t^d and Y_t^d. Applying (48), one gets that Z^d satisfies (4).

(ic) Resource clearing. I will first show by induction that for any function ϑ : $\mathcal{L}^\Theta \to \mathcal{R}$ and for any $t > 1$,

$$\int_{\mathcal{L}^\Theta} \vartheta[\mathbf{w}(\cdot)] d\tilde{\varphi}_t = \int_{\Theta^{t-1}} \vartheta[\hat{W}_t^d(w_1, \theta^{t-1}, \cdot)] d\mu^{t-1}. \tag{49}$$

The relationship clearly holds for $t = 2$ since

$$\int_{\mathcal{L}^\Theta} \vartheta[\mathbf{w}(\cdot)] d\tilde{\varphi}_2 = \int_\Theta \vartheta[\tilde{w}_2(w_1, \theta_0, \theta_1, \cdot)] \Pi(\theta_0, d\theta_1)$$

$$= \int_\Theta \vartheta[\hat{W}_2^d(w_1, \theta_1, \cdot)] \mu^1(\theta_0, d\theta_1).$$

where the first line follows from the definition of $\tilde{\varphi}_2$ and the second one follows from the definition of $\hat{W}_2^d$. For $t > 2$ one has

$$\int_{\mathcal{L}^\Theta} \vartheta[\mathbf{w}(\cdot)] d\tilde{\varphi}_{t+1} = \int_{\Theta^2 \times \mathcal{L}^\Theta} \vartheta\{\tilde{w}_{t+1}[\mathbf{w}(\cdot), \theta_{t-1}, \theta_t, \cdot]\} d\tilde{\varphi}_t d\mu_{t-1}^t$$

$$= \int_{\Theta^t} \vartheta\{\tilde{w}_{t+1}[\hat{W}_t^d(w_1, \theta^{t-1}, \cdot), \theta_{t-1}, \theta_t, \cdot]\} d\mu^t$$

$$= \int_{\Theta^t} \vartheta\{\hat{W}_t^d(w_1, \theta^t, \cdot)\} d\mu^t$$

where the first line follows from definition of $\tilde{\varphi}_{t+1}$, the second one from the induction hypothesis and the last one from the definition of $\hat{W}_t^d$. Applying the equality for $\vartheta[\mathbf{w}(\cdot)] = \int_{\Theta^2} \tilde{c}_t[\mathbf{w}(\cdot), \theta_{t-1}, \theta_t] d\mu_{t-1}^t$ one gets that

$$\int_{\Theta^2 \times \mathcal{L}^\Theta} \tilde{c}_t[\mathbf{w}(\cdot), \theta_{t-1}, \theta_t] d\tilde{\varphi}_t d\mu_{t-1}^t = \int_{\Theta^t} \tilde{c}_t[\hat{W}_t^d(w_1, \theta^{t-1}, \cdot), \theta_{t-1}, \theta_t] d\mu^t$$

$$= \int_{\Theta^t} C_t^d(w_1, \theta^t) d\mu^t.$$

Applying the inequality again for $\vartheta[\mathbf{w}(\cdot)] = \int_{\Theta^2} \tilde{y}_t[\mathbf{w}(\cdot), \theta_{t-1}, \theta_t] d\theta_{t-1} d\mu_{t-1}^t$ one gets a similar equality. Substracting both equations, discounting by Q_t and summing over all t yields that the present value of resources is the same for both the allocation rule $\tilde{z}$ and the allocation Z^d.

ii.) I first define a candidate allocation rule $\tilde{z}$. By Lemma (12) one can take Z^d such that it satisfies the equal treatment property. For any function $\mathbf{w}(\cdot) \in \mathcal{L}^\Theta$ and $\theta_- \in \Theta$ construct a set

$$H_t[\mathbf{w}(\cdot), \theta_-] = \{\hat{\theta}^{t-1} \in \Theta^{t-1} : W_t(Z^d, \hat{\theta}^{t-1}, \cdot) = \mathbf{w}(\cdot), \theta_- = \hat{\theta}_{t-1}\}$$

for $t > 1$. H_t is the set of all histories and initial utility entitlements such that the promised utility function is $\mathbf{w}(\cdot)$ and last period shock was truthfully reported to be θ_-. If the set $H_t[\mathbf{w}(\cdot), \theta_-]$ is empty, set $\tilde{y}_t(\mathbf{w}(\cdot), \theta_{t-1}, \theta_-) = 0$, $\tilde{c}_t(\mathbf{w}(\cdot), \theta_{t-1}, \theta_-) = U^{-1}[(1-\beta)w(\theta_-), 0]$ and $\tilde{w}_{t+1}(\mathbf{w}(\cdot), \theta_{t-1}, \hat{\theta}, \cdot) = \mathbf{w}(\cdot)$ for all $\hat{\theta}_t \in \Theta$. If not, let

$$\tilde{c}_t(\mathbf{w}(\cdot), \theta_{t-1}, \hat{\theta}_t) = C_t^d(\hat{\theta}^{t-1}, \hat{\theta}_t) \text{ for all } \hat{\theta}^{t-1} \in H_t[\mathbf{w}(\cdot), \theta_{t-1}]$$

$$\tilde{y}_t(\mathbf{w}(\cdot), \theta_{t-1}, \hat{\theta}_t) = Y_t^d(\hat{\theta}^{t-1}, \hat{\theta}_t) \text{ for all } \hat{\theta}^{t-1} \in H_t[\mathbf{w}(\cdot), \theta_{t-1}]$$

$$\tilde{w}_{t+1}(\mathbf{w}(\cdot), \theta_{t-1}, \hat{\theta}_t, \theta_t) = W_{t+1}[Z^d, (\hat{\theta}^{t-1}, \hat{\theta}_t), \theta_t] \text{ for all } \hat{\theta}^{t-1} \in H_t[\mathbf{w}(\cdot), \theta_{t-1}]$$

In the first period, set $\tilde{c}_1(w_1, \hat{\theta}_1) = C_1^d(\hat{\theta}_1)$, $\tilde{y}_1(w_1, \hat{\theta}_1) = Y_1^d(\hat{\theta}_1)$ and $\tilde{w}_2(w_1, \hat{\theta}_1, \theta_1) = W_2(Z, \hat{\theta}_1, \theta_1)$ if $w_1 = W(Z, \theta_0)$. Otherwise, set $\tilde{y}_1(w_1, \hat{\theta}_1) = 0$, $\tilde{c}_1(w_1, \hat{\theta}_1) = U^{-1}[(1-\beta)w, 0]$, and $\tilde{w}_2(w_1, \hat{\theta}_1, \theta_1) = w$.

One needs to show that $\tilde{z}$ satisfies the promise keeping constraints (5) and (6), the temporary incentive compatibility constraint (7), that $\tilde{w}_{t+1}$ belongs to set $\tilde{\mathcal{U}}^*$ and that $\tilde{z}$ implies the same present value of resources as Z^d.

(iia) Promise keeping. For any w_1 and all $\hat{\theta}^{t-1} \in H_t[\mathbf{w}(\cdot), \theta_{t-1}]$,

$$
\begin{aligned}
\mathbf{w}(\cdot) &= W_t(Z^d, \hat{\theta}^{t-1}, \cdot) \\
&= \int_\Theta \{ U[C_t^d(w_1, \hat{\theta}^{t-1}, \theta_t), Y_t^d(w_1, \hat{\theta}^{t-1}, \theta_t), \theta_t] + \beta W_{t+1}[Z^d, (\hat{\theta}^{t-1}, \theta_t), \theta_t] \} \Pi(\cdot, d\theta_t) \\
&= \int_\Theta \{ U\{\tilde{c}_t[\mathbf{w}(\cdot), \theta_{t-1}, \theta_t], \tilde{y}_t[\mathbf{w}(\cdot), \theta_{t-1}, \theta_t], \theta_t\} + \beta \tilde{w}_{t+1}[\mathbf{w}(\cdot), \theta_{t-1}, \theta_t, \theta_t] \} \Pi(\cdot, d\theta_t).
\end{aligned}
$$

The first equality uses the definition of H_t, the second one property (35) and the third one the definition of H_t again. Hence $\tilde{z}$ satisfies (5) for any $t > 1$. For the first period one gets that

$$
w_1 = \int_\Theta \{ U[\tilde{c}_1(w_1, \theta_1), \tilde{y}_1(w_1, \theta_1), \theta_1] + \beta \tilde{w}_2(w_1, \theta_1, \theta_1) \} \Pi(\theta_0, d\theta_1),
$$

which is (6). Hence, $\tilde{z}$ satisfies the promise keeping constraint for all periods.

(iib) Incentive compatibility. Take any $w_1, \hat{\theta}^{t-1} \in H_t(\mathbf{w}(\cdot), \theta_{t-1})$. For any $\hat{\theta}_t \in \Theta$,

$$
\begin{aligned}
&U[\tilde{c}_t(\mathbf{w}(\cdot), \theta_{t-1}, \theta_t), \tilde{y}_t(\mathbf{w}(\cdot), \theta_{t-1}, \theta_t), \theta_t] + \beta \tilde{w}_{t+1}[\mathbf{w}(\cdot), \theta_{t-1}, \theta_t, \theta_t] \\
={}& U[C_t^d(w_1, (\hat{\theta}^{t-1}, \theta_t)), Y_t^d(w_1, (\hat{\theta}^{t-1}, \theta_t)), \theta_t] + \beta W_{t+1}[Z^d, (\hat{\theta}^{t-1}, \theta_t), \theta_t] \\
\geq{}& U[C_t^d(w_1, (\hat{\theta}^{t-1}, \hat{\theta}_t)), Y_t^d(w_1, (\hat{\theta}^{t-1}, \hat{\theta}_t)), \theta_t] + \beta W_{t+1}[Z^d, (\hat{\theta}^{t-1}, \hat{\theta}_t), \theta_t] \\
={}& U[\tilde{c}_t(\mathbf{w}(\cdot), \theta_{t-1}, \hat{\theta}_t), \tilde{y}_t(\mathbf{w}(\cdot), \theta_{t-1}, \hat{\theta}_t), \theta_t] + \beta \tilde{w}_{t+1}[\mathbf{w}(\cdot), \theta_{t-1}, \hat{\theta}_t, \theta_t],
\end{aligned}
$$

where the equalities follow from the definition of $\tilde{z}$ and the inequality follows from Lemma (4). Thus, $\tilde{z}$ is incentive compatible. Similar arguments show that $\tilde{z}$ satisfies the incentive compatibility constraint in period $t = 1$.

(iic) Inclusion in $\tilde{\mathcal{U}}^*$. I will now show that $\tilde{w}_{t+1}(\mathbf{w}(\cdot), \theta_{t-1}, \theta_t, \cdot) \in \tilde{\mathcal{U}}^*$ for all t, all $\mathbf{w}(\cdot)\mathcal{L}^\Theta$, all $(\theta_{t-1}, \theta_t) \in \Theta^2$.

Fix T. Let $\tilde{\mathcal{U}}_T^T = \mathcal{L}^\Theta$. Define for $t = 1..T$

$$\tilde{\mathcal{U}}_{t-1}^T = \{\mathbf{w}(\cdot) \in \mathcal{L}^\Theta : \exists \tilde{c} : \Theta \to \mathcal{R}_+, \tilde{y} : \Theta \to \mathcal{R}_+, \tilde{w}' : \Theta^2 \to \mathcal{R}$$

such that

$$\mathbf{w}(\cdot) = \int_\Theta [U(\tilde{c}(\theta), \tilde{y}(\theta), \theta) + \beta \tilde{w}'(\theta, \theta)] \Pi(\cdot, d\theta)$$

$$U(\tilde{c}(\theta), \tilde{y}(\theta), \theta) + \beta \tilde{w}'(\theta, \theta) \geq U(\tilde{c}(\hat{\theta}), \tilde{y}(\hat{\theta}), \theta) + \beta \tilde{w}'(\hat{\theta}, \theta) \qquad \forall \hat{\theta} \text{ s.t. } (\tilde{y}(\hat{\theta}), \theta) \in \mathcal{Y}, \text{ all } \theta$$

$$\tilde{w}'(\theta, \cdot) \in \tilde{\mathcal{U}}_t^T \}$$

It follows from parts (iia) and (iib) that $\tilde{w}_T'(\mathbf{w}(\cdot), \theta_{t-1}, \theta_t, \cdot) \in \tilde{\mathcal{U}}_{T-1}^T$ for all $\mathbf{w}(\cdot) \in \mathcal{L}^\Theta$, all $(\theta_{t-1}, \theta_t) \in \Theta^2$.[20]An induction argument implies that $\tilde{w}_{t+1}(\mathbf{w}(\cdot), \theta_{t-1}, \theta_t, \cdot) \in \tilde{\mathcal{U}}_t^T$, all $\mathbf{w}(\cdot) \in \mathcal{L}^\Theta$, all $(\theta_{t-1}, \theta_t) \in \Theta^2$, all $t = 1...T$.

The proof is complete when one shows that $\lim_{T \to \infty} \tilde{\mathcal{U}}_t^T = \tilde{\mathcal{U}}^*$ for all t. To show this, note first that $\tilde{\mathcal{U}}_t^T = \tilde{\mathcal{U}}_1^{T-t+1}$ by construction. Thus, it is enough to show that $\lim_{T \to \infty} \tilde{\mathcal{U}}_1^T = \tilde{\mathcal{U}}$. This follows from the fact that $\tilde{\mathcal{U}}^*$ is a fixed point of (8).

(iid) Resource clearing. A proof analogous to (ic) shows that the allocation rule $\tilde{z}$ implies the same present value of the resource costs as Z^d. ∎

Theorem 14 *i) Suppose that $\tilde{z}$ is an allocation rule and that $\tilde{z} \in \tilde{\mathfrak{z}}(w_1, \theta_0)$. Let $e = \tilde{E}^q(\tilde{z}, w_1, \theta_0)$ be the implied resource costs. Then the allocation Z that is generated by the allocation rule satisfies $Z \in \mathfrak{Z}^q(e, \theta_0)$ and delivers lifetime utility $W(Z, \theta_0) = w_1$.*

ii) Suppose that Z is an allocation and that $Z \in \mathfrak{Z}^q(e, \theta_0)$. Let $w_1 = W(Z, \theta_0)$ be the lifetime utility it delivers. Then there exists an allocation rule $\tilde{z} \in \tilde{\mathfrak{z}}(w_1, \theta_0)$ that has the implied resource costs $\tilde{E}^q(\tilde{z}, w_1, \theta_0) = e$.

Proof of Theorem (14).

i) By Lemma (13) part (i) there exists an allocation in the sequence cost minimization problem Z^d that satisfies (4) and (41). By Lemma (13) part (ic) the present value of resources implied by Z^d is $\tilde{E}^q(\tilde{z}, w_1, \theta_0)$.

---

[20]For instance, $\tilde{c}_T(\mathbf{w}(\cdot), \theta_{T-1}, \cdot)$, $\tilde{y}_T(\mathbf{w}(\cdot), \theta_{T-1}, \cdot)$.and $\tilde{w}_{T+1}'(\mathbf{w}(\cdot), \theta_{T-1}, \theta_T, \cdot)$ satisfy the requirements of the definition of $\tilde{\mathcal{U}}_{T-1}^T$.

Define $Z(\tilde{E}^q(\tilde{z}, w_1, \theta_0)) = Z^d(w_1)$. Then Z is generated by $\tilde{z}$, satisfies (4), and implies the same present value of costs as Z^d, which are e^d. Hence $Z \in \mathfrak{Z}^q(e^d, \theta_0)$. By construction, $W[Z(e^d), \theta_0] = W[Z^d(w_1), \theta_0] = w_1$.

ii) Conversely, construct an allocation in the sequence cost minimization problem by $Z^d[W(Z, \theta_0)] = Z(e)$. It is straightforward to see that both have the same present value of costs e, that Z^d satisfies (4) and that Z^d delivers promised utility $W(Z, \theta_0)$. By Lemma (13) part (ii) there exists an allocation rule $\tilde{z} \in \tilde{\mathfrak{z}}(W(Z, \theta_0), \theta_0)$ and has the implied resource costs $\tilde{E}^q(\tilde{z}, W(Z, \theta_0), \theta_0) = e$. ∎