Inequality and Social Discounting*

Emmanuel Farhi
MIT
efarhi@mit.edu

Iván Werning
MIT and NBER
iwerning@mit.edu

October 5, 2006 (1:24pm)

* For useful comments and suggestions we thank the editor, Nancy Stokey, two anonymous referees, as well as Daron Acemoglu, Laurence Ales, Fernando Alvarez, George-Marios Angeletos, Abhijit Banerjee, Gary Becker, Bruno Biais, Olivier Blanchard, Ricardo Caballero, Francesca Carapella, Dean Corbae, Mikhail Golosov, Bengt Holmstrom, Chad Jones, Narayana Kocherlakota, Robert Lucas, Pricila Maziero, Casey Mulligan, Roger Myerson, Chris Phelan, Thomas Philippon, Gilles Saint-Paul, Robert Shimer, Jean Tirole and various seminar and conference participants This work begun sparked by a seminar presentation by Chris Phelan at MIT. We also gained insight from a manuscript by Michael Sadler and the late Scott Freeman that Dean Corbae brought to our attention.

Abstract

We explore steady-state inequality in an intergenerational model with altruistically-linked individuals that experience privately observed taste shocks. When the welfare function only depends on the initial generation, efficiency requires immiseration: inequality grows without bound and everyone's consumption converges to zero. We study other efficient allocations, where the welfare function values future generations directly, placing a positive but vanishing weight on their welfare. The social discount factor is then higher than the private one and for any such difference we find that consumption exhibits mean-reversion and that a steady-state, cross-sectional distribution for consumption and welfare exists, with no one trapped at misery.

Introduction

Societies inevitably *choose* the inheritability of inequality. Some balance between equality of opportunity for newborns and incentives for altruistic parents is struck. In this paper, we explore how this balancing act plays out to determine long-run inequality.

In a model with infinitely-lived agents featuring private information, Atkeson and Lucas (1992) reached an extreme and surprising conclusion. They proved an *immiseration* result: inequality should be perfectly inheritable and rise steadily without bound, with everyone converging to absolute misery and a vanishing lucky fraction to bliss.[1] We depart minimally from this contribution by adopting the same positive economic model, but using a slightly different normative criterion. In a generational context, efficient allocations for infinitely-lived agents only characterize the instance where future generations are not considered *directly*, but only *indirectly* through the altruism of earlier ones. On the opposite side of the spectrum, Phelan (2006) proposes a social planner with equal weights on all generations that avoids the immiseration result because any allocation that leads everyone to misery actually minimizes the welfare criterion. Our interest here is in exploring a large class of Pareto-efficient allocations that also value future generations, but not equally. We

[1] This immiseration result is robust; it obtains invariably in partial equilibrium (Green, 1987; Thomas and Worrall, 1990), in general equilibrium (Atkeson and Lucas, 1992), across environments with moral-hazard regarding work effort or with private information regarding preferences or productivity (Aiyagari and Alvarez, 1995), and requires very weak assumptions on preferences (Phelan, 1998).

place a positive and vanishing Pareto weight on the expected welfare of future generations, which allows us to remain arbitrarily close to Atkeson-Lucas.[2]

Our welfare criterion captures the idea that it is desirable to insure the unborn against the luck of their ancestors, or, equivalently, insure the risk of which family they are born into—arguably, the biggest risk in life. Formally, we show that it is equivalent to a social discount factor that is higher than the private one. This relatively small change relative to Atkeson-Lucas produces a drastically different result: long-run inequality remains bounded, in the sense that a steady-state, cross-sectional distribution exists for consumption and welfare. At the steady state, there is social mobility and welfare remains above an endogenous lower bound, which is strictly better than misery. This outcome holds, however small the difference between social and private discounting.

Our positive economy is identical to Atkeson-Lucas' taste-shock setup. Each generation is composed of a continuum of individuals who live for one period and are altruistic towards a single descendant. There is a constant aggregate endowment of the only consumption good in each period. Individuals are ex-ante identical, but experience idiosyncratic shocks to preferences that are privately observed. Feasible allocations must be incentive compatible and must satisfy the aggregate resource constraint in all periods.

---

[2] To make the notion more familiar, note that in overlapping-generation models without altruistic links, all market equilibria that are Pareto efficient place positive direct weight on future generations. Bernheim (1989) pointed out that in extensions of these models that incorporate altruism many Pareto efficient allocations are not attainable by the market.

When only the welfare of the first generation is considered, the planning problem is equivalent to that of an economy with infinite-lived individuals. Intuitively, immiseration then results because rewards and punishments, required for incentives, are best delivered permanently to smooth dynastic consumption over time. As a result, the consumption process inherits a random-walk component that leads cross-sectional inequality to grow without bound. No steady-state, cross-sectional distribution with positive consumption exists. Moreover, indefinite spreading of the distribution is consistent with a constant aggregate endowment only if everyone's consumption converges to zero.

Interpreted in the intergenerational context, this solution requires a lock-step link between the welfare of parent and child. This perfect intergenerational transmission of welfare improves parental incentives, but it exposes future generations to the risk of their dynasty's history. Future descendants value insurance against the uncertainty of their ancestors' past shocks.

By contrast, when future generations are weighted in the social welfare function it remains optimal to link the fortunes of parents and children, but no longer in lock-step. Rewards and punishments are distributed over all future descendants, but in a front-loaded manner. This creates a mean-reverting tendency in consumption—instead of a random walk—that is strong enough to bound long-run inequality. The result is a steady-state distribution for the cross section of consumption and welfare, with no one at misery. Moreover, mean reversion ensures a form of social mobility, so that families rise and fall through the ranks incessantly.

It is worth emphasizing that our exercise is not predicated on any paternalistic concern that individuals do not discount the future appropriately. Rather, the difference between social and private discounting arises because the social welfare function gives direct weight to future generations. However, our formal analysis can be applied whatever the motivation, paternalistic or not, for a difference in social and private discounting. For example, Caplin and Leahy (2004) make a case for a higher social discount factor within a lifetime.

A methodological contribution of this paper is to reformulate the social planning problem recursively in a way that extends the ideas introduced by Spear and Srivastava (1987) to a general-equilibrium situation where private and social objectives potentially differ. We are able to reduce the dynamic program to a one-dimensional state variable, and our analysis and results heavily exploit the resulting Bellman equation.

The paper most closely related to ours is Phelan (2006), who considered a social planning problem with no discounting of the future. He shows that if a steady state for the planning problem exists then it must solve a static maximization problem, and that solutions to this problem have strictly positive inequality and social mobility. Our paper establishes the existence of a steady-state distribution for any difference in social and private discounting. Unlike the case with no discounting, there is no valid static problem, so our methods are necessarily quite different. Our work is also indirectly related to Sleet and Yeltekin (2004), who study a utilitarian planner that lacks commitment and always cares for the current generation only. The best equilibrium allocation *without* commitment is equivalent to the optimal one *with* commitment but with a more patient welfare criterion. Thus, our approach and results provide an indirect, but effective, way of characterizing the problem without commitment and establishing the existence of a steady-state distribution. This illustrates that lack of commitment or political economy considerations, provide a motivation for the positive Pareto weights that future generations command.

The rest of the paper is organized as follows. Section 1 contains some simple examples to illustrate why weighing future generations leads to a higher social discount factor, and why mean reverting forces emerge from any difference between social and private discounting. Section 2 introduces the economic environment and sets up the social planning problem. In Section 3, we develop a recursive version of the planning problem and establish its relation to the original formulation. The resulting Bellman equation is then used in Section 4 to characterize the mean-reversion in the solution. Section 5 contains the main results on the existence of a steady state. Section 6 offers some conclusions from the analysis. Proofs are contained in the Appendix.

1. Social Discounting and Mean Reversion

In this section, we preview the main forces at work in the full model using a simple deterministic example. We first explain why weighing future generations maps into lower social discounting. We then show how this affects the optimal inheritability of welfare across generations. Finally, we relate the latter to the mean-reversion force, which guarantees a steady state distribution with social mobility in the full model. Our discussion also provides intuition for the immiseration result in Atkeson-Lucas.

Social Discounting. Imagine a two period deterministic economy. The parent is alive in the first period, $t = 0$, and replaced by a single child in the next, $t = 1$. The child derives utility from his own consumption, so that $v_1 = u(c_1)$. The parent cares about her own consumption but is also altruistic towards the child, so that her welfare is $v_0 = u(c_0) + \beta v_1 = u(c_0) + \beta u(c_1)$.

A welfare criterion that weighs both agents, and serves to trace out the Pareto frontier between v_0 and v_1, is $W \equiv v_0 + \alpha v_1$, for some weight $\alpha \geq 0$. Equivalently,

$$W = u(c_0) + (\beta + \alpha)u(c_1) = u(c_0) + \hat{\beta} u(c_1),$$

with the social discount factor given by $\hat{\beta} \equiv \beta + \alpha$.

The only difference between the welfare criterion and the objective of the parent is the rate of discounting. Social discounting depends on the weight on future generations α. When no direct weight is placed on children, so that $\alpha = 0$, social and private discounting coincide, $\hat{\beta} = \beta$, which is the case covered by Atkeson and Lucas (1992). Whenever children are counted directly in the welfare criterion, $\alpha > 0$, society discounts less than parents do privately, $\hat{\beta} > \beta$. The child's consumption it is a public good that both generations enjoy, so the utility from it gets more weight.

A Planning Problem. In Section 2 we show that the calculations above generalize to an infinite horizon economy and lead to an objective with more patient geometric discounting: $\sum_{t=0}^{\infty} \hat{\beta}^t u(c_t)$. We now consider a simple planning problem for such an infinite-horizon version.

Now, suppose there are two dynasties, A and B. In each period, a planner must divide a fixed endowment $2e$ between the two dynasties, giving $c_{A,t}$ to A and $c_{B,t} = 2e - c_{A,t}$ to B. Suppose that, for some reason, the heads of dynasties are promised differential treatment, so that the difference in their welfare must be Δ. The planner's problem is

$$\max_{\{c_{A,t}\}} \sum_{t=0}^{\infty} \hat{\beta}^t \left(\tfrac{1}{2}u(c_{A,t}) + \tfrac{1}{2}u(2e - c_{A,t})\right)$$

subject to

$$\sum_{t=0}^{\infty} \beta^t u(c_{A,t}) - \sum_{t=0}^{\infty} \beta^t u(2e - c_{A,t}) = \Delta.$$

The first-order conditions for an interior optima are

$$\frac{u'(c_{A,t})}{u'(c_{B,t})} = \frac{1 + \lambda(\beta/\hat{\beta})^t}{1 - \lambda(\beta/\hat{\beta})^t} \qquad t = 0, 1, \ldots,$$

where λ is the Lagrange multipliers on the constraint.

Imperfect Inheritability. Suppose that the founder of dynasty A has been promised higher welfare $\Delta > 0$ so that $\lambda > 0$. The first-order condition then reveals that every member of dynasty A enjoys higher consumption, $c_{A,t} > c_{B,t}$. If $\hat{\beta} = \beta$, as in Atkeson and Lucas (1992), consumption is constant over time for both groups, and initial differences persist forever. The unequal promises to the first generation have a permanent impact on their descendants. The inheritability of welfare across generations is perfect: the consumption and welfare of the child moves one-to-one with the parent's welfare.

In contrast, when $\hat{\beta} > \beta$ the consumption differences between the two dynasties shrinks over time. Consumption declines across generations for group A, and rises for the group B. The inheritability of welfare across generations is imperfect: a child's consumption rises less than one-for-one with the parent's. Indeed, initial differences completely vanishes asymptotically—initial inequality dies out. Figure 1 illustrates these dynamics for consumption.

In this simple deterministic example, initial inequality Δ was taken as given. However, the model below generates similar responses in future consumption in order to provide incentives. That is, the dynamics after a shock are similar to those illustrated here, and Figure 1 can be loosely interpreted as an "impulse response" function. If $\hat{\beta} = \beta$, shocks have a permanent effects on inequality; consumption inherits a random-walk-like property. If the social welfare function puts weight directly on all generations so that $\hat{\beta} > \beta$, then there is mean reversion: the effects of shocks decay over time. Moreover the asymptotic rate of decay is $\hat{\beta}/\beta$.[3]

In our deterministic example, we found that as long as $\hat{\beta} > \beta$ inequality vanishes in the long run—the unique steady state has no inequality. However, in our full model with ongoing taste shocks, inequality remains positive in the long-run steady state. The mean-reverting force illustrated here ensures that long-run inequality remains bounded. By contrast, when $\hat{\beta} = \beta$, as in Atkeson and Lucas (1992), there is no mean-reversion. The random walk like property of the allocation implies that shocks accumulate indefinitely and inequality increases without bound.

2. An Intergenerational Insurance Problem

At any point in time, the economy is populated by a continuum of individuals who have identical preferences, live for one period, and are replaced by a single descendant in the next. Parents born in period t are altruistic towards their only child and their welfare v_t satisfies

$$v_t = \mathbb{E}_{t-1}\left[\theta_t U(c_t) + \beta v_{t+1}\right],$$

[3] A simple Taylor expansion yields $c_{A,t}/c_{B,t} - 1 \simeq \frac{u'(e)}{-eu''(e)}(\lambda^A - \lambda^B)(\beta/\hat{\beta})^t$. Hence, the relative difference in consumption goes to zero at asymptotic rate $\beta/\hat{\beta}$.

where $c_t \geq 0$ is the parent's own consumption and $\beta \in (0,1)$ is the altruistic weight placed on the descendant's welfare v_{t+1}, and $\theta_t \in \Theta$ is a taste shock that is assumed to be identically and independently distributed across individuals and time. We assume the following throughout.

Assumption 1 *(a) the set of taste shocks Θ is finite; (b) the utility function $U(c)$ is concave and continuously differentiable for $c > 0$ with $\lim_{c\to 0} U'(c) = \infty$ and $\lim_{c\to\infty} U'(c) = 0$.*

This specification of altruism is consistent with individuals having a preference over the entire future consumption of their dynasty given by

$$v_t = \sum_{s=0}^{\infty} \beta^s \, \mathbb{E}_{t-1}\left[\theta_{t+s} U\left(c_{t+s}\right)\right]. \tag{1}$$

In each period, a resource constraint limits aggregate consumption to be no greater than some constant aggregate endowment $e > 0$.

The following notation and conventions will be used. We refer to $u_t = U(c_t)$ as utility and the discounted, expected utility v_t as welfare. Let $p(\theta)$ denote the probability of $\theta \in \Theta$; we adopt the normalization that $\mathbb{E}[\theta] = \sum_{\theta\in\Theta} \theta p(\theta) = 1$. The cost function $c(u)$ is defined as the inverse of the utility function $c \equiv U^{-1}$. Dynastic welfare v_t belongs to the set $u(\mathbb{R}_+)/(1-\beta)$ with extremes $\underline{v} \equiv U(0)/(1-\beta)$ and $\bar{v} \equiv \lim_{c\to\infty} U(c)/(1-\beta)$, which may be finite or infinite.

Taste shock realizations are privately observed, so any mechanism for allocating consumption must be incentive compatible. The revelation principle allows us to restrict attention to mechanisms that rely on truthful reports of these shocks. Thus, each dynasty faces a sequence of consumption functions $\{c_t\}$, where $c_t(\theta^t)$ represents an individual's consumption after reporting the history $\theta^t \equiv (\theta_0, \theta_1, \ldots, \theta_t)$. It is more convenient to work with the implied allocation for utility $\{u_t\}$ with $u_t(\theta^t) \equiv U(c_t(\theta^t))$. A dynasty's reporting strategy $\sigma \equiv \{\sigma_t\}$ is a sequence of functions $\sigma_t : \Theta^{t+1} \to \Theta$ that maps histories of shocks θ^t into a current report $\hat{\theta}_t$. Any strategy σ induces a history of reports $\sigma^t : \Theta^{t+1} \to \Theta^{t+1}$. We use σ^* to denote the truth-telling strategy with $\sigma_t^*(\theta^t) = \theta_t$ for all $\theta^t \in \Theta^{t+1}$.

Following Atkeson-Lucas, we identify each dynasty with a number v, which we interpret as its founder's welfare entitlement, $v_0 = v$. We assume that all dynasties with the same entitlement v receive the same treatment. We then let ψ denote a distribution of v across dynasties. An allocation is a sequence of functions $\{u_t^v\}$ for each v. For any given initial distribution of entitlements ψ and resources e, we say that an allocation $\{u_t^v\}$ is *feasible* if: (i) it delivers expected utility of v to all initial dynasties entitled to v:

$$v = \sum_{t=0}^{\infty} \sum_{\theta^t \in \Theta^{t+1}} \beta^t \theta_t u_t^v\big(\sigma^t(\theta^t)\big) \Pr(\theta^t) \tag{2}$$

(ii) it is incentive compatible for all v:

$$\sum_{t=0}^{\infty} \sum_{\theta^t \in \Theta^{t+1}} \beta^t \theta_t \Big(u_t^v \big(\sigma^t(\theta^t)\big) - u_t^v \big(\sigma^t(\theta^t)\big) \Big) \Pr(\theta^t) \geq 0 \quad \text{for all } \sigma; \tag{3}$$

and (iii) total consumption does not exceed the fixed endowment e in all periods:

$$\int \sum_{\theta^t \in \Theta^{t+1}} c\big(u_t^v(\theta^t)\big) \Pr(\theta^t) \, d\psi(v) \leq e \qquad t = 0, 1, \ldots \tag{4}$$

Define $e^*(\psi)$ to be the lowest endowment e such that there exists an allocation satisfying (2)–(4), which is precisely the efficiency problem studied in Atkeson and Lucas (1992).

Social Discounting. We have adopted the same preferences, technology and informational assumptions as in Atkeson and Lucas (1992). Our only departure is to introduce the planning objective

$$\sum_{t=0}^{\infty} \hat{\beta}^t \, \mathbb{E}_{-1}[\theta_t U(c_t)], \tag{5}$$

for each dynasty, which is equivalent to the preferences in (1), except for the discount factor $\hat{\beta} > \beta$. Our motivation for this objective is that it can be derived from a welfare criterion that places direct weight on the welfare of future generations. To see this, consider the sum of expected utilities using strictly positive weights $\{\alpha_t\}$:

$$\sum_{t=0}^{\infty} \alpha_t \, \mathbb{E}_{-1} v_t = \sum_{t=0}^{\infty} \delta_t \, \mathbb{E}_{-1} \left[\theta_t U\left(c_t\right)\right], \tag{6}$$

where $\delta_t \equiv \beta^t \alpha_0 + \beta^{t-1}\alpha_1 + \cdots + \beta\alpha_{t-1} + \alpha_t$. Then the discount factor satisfies

$$\frac{\delta_{t+1}}{\delta_t} = \beta + \frac{\alpha_{t+1}}{\delta_t} > \beta,$$

so that social preferences are more patient. Future generations are already indirectly valued through the altruism of the current generation. If, in addition, they are also *directly* included in the welfare function the social discount factor must be higher than β.[4]

In particular, for geometric Pareto weights $\alpha_t = \hat{\beta}^t$ with $\hat{\beta} > \beta$:

$$\sum_{t=1}^{\infty} \alpha_t \, \mathbb{E}_{-1} v_t = \frac{1}{\hat{\beta} - \beta} \left(\sum_{t=0}^{\infty} \hat{\beta}^t \, \mathbb{E}_{-1} [\, \theta_t U(c_t) \,] - v_0 \right). \tag{7}$$

[4] Bernheim (1989) performs similar intergenerational discount factor calculations in his welfare analysis of a deterministic dynastic saving model. Caplin and Leahy (2004) argue that these ideas also apply to intra-personal discounting within a lifetime, leading to a social discount factor that is greater than the private one not only across generations, but within generations as well.

The first term is identical to the expression in (5); when initial welfare promises v_0 for the founding generation are given the second term is a constant.

Planning Problem. Define the social optimum as a feasible allocation that maximizes the integral of the welfare function (5) with respect to distribution ψ. That is, the *social planning problem* given an initial distribution of welfare entitlements ψ and an endowment level e is

$$S(\psi; e) \equiv \sup_{\{u_t^v\}} \int \sum_{t=0}^{\infty} \sum_{\theta^t \in \Theta^{t+1}} \hat{\beta}^t \theta_t u_t(\theta^t) \Pr(\theta^t)\, d\psi(v) \quad \text{subject to (2)–(4).} \tag{8}$$

The constraint set is nonempty as long as $e \geq e^*(\psi)$. If $e = e^*(\psi)$, then the only feasible allocation is the one characterized by Atkeson-Lucas; we study cases with $e > e^*(\psi)$.

The way we have defined the social planning problem imposes that initial welfare entitlements v be delivered exactly, in the sense that the promise-keeping constraints (2) are equalities, instead of inequalities. Alternatively, suppose that the founder of each dynasty is indexed by some minimum welfare entitlement $\tilde{v}$, with distribution $\tilde{\psi}$. The *Pareto problem* maximizes the average welfare criterion (7) subject to delivering $\tilde{v}$ or more to the founders and incentive compatibility. The two problems are related: the solution to the Pareto problem solves the social planning problem for some distribution ψ that first-order stochastically dominates $\tilde{\psi}$. In particular, it chooses ψ to maximize

$$S(\psi; e) - \int v \, d\psi(v)$$

subject to $\psi(v) \leq \tilde{\psi}(v)$ for all v. In general, depending on $\tilde{\psi}$, the constraints of delivering the initial welfare entitlements $\tilde{v}$ or more may be slack, so that $\psi \neq \tilde{\psi}$ may be optimal. However, we shall show that setting $\psi = \tilde{\psi}$ is optimal for initial distributions of entitlements $\tilde{\psi}$ that are steady states, as defined below. Our strategy is to work with the social planning problem and then return to its connection with the Pareto problem.

Steady States. The social planning problem takes the initial distribution of welfare entitlements ψ as given. In later periods the current cross-sectional distribution of continuation welfare ψ_t is a sufficient statistic for the remaining planning problem: the social planning problem is recursive with state variable ψ_t. It follows that the solution to the planning problem from any period t onward, $\{u^v_{t+s}\}_{s=0}^{\infty}$, is a time independent function of the current distribution ψ_t which evolves according to a stationary recursion $\psi_{t+1} = \Psi\psi_t$, for some fixed mapping Ψ.

Our focus is on distributions of welfare entitlements ψ^* such that the solution to the planning problem features, in each period, a cross-sectional distribution of continuation utilities v_t that is also distributed according to ψ^*. In this case, the cross-sectional distribution of consumption also replicates itself over time. We term any distribution of entitlements ψ^* with these properties a *steady state*. A steady state corresponds to a fixed point of this mapping, $\psi^* = \Psi\psi^*$.

In the Atkeson-Lucas case, with $\beta = \hat{\beta}$, the non-existence of a steady state with positive consumption is a consequence of the immiseration result: starting from any non-trivial initial distribution ψ and resources $e = e^*(\psi)$ the sequence of distributions converges weakly to the distribution having full mass at misery $\underline{v} = U(0)/(1-\beta)$, with zero consumption for everyone. We seek non-trivial steady states ψ^* that exhaust a strictly positive aggregate endowment e in all periods.

Using the entire distribution ψ_t as a state variable is one way to approach the planning problem. Indeed, this is the method adopted by Atkeson and Lucas (1992). They were able to keep the analysis manageable, despite the large dimensionality of the state variable, by exploiting the homogeneity of the problem with constant relative risk aversion (CRRA) preferences. In contrast, even in the CRRA case, our model lacks this homogeneity, making such a direct approach intractable.[5] Consequently, in the next section, we attack the problem differently, using a dynamic program with a one-dimensional state variable.[6] The idea is that the continuation welfare v_t of each dynasty follows a Markov process, and that steady states are invariant distributions of this process.

---

[5] Indeed, the homogeneity in Atkeson-Lucas is intimately linked to their immiseration result: it implies that rewards and punishments are permanent and delivered by shifting the entire sequence of consumption up or down multiplicatively. In our case, even with CRRA preferences, homogeneity breaks down and mean reversion emerges, preventing immiseration.

[6] A similar approach is taken in Atkeson and Lucas (1995), Aiyagari and Alvarez (1995), and others.

3. A Bellman Equation

In this section we approach the planning problem by studying a relaxed version of it, whose solution coincides with that of the original problem at steady states. The relaxed problem has two important advantages. First, it can be solved by studying a set of subproblems, one for each dynasty, thereby avoiding the need to keep track of the entire distribution ψ_t in the population. Second, each subproblem admits a one-dimensional recursive formulation, which we are able to characterize quite sharply. We believe that the general approach we develop here may be useful in other contexts.

Define a *relaxed planning problem* by replacing the sequence of resource constraints (4) in the social planning problem (8) with a single intertemporal constraint

$$\int \sum_{t=0}^{\infty} Q_t \sum_{\theta^t \in \Theta^{t+1}} c\big(u_t^v(\theta^t)\big) \Pr(\theta^t)\, d\psi(v) \leq e \sum_{t=0}^{\infty} Q_t, \tag{9}$$

for some positive sequence $\{Q_t\}$ with $\sum_{t=0}^{\infty} Q_t < \infty$. One can interpret this problem as representing a small open economy facing intertemporal prices $\{Q_t\}$. The original and relaxed versions of the social planning problem are related in that any solution to the latter that happens to satisfy the resource constraints (4) is also a solution to the former.[7,8] It follows that any steady-state solution

[7] This is related to the decentralization result in Atkeson and Lucas (1992, Theorem 1, Section 7), although they do not use it, as we do here, to characterize the solution.

[8] Since the problem is convex, a Lagrangian argument establishes the converse: there must exist some positive

to the relaxed problem is a steady-state solution to the original one, since at a steady state the intertemporal constraint (9) implies the resource constraints (4). A steady state requires $Q_t = \hat{\beta}^t$.

Consider then the intertemporal resource constraint (9) with $Q_t = \hat{\beta}^t$:

$$\int \sum_{t=0}^{\infty} \hat{\beta}^t \sum_{\theta^t \in \Theta^{t+1}} c\big(u_t^v(\theta^t)\big) \Pr(\theta^t)\, d\psi(v) \le e \sum_{t=0}^{\infty} \hat{\beta}^t. \tag{10}$$

Letting η denote the multiplier on this constraint we form the Lagrangian (omitting the constant term due to e)

$$L \equiv \int L^v \, d\psi(v) \qquad \text{where} \qquad L^v \equiv \sum_{t=0}^{\infty} \sum_{\theta^t \in \Theta^{t+1}} \hat{\beta}^t \Big(\theta_t u_t^v(\theta^t) - \eta c\big(u_t^v(\theta^t)\big)\Big) \Pr(\theta^t). \tag{11}$$

We study the maximization of L subject to (2) and (3).

The advantage of working with the Lagrangian is that maximizing L is equivalent to the point-

sequence $\{Q_t\}$ such that the solution to the original social planning problem also solves the relaxed problem. This is analogous to the second theorem of welfare economics for our environment. However, we will not require this converse result to construct a steady state solution.

wise optimization, for each v, of the subproblem:

$$k(v) \equiv \sup_{\{u_t^v\}} L^v \quad \text{subject to (2) and (3).} \tag{12}$$

We call this subproblem, for a given v and η, the *component planning problem*. Its connection with the relaxed problem is that for any e there exists a positive multiplier η such that an allocation $\{u_t^v\}$ solves the relaxed planning problem with endowment e if and only if for each v the allocation $\{u_t^v\}$ solves the component planning problem given v and η.

Our first result characterizes the value function $k(v)$, defined from a sequence problem, showing that it satisfies a Bellman equation.

Theorem 1 *The value function of the component planning problem $k(v)$ defined by equation (12) is continuous, concave, and satisfies the Bellman equation*

$$k(v) = \max_{u,w} \mathbb{E}\big[\theta u(\theta) - \eta c\big(u(\theta)\big) + \hat{\beta} k\big(w(\theta)\big)\big] \tag{13}$$

subject to

$$v = \mathbb{E}[\theta u(\theta) + \beta w(\theta)] \tag{14}$$

$$\theta u(\theta) + \beta w(\theta) \geq \theta u(\theta') + \beta w(\theta') \qquad \textit{for all} \quad \theta, \theta' \in \Theta. \tag{15}$$

This recursive formulation imposes a promise-keeping constraint (14) and an incentive constraint (15). Intuitively, the latter rules out one-shot deviations from truth-telling, guaranteeing that telling the truth today is optimal if the truth is told in future periods. Of course, this is necessary to satisfy the full incentive-compatibility condition (3). The rest is implicitly taken care of in (13) by evaluating the value function, defined from the sequence problem, at the continuation welfare: for any given continuation value $w(\theta)$, envision the planner in the next period solving the remaining sequence problem by selecting an entire allocation that is incentive compatible from then on; $k(w(\theta))$ represents the value to the planner of this continuation allocation. For any given v, a pair $u(\theta)$ and $w(\theta)$ that satisfies (14)–(15), pasted with the corresponding continuation allocations for each $w(\theta)$, describes an allocation that satisfies (2)–(3).

Among other things, Theorem 1 shows that the maximization on the right hand side of the Bellman equation is uniquely attained by some continuous policy functions $g^u(\theta, v)$ and $g^w(\theta, v)$, for u and w, respectively. We emphasize that these policy functions solve the maximization in the Bellman equation (13) using the value function $k(v)$ defined from the sequence problem (12).[9] For any initial welfare entitlement v, an allocation $\{u_t\}$ can then be generated from the policy

[9] If one assumes bounded utility, a standard contraction-mapping argument could be used to establish that the Bellman equation has a unique solution, which must then coincide with that defined from the sequence problem. However, we do not assume bounded utility and for our purposes such arguments are not required. Indeed, we never require solving fixed points of the Bellman equation or showing that it has a unique solution. We work directly with $k(v)$ defined from the sequence problem (12).

functions (g^u, g^w) by $u_t(\theta^t) = g^u(\theta_t, v_t(\theta^{t-1}))$, with $v_0 = v$ and $v_{t+1}(\theta^t) = g^w(\theta_t, v_t(\theta^{t-1}))$. Our next result provides a connection between allocations generated this way and solutions to the component planning problem (12).

Theorem 2 *For any (v, η), if an allocation $\{u_t\}$ attains the supremum in (12) then it is generated by (g^u, g^w). Conversely, if an allocation $\{u_t\}$ generated by (g^u, g^w) satisfies*

$$\limsup_{t\to\infty} \mathbb{E}_{-1}\beta^t v_t\big(\sigma^{t-1}(\theta^{t-1})\big) \geq 0 \qquad \textit{for all } \sigma,$$

then it attains the supremum in (12).

The first part of Theorem 2 implies either that the solution to the relaxed planning problem is generated by the policy functions, or that there is no solution at all. The second part of Theorem 2 shows that a solution can be guaranteed if we can verify a certain limiting condition. Note that this condition is trivially satisfied for all utility functions that are bounded below.[10]

[10] Theorem 2 involves various applications of versions of the Principle of Optimality. For example, for any given policy functions (g^u, g^w) and an initial value v_0, the individual dynasty faces a recursive dynamic programming problem with state variable v_t. Conditions (14) and (15) then amount to guessing and verifying a solution to the Bellman equation of the agent's problem—in particular, that the value function that satisfies the Bellman equation, with truth telling, is the identity function. However, one also needs to verify that this value function represents the true optimal value for the dynasty from the sequential problem. This verification is accomplished by the limiting condition of Theorem 2.

The next result establishes that $k(v)$ is differentiable and strictly concave with an interior peak.

Proposition 1 *The value function $k(v)$ is strictly concave; it is differentiable on the interior of its domain, with $\lim_{v \to \bar{v}} k'(v) = -\infty$. If utility is unbounded below, then $\lim_{v \to \underline{v}} k'(v) = 1$. Otherwise $\lim_{v \to \underline{v}} k'(v) = \infty$.*

The shape of the value function is important because mean-reversion occurs towards the interior peak, as we show next.

4. Mean Reversion

In this section we use the Bellman equation to characterize the solution to the planning problem and to show that it displays mean-reversion. Let λ be the multiplier on the the promise-keeping constraint (14) and let $\mu(\theta, \theta')$ be the multipliers on the incentive constraints (15). The first-order condition for $u(\theta)$ is

$$\Big(1 - \eta c'\big(g^u(\theta, v)\big)\Big)p(\theta) - \theta\lambda p(\theta) + \sum_{\theta' \in \Theta} \theta\mu(\theta, \theta') - \sum_{\theta' \in \Theta} \theta'\mu(\theta', \theta) \leq 0,$$

with equality if $g^u(\theta, v)$ is interior. Given the Inada conditions for $k(v)$ derived in Proposition 1, the solution for $w(\theta)$ must be interior and satisfy the first-order condition

$$\hat{\beta} k'\big(g^w(\theta, v)\big)p(\theta) - \beta\lambda p\big(\theta\big) + \beta \sum_{\theta' \in \Theta} \mu(\theta, \theta') - \beta \sum_{\theta' \in \Theta} \mu(\theta', \theta) = 0.$$

Incentive compatibility implies that $g^u(\theta, v)$ is nondecreasing as a function of θ; similarly, $g^w(\theta, v)$ is nonincreasing in θ. Using the envelope condition $k'(v) = \lambda$ and summing over θ, this becomes

$$\sum_{\theta \in \Theta} k'\big(g^w(\theta, v)\big)p(\theta) = \frac{\beta}{\hat{\beta}} k'(v). \tag{16}$$

In sequential notation, this condition is

$$\mathbb{E}_{t-1}\big[k'\big(v_{t+1}\big)\big] = \frac{\beta}{\hat{\beta}} k'\big(v_t\big), \tag{17}$$

where $\{v_t(\theta^{t-1})\}$ is generated by the policy function g^w. Since $\beta/\hat{\beta} < 1$ the Markov process $\{k'(v_t)\}$ regresses towards zero. By Proposition 1, the value function $k(v)$ has an interior maximum at v^*, where $k'(v^*) = 0$, that is strictly higher than misery $\underline{v}$. Reversion occurs towards this interior point.

Economically, this mean-reversion implies an interesting form of social mobility. Divide the population into two, those above and those below v^*. Then mobility is ensured between these groups: descendants of individuals with current welfare above v^* will eventually fall below it, and viceversa. This rise and fall of families illustrates one form of intergenerational mobility.

To provide incentives, the planner rewards the descendants of an individual reporting a low taste shock. Rewards can take two forms and it is optimal to makes use of both. The first is standard and involves spending more on a dynasty in present-value terms. The second is more subtle and exploits differences in preferences: it is to allow an adjustment in the pattern of consumption, for a given present value, in the direction preferred by individuals relative to the planner.[11] Since individuals are more impatient than the planner, this form of reward is delivered by tilting the consumption profile toward the present. Earlier consumption dates are used more intensively to provide incentives: rewards and punishments are front-loaded.

Our next result derives upper and lower bounds for the $\{k'(v_t)\}$ process.

Proposition 2 *There exists constants $\underline{\gamma} \leq \bar{\gamma}$ so that*

(a) *if utility is unbounded below then for all $\theta \in \Theta$:*

$$\underline{\gamma}\big(1 - k'(v)\big) + \left(1 - \frac{\beta}{\hat{\beta}}\right) \leq 1 - k'\big(g^w(\theta, v)\big) \leq \bar{\gamma}\big(1 - k'(v)\big) + \left(1 - \frac{\beta}{\hat{\beta}}\right) \tag{18}$$

[11] Some readers may recognize this last method as the time-honored system of rewards and punishments used by parents when conceding their child's favorite snack or reducing their TV-time. In these instances, the child values some goods more than the parent wishes, and the parent uses them to provide incentives.

(b) *if utility is bounded below, then for* $k'(v) \leq 1$*, the lower bound on* $1 - k'\big(g^w(\theta, v)\big)$ *in (18) holds; the upper bound holds for sufficiently high* v*. For* $k'(v) > 1$*,* $g^u(\theta, v) = U(0)$ *and* $g^w(\theta, v) = (v - U(0))/\beta > v$ *for all* $\theta \in \Theta$*.*

Moreover, in the limit $\underline{\gamma}, \bar{\gamma} \to \beta/\hat{\beta}$ *as* $\underline{\theta}, \bar{\theta} \to 1$*.*

These bounds are instrumental in proving the existence of an invariant distribution. They also illustrate a powerful tendency away from misery. For example, they imply that, when utility is unbounded below, continuation welfare $g^w(\theta, v)$ remains bounded even as $v \to -\infty$. No matter how much a parent is to be punished, his child is somewhat spared. Given the result that $\bar{\gamma} \to \beta/\hat{\beta}$ as $\underline{\theta}, \bar{\theta} \to 1$, the bounds also imply that the ergodic set for $1 - k'(v_t)$ is bounded above if the amplitude of taste shocks is not too wide. This, provides conditions for v_t to be bounded away from $\bar{v}$.

5. Existence of an Invariant Distribution

In this section we show that a steady-state invariant distribution exists. The proof relies on the mean-reversion in equation (17) and the bounds in Proposition 2.

Proposition 3 *The Markov process* $\{v_t\}$ *implied by* g^w *has an invariant distribution* ψ^* *with* $\int k'(v)\psi^*(v) = 0$ *and no mass at misery,* $\psi^*(\{\underline{v}\}) = 0$*, if any of the following holds: utility is unbounded below, utility is bounded above,* $\bar{\gamma} < 1$*, or* $\underline{\gamma} > 0$.

When combined with the first part of Theorem 2, this proposition leaves only two possibilities. Either the social planning problem admits a steady-state invariant distribution, or no solution exists. This contrasts with the Atkeson-Lucas case, with $\beta = \hat{\beta}$, where a solution exists but does not admit a steady state. Later in this section we verify the second part of Theorem 2 to confirm that a solution to the planning problem can be guaranteed and a steady state exists.

Our Bellman equation also provides an efficient way of solving the planning problem. We illustrate this with two examples, one analytical and another numerical.

Example 1. Suppose utility is CRRA with $\sigma = 1/2$, so that $U(c) = 2^{1/2}c^{1/2}$, so that $c \geq 0$ and $c(u) = u^2/2$ for $u \geq 0$. For $\beta = \hat{\beta}$ Atkeson-Lucas show that the optimum involves consumption inequality growing without bound and leading to immiseration.

Consider the Bellman equation for the problem that ignores the nonnegativity constraints on u and w:

$$k(v) = \max_{u,w} \mathbb{E}\left[\theta u(\theta) - \tfrac{\eta}{2} u(\theta)^2 + \hat{\beta} k\big(w(\theta)\big)\right]$$

subject to (14) and (15). This is a linear-quadratic dynamic programming problem, so the value function $k(v)$ is quadratic function and the policy functions are linear in v:

$$g^u(\theta, v) = \gamma_1^u(\theta)v + \gamma_0^u(\theta),$$
$$g^w(\theta, v) = \gamma_1^w(\theta)v + \gamma_0^w(\theta).$$

Furthermore, for taste shocks with sufficiently small amplitude we can guarantee strictly positive consumption $g^w(\theta, v)$ on a unique bounded ergodic set for welfare v. The nonnegativity constraints are then satisfied. It follows that this allocation also solves the problem that imposes these additional constraints.

Example 2. To illustrate the numerical value of our recursive formulation, we now compute the solution for the logarithmic utility case $U(c) = \log(c)$ with $\beta = 0.9$, $e = \eta^{-1} = 0.6$, $\theta_h = 1.2$, $\theta_l = 0.75$, $p = 0.5$ and several values of $\hat{\beta}$. Figure 2 displays steady-state distributions of welfare, measured in consumption-equivalent units $c(v(1-\beta))$. The distributions have a smooth bell-curve shape. This must be due to the smooth, mean-reverting dynamics of the model, since it cannot be a direct consequence of our two-point distribution of taste shocks. Dispersion appears to increase for lower values of $\hat{\beta}$, supporting the natural conjecture that as we approach $\hat{\beta} \to \beta$, the Atkeson-Lucas case, the invariant distributions diverge.

We now briefly discuss to issues of uniqueness and stability of the invariant distribution guaranteed by Proposition 3. This question is of economic interest because it represents an even stronger notion of social mobility than that implied by the mean-reversion condition (17) discussed in the previous subsection. Suppose the economy finds itself at a steady state ψ^*. Then convergence from any initial v_0 toward the distribution ψ^* means that the distribution of welfare for distant descendants is independent of an individual's present condition. The past exerts some influence on the present, but its influence dies out over time. The inheritability of welfare is imperfect and the advantages or disadvantages of distant ancestors are eventually wiped out.

Indeed, the solution may display this strong notion of social mobility. To see this, suppose the ergodic set for $\{k'(v_t)\}$ is compact, which is guaranteed if $\bar{\gamma} < 1$. Then, if the policy function $g^w(\theta, v)$ is monotone in v, the invariant distribution ψ^* is unique and stable: starting from any initial distribution ψ_0, the sequence of distributions $\{\psi_t\}$, generated by g^w, converges weakly to ψ^*.[12] The required monotonicity of the policy functions was satisfied by Examples 1 and 2 and seems plausible more generally.[13] Another approach suggests uniqueness and convergence without relying on monotonicity of the policy functions. Grunwald, Hyndman, Tedesco, and Tweedie (2000) show that one-dimensional, irreducible Markov processes with the Feller property that are bounded below

[12] This follows since the conditional-expectation equation (17) ensures enough mixing to apply Hopenhayn-Prescott's Theorem. See pg. 382-383 in Stokey and Lucas (1989).

[13] Indeed, it can be shown that $g^w(\underline{\theta}, v)$ must be strictly increasing in v. However, although we know of no counterexample, we have not found conditions that ensure the monotonicity of $g^w(\theta, v)$ in v for $\theta \neq \underline{\theta}$.

and are conditional linear autoregressive, as implied by (17), have a unique and stable invariant distribution. All their hypotheses have been verified here except for the technical condition of irreducibility.[14] We do not pursue this formally other than to note that the forces for reversion in (17) might be further exploited to establish uniqueness and convergence.

We have focused on steady states where the distribution of welfare replicates itself over time. However, for the logarithmic utility case we can also characterize transitional dynamics.

Proposition 4 *If utility is logarithmic* $U(c) = \log(c)$, *then for any initial distribution of entitlements* ψ *there exists an endowment level* $e = \hat{e}(\psi)$ *such that the solution to the social planning problem is generated by the policy functions* (g^u, g^w). *The function* $\hat{e}$ *is increasing, in that if* ψ^b *first-order stochastic dominates* ψ^a *then* $\hat{e}(\psi^a) < \hat{e}(\psi^b)$.

One can apply this result to the case with no initial inequality, where dynasties are all started at v^* solving $k'(v^*) = 0$. The cross-sectional distribution of welfare and consumption fans out over time starting from this initial egalitarian position. The issues of convergence and uniqueness discussed above now acquire an additional economic interpretation. It implies that the transition is stable, with the cross-sectional distributions of welfare and consumption converging over time to the steady state.

---

[14] We conjecture that this condition could be guaranteed in an extension with a continuous distribution of taste shocks.

As mentioned in Section 3, for any utility function one can characterize the solution for any (ψ, e) as the solution to a relaxed problem with some sequence $\{Q_t\}$, that are not necessarily exponential. That is, imposing the general intertemporal constraint (9) instead of (10). Proposition 4 identifies the distributions and endowment pairs (ψ, e) that lead to exponential $\{Q_t\}$ in the logarithmic case. More generally, with logarithmic utility for *any* pair (ψ, e), we can show that $Q_t = \hat{\beta}^t + a\beta^t$ for some constant a. The entire optimal allocation can then be characterized by the policy functions from a non-stationary Bellman equation. Since $\{Q_t\}$ is asymptotically exponential (i.e. $\lim_{t\to\infty} \hat{\beta}^{-t} Q_t = 1$) the long-run dynamics are dominated by the policy functions (g^u, g^w) from the problem with $Q_t = \hat{\beta}^t$ that we have characterized.

We have shown the existence of a steady state ψ^* generated by the policy function g^w. We now provide sufficient conditions to ensure that ψ^* is also a steady state for the social planning problem. This involves two things. We first establish that allocations generated by the policy functions are indeed incentive compatible by verifying the limiting condition in Theorem 2; this guarantees that, given ψ^* and η, the allocation maximizes the Lagrangian (11). Second, we verify that average consumption is finite under ψ^*, so that there exists some endowment e for which the resource constraints (4) and (9) hold. It follows that the allocations generated by (g^u, g^w) solve the social planning problem, given e and ψ^*.

Proposition 5 *The allocation generated from the policy functions (g^u, g^w), starting from any v_0, is guaranteed to be incentive compatible in the following cases: (a) utility is bounded above; (b) utility* *is bounded below; (c) utility is logarithmic; or (d)* $\bar{\gamma} < 1$ *or* $\underline{\gamma} > 0$.

Next, we give sufficient conditions to guarantee that total consumption is finite at the steady state ψ^*. If the ergodic set for welfare v is bounded away from the extremes, then consumption is bounded and total consumption is finite. Even when a bounded ergodic set for welfare v cannot be ensured, we can guarantee that total consumption is finite for a large class of utility functions.

Proposition 6 *Total consumption is finite under the invariant distribution* ψ^*,

$$\int \sum_{\theta \in \Theta} c\big(g^u(\theta, v)\big) p(\theta)\, d\psi^*(v) < \infty,$$

if either (a) the ergodic set for v *is bounded; or (b) utility is such that* $c'(U(c))$ *is a convex function of* c.

Condition (a) is guaranteed if the amplitude of taste shocks is not too wide, so that $\bar{\gamma} < 1$. Condition (b) holds, for example, with constant relative risk aversion utility functions with $\sigma \geq 1$.

The value of average consumption depends on the value of η. For instance, in the case of constant relative risk aversion utility, average steady state consumption is a power function of η, and thus has full range. In fact, in this case the entire solution for consumption is homogenous of degree one in the value of the endowment e. This ensures a steady state solution to the social planning problem for any endowment level.

We now return to the Pareto problem and its relation to the social planning problem. Recall that the former is exactly as the latter except that the promise-keeping constraints are inequalities, instead of equalities. The next result establishes that these inequality constraints bind for steady state distributions ψ^*. Thus, the solutions to the Pareto and social planning problems coincide. The proof relies on the fact that a marginal increase in v contributes $k'(v)-1$ to the welfare criterion (7) and that $k'(v) < 1$.

Proposition 7 *Let ψ^* denote an invariant distribution of the process $\{v_t\}$ generated by the policy function g^w. Then $\psi = \psi^*$ solves the Pareto problem*

$$\max_{\psi} \left\{ S(\psi; e) - \int v \, d\psi(v) \right\} \quad \textit{subject to} \quad \psi(v) \leq \psi^*(v) \quad \textit{for all } v,$$

if consumption is strictly positive on the support of ψ^*, i.e. $c\big(g^u(\theta, v)\big) > 0$ for all v in the support of ψ^*.

The condition that consumption be strictly positive is automatically satisfied if utility is unbounded below; otherwise it is ensured if the amplitude of the taste shocks is not too large, so that $\bar{\gamma} > 0$.

It follows that if the Pareto problem is started with the distribution $\tilde{\psi} = \psi^*$, then this distribution is replicated over time: $\psi_t = \psi^*$ for all $t = 0, 1, \ldots$ It also implies that the steady state distribution

ψ^* is time-consistent, in the sense that in any period t the initial allocation also solves the Pareto problem, and is, thus, ex-post Pareto efficient.

In our analysis, future generations were valued by attaching a declining Pareto weight on their expected welfare. The mapping from Pareto weights to their expected welfare is as follows. At a steady state, the weights $\alpha_t = \hat{\beta}^t$ for $t \geq 1$ correspond to the constraints that average expected welfare be no lower than that obtained under the steady state distribution ψ^*:

$$\int v_t \, d\psi_t(v_t) \geq \underline{V} \quad \text{for all } t = 1, \ldots \tag{19}$$

where $\underline{V} \equiv \int v \, d\psi^*(v)$. In other words, we have characterized the maximization of the expected welfare of generation $t = 1$ subject to (19) for all other future generations $t = 2, 3, \ldots$, as well as the promise-keeping constraints of delivering $\tilde{v}$ or more to the founder of each dynasty, given some distribution of initial promises $\tilde{\psi}$, and satisfying incentive compatibility. We have constructed steady states for such a problem: initial entitlement distributions $\tilde{\psi} = \psi^*$ that replicate themselves over time. Different values of $\hat{\beta}$ translate into different steady states ψ^* and trace out associated values for $\underline{V}$ in this Pareto problem.

6. Conclusions

How should privately-felt parental altruism affect the social contract? What are the long-run implications for inequality? To address these questions, we modeled the tradeoff between equality of opportunity for newborns and incentives for altruistic parents. In our model, society should exploit altruism to motivate parents, linking the welfare of children to that of their parents. If future generations are included in the welfare function this inheritability should be tempered and the existence of a steady-state is ensured, where welfare and consumption are mean-reverting, long-run inequality is bounded, social mobility is possible and misery is avoided by everyone.

The backbone of our model requires a tradeoff between insurance and incentives. The source for this tradeoff is inessential. In this paper, we adopted the Atkeson-Lucas taste-shock specification for purposes of comparison. Farhi and Werning (2006) and Farhi, Kocherlakota, and Werning (2006) study dynamic Mirrleesian models—with productivity shocks, instead of taste shocks—and find that a progressive estate tax implements efficient allocations by providing the necessary mean-reversion across generations.

Appendix

Proof of Theorem 1

Weak concavity of the value function $k(v)$ follows because the component planning sequence problem has a concave objective and a convex constraint set. The weak concavity of the value function $k(v)$ implies its continuity over the interior of its domain. If utility is bounded, continuity at the extremes can also be established as follows. Define the first-best value function

$$k^*(v) \equiv \max_{\{u_t\}} \sum_{t=0}^{\infty} \hat{\beta}^t \, \mathbb{E}_{-1}\left[\theta_t u_t(\theta^t) - \eta c\big(u_t(\theta^t)\big)\right]$$

subject to $v = \sum_{t=0}^{\infty} \beta^t \, \mathbb{E}_{-1}[\theta_t u_t(\theta^t)]$. Then $k^*(v)$ is continuous and $k(v) \leq k^*(v)$, with equality at any finite extremes $\bar{v}$ and $\underline{v}$. Then continuity of $k(v)$ at finite extremes follows. Thus, $k(v)$ is continuous.

We first show that the constraint (10) with $q = \hat{\beta}$ implies that utility and continuation welfare are well-defined. Toward a contradiction, suppose

$$\lim_{T\to\infty} \sum_{t=0}^{T} \beta^t E_s \theta_t u(c_t)$$

is not defined, for some $s \geq -1$. This implies that $\lim_{T\to\infty} \sum_{t=0}^{T} \beta^t \max\{E_s\theta_t u(c_t), 0\} = \infty$. Since utility is concave $\theta U(c) \leq Ac + B$ for some $A, B > 0$, so it follows that

$$\sum_{t=0}^{T} \beta^t \max\{E_s\theta_t u(c_t), 0\} \leq A \sum_{t=0}^{T} \beta^t \, \mathbb{E}_s[c_t] + B \leq A \sum_{t=0}^{T} \hat{\beta}^t \, \mathbb{E}_s[c_t] + B$$

Taking the limit yields $\lim_{T\to\infty} \sum_{t=0}^{T} \hat{\beta}^t \, \mathbb{E}_s[c_t] = \infty$. Since there are finitely many histories $\theta^s \in \Theta^{s+1}$ this implies $\lim_{T\to\infty} \sum_{t=0}^{T} \hat{\beta}^t \, \mathbb{E}_{-1}[c_t] = \infty$. If there is a non-zero measure of such agents this implies a contradiction of the intertemporal constraint (10), and thus of at least one resource constraint in (4). Thus, for both the relaxed and original planning problems utility and continuation welfare are well defined given the other constraints on the problem. This is important for our recursive formulation below.

We next prove two lemmas that imply the rest of the theorem. Consider the optimization problem on the right hand side of the Bellman equation:

$$\sup_{u,w} \mathbb{E}\big[\theta u(\theta) - \eta c\big(u(\theta)\big) + \hat{\beta} k\big(w(\theta)\big)\big] \tag{20}$$

$$v = \mathbb{E}\big[\theta u(\theta) + \beta w(\theta)\big] \tag{21}$$

$$\theta u(\theta) + \beta w(\theta) \geq \theta u(\theta') + \beta w(\theta') \text{ for all } \theta, \theta' \in \Theta \tag{22}$$

Define $m \equiv \max_{c\geq 0,\theta\in\Theta}(\theta U(c) - \eta c)$ and $\hat{k}(v) \equiv k(v) - m/(1-\hat{\beta}) \leq 0$. The problem in (20) is equivalent to the following optimization with non-positive objective:

$$\sup_{u,w} \mathbb{E}\big[\theta u(\theta) - \eta c\big(u(\theta)\big) - m + \hat{\beta}\hat{k}\big(w(\theta)\big)\big] \tag{23}$$

subject to (21) and (22).

Lemma A.1 *The supremum in (20), or equivalently (23), is attained.*

Proof. If utility is bounded the result follows immediately by continuity of the objective function and compactness of the constraint set. So suppose utility is unbounded above and below — similar arguments apply when utility is only unbounded below or only unbounded above. We first show that

$$\lim_{v\to\infty} \hat{k}(v) = \lim_{v\to-\infty} \hat{k}(v) = -\infty \tag{24}$$

and then use this result to restrict, without loss, the optimization within a compact set, ensuring a maximum is attained.

To establish these limits, define

$$h(v;\hat{\beta}) \equiv \sup_{u} \sum_{t=0}^{\infty} \hat{\beta}^t\, \mathbb{E}_{-1}\big[\theta_t u(\theta^t) - \eta c\big(u(\theta^t)\big) - m\big]$$

subject to $v = \mathbb{E}_{-1} \sum_{t=0}^{\infty} \beta^t \theta_t u(\theta^t)$. Since this corresponds to the same problem but without the incentive constraints it follows that $\hat{k}(v) \leq h(v, \hat{\beta})$. If $\lim_{v \to \infty} h(v, \hat{\beta}) = \lim_{v \to -\infty} h(v, \hat{\beta}) = -\infty$, then the desired limits (24) follow. Since $\theta u - \eta c(u) - m \leq 0$ and $\beta < \hat{\beta}$ it follows that

$$h(v, \hat{\beta}) \leq h(v, \beta) = v - \eta C(v, \beta) - \frac{m}{1-\beta}, \tag{25}$$

where

$$C(v, \beta) \equiv \inf_{u} \sum_{t=0}^{\infty} \beta^t \, \mathbb{E}_{-1} \left[c\big(u(\theta^t)\big) \right]$$

subject to $v = \sum_{t=0}^{\infty} \beta^t \, \mathbb{E}_{-1}[\theta_t u(\theta^t)]$. Note that $C(v, \beta)$ is a standard convex first-best allocation problem, with solution $u(\theta^t) = (c')^{-1}(\theta_t \gamma(v))$ for some positive multiplier $\gamma(v)$, increasing in v and such that $\lim_{v \to -\infty} \gamma(v) = 0$ and $\lim_{v \to \infty} \gamma(v) = \infty$. Then

$$C(v, \beta) = \frac{1}{1-\beta} \mathbb{E}\left[c\Big((c')^{-1} \big(\theta \gamma(v)\big) \Big) \right],$$

so that $\lim_{v \to -\infty} h(v, \beta) = -\infty$ and $\lim_{v \to \infty} h(v, \beta) = -\infty$. Using the inequality (25) this establishes $\lim_{v \to -\infty} h(v, \hat{\beta}) = -\infty$ and $\lim_{v \to \infty} h(v, \hat{\beta}) = -\infty$, which, in turn, imply the limits (24).

Fix a v. Take any allocation that verifies the constraints (21) and (22) and let $\bar{k} < \infty$ be the corresponding value of (23). Then, since the objective is non-positive, we can restrict the maximization to $w(\theta)$ such that $\hat{k}(w(\theta)) \geq \bar{k}/(\hat{\beta}p(\theta))$. Since $\hat{k}(w(\theta))$ is concave with the limits (24), this defines a closed, bounded interval for $w(\theta)$, for each θ. It follows that there exists $M_{v,w} < \infty$ such that we can restrict the maximization to $|w(\theta)| \leq M_{v,w}$.

Similarly, we can restrict the maximization over $u(\theta)$ so that $\theta u(\theta) - \eta c(u(\theta)) - m \geq \bar{k}/p(\theta)$. Since $(\theta u - \eta c(u))$ is strictly concave, with $(\theta u - \eta c(u)) \to -\infty$ when either $u \to \infty$ or $u \to -\infty$, this defines a closed, bounded interval for $u(\theta)$, for each θ. Thus, there exists an $M_{v,u} < \infty$ such that we can restrict the maximization to $|u(\theta)| \leq M_{v,u}$.

Hence, we can restrict the maximization in (23) to a compact set. Since the objective function is continuous over this restricted set, the maximum must be attained. ■

Lemma A.2 *The value function $k(v)$ satisfies the Bellman equation (13)–(15).*

Proof. Suppose that for some v

$$k(v) > \max_{u,w} \mathbb{E}\big[\theta u(\theta) - \eta c\big(u(\theta)\big) + \hat{\beta}k\big(w(\theta)\big)\big]$$

where the maximization is subject to (21) and (22). Then there exists $\Delta > 0$ such that

$$k(v) \geq \mathbb{E}\big[\theta u(\theta) - \eta c\big(u(\theta)\big) + \hat{\beta}k\big(w(\theta)\big)\big] + \Delta$$

for all (u, w) that satisfy (21) and (22). But then by definition

$$k\big(w(\theta)\big) \geq \sum_{t=0}^{\infty} \hat{\beta}^t \, \mathbb{E}_{-1}\big[\theta_t \tilde{u}_t(\theta^t) - \eta c\big(\tilde{u}_t(\theta^t)\big)\big]$$

for *all* allocations $\tilde{u}$ that yield $w(\theta)$ and are incentive compatible. Substituting, we find that

$$k(v) \geq \sum_{t=0}^{\infty} \hat{\beta}^t \, \mathbb{E}_{-1}\big[\theta_t u_t(\theta^t) - \eta c\big(u_t(\theta^t)\big)\big] + \Delta$$

for *all* incentive-compatible allocations that deliver v, a contradiction with the definition of $k(v)$. Namely, that there should be a plan with value arbitrarily close to $k(v_0)$. We conclude that $k(v) \leq \max_{u,w} \mathbb{E}[\theta u(\theta) - \eta c(u(\theta)) + \hat{\beta} k(w(\theta))]$ subject to (21) and (22).

By definition, for every v and $\varepsilon > 0$ there exists a plan $\{\tilde{u}_t(\theta^t; v, \varepsilon)\}$ that is incentive compatible and delivers v with value

$$\sum_{t=0}^{\infty} \hat{\beta}^t \, \mathbb{E}_{-1}\big[\theta_t \tilde{u}_t(\theta^t; v, \varepsilon) - \eta c\big(\tilde{u}_t(\theta^t; v, \varepsilon)\big)\big] \geq k(v) - \varepsilon.$$

Let $(u^*(\theta), w^*(\theta)) \in \arg\max_{u,w} \mathbb{E}[\theta u(\theta) - \eta c(u(\theta)) + \hat{\beta} k(w(\theta))]$. Consider the plan $u_0(\theta_0) = u^*(\theta_0)$

and $u_t(\theta^t) = \tilde{u}_{t-1}((\theta_1, \ldots, \theta_t); w^*(\theta_0), \varepsilon)$ for $t \geq 1$. Then

$$
\begin{aligned}
k(v) &\geq \sum_{t=0}^{\infty} \hat{\beta}^t \, \mathbb{E}_{-1}\left[\theta_t u_t(\theta^t) - \eta c\big(u_t(\theta^t)\big)\right] \\
&= \mathbb{E}_{-1}\Big[\theta_0 u^*(\theta_0) - \eta c\big(u^*(\theta_0)\big) + \hat{\beta} \sum_{t=0}^{\infty} \hat{\beta}^t \, \mathbb{E}_0\left[\theta_{t+1} u_{t+1}(\theta^{t+1}) - \eta c\big(u_{t+1}(\theta^{t+1})\big)\right]\Big] \\
&\geq \max_{u,w} \mathbb{E}\left[\theta u(\theta) - \eta c\big(u(\theta)\big) + \hat{\beta} k\big(w(\theta)\big)\right] - \hat{\beta}\varepsilon.
\end{aligned}
$$

Since $\varepsilon > 0$ was arbitrary it follows that $k(v) \geq \max_{u,w} \mathbb{E}[\theta u(\theta) - \eta c(u(\theta)) + \hat{\beta} k(w(\theta))]$ subject to (21) and (22).

Finally, together both inequalities imply $k(v) = \max_{u,w} \mathbb{E}[\theta u(\theta) - \eta c(u(\theta)) + \hat{\beta} k(w(\theta))]$ subject to (21) and (22). ■

Proof of Theorem 2

We establish the following results from which the Theorem follows: **(a)** An allocation $\{u_t\}$ is optimal for the component planning problem with multiplier $\hat{\lambda}$, given $v_0 = v$, if and only if it is generated by the policy functions (g^u, g^w) starting at v_0, is incentive compatible, and delivers welfare v_0; **(b)** an allocation $\{u_t\}$ generated by the policy functions (g^u, g^w), starting at v_0, has

$\lim_{t\to\infty} \beta^t \, \mathbb{E}_{-1}[v_t(\theta^{t-1})] = 0$ and delivers welfare v_0; **(c)** an allocation $\{u_t\}$ generated by the policy functions (g^u, g^w), starting from v_0, is incentive compatible if

$$\limsup_{t\to\infty} \mathbb{E}_{-1}\beta^t v_t\big(\sigma^{t-1}(\theta^{t-1})\big) \geq 0$$

for all reporting strategies σ.

Part (a). Suppose the allocation $\{u_t\}$ is generated by the policy functions starting from v_0, is incentive compatible and delivers welfare v_0. After repeated substitutions of the Bellman equation (13), we arrive at

$$k(v_0) = \sum_{t=0}^{T} \hat{\beta}^t \, \mathbb{E}_{-1}[\theta_t u_t(\theta^t) - \eta c(u_t(\theta^t))] + \hat{\beta}^{T+1} \, \mathbb{E}_{-1} k(v_{T+1}(\theta^T)). \tag{26}$$

Since $k(v_0)$ is bounded above this implies that

$$k(v_0) \leq \sum_{t=0}^{\infty} \hat{\beta}^t \, \mathbb{E}_{-1}[\theta_t u_t(\theta^t) - \eta c(u_t(\theta^t))],$$

so $\{u_t\}$ is optimal, by definition of $k(v_0)$.

Conversely, suppose an allocation $\{u_t\}$ is optimal given v_0. Then, by definition it must be incentive compatible and deliver welfare v_0. Define the continuation welfare implicit in the allocation

$$w_0(\theta_0) \equiv \sum_{t=1}^{\infty} \beta^{t-1}\,\mathbb{E}_0[\theta_t u_t(\theta^t)].$$

and suppose that either $u_0(\theta) \neq g^u(\theta; v_0)$ or $w_0(\theta) \neq g^w(\theta; v_0)$, for some $\theta \in \Theta$. Since the original plan $\{u_t\}$ is incentive compatible, $u_0(\theta)$ and $w_0(\theta)$ satisfy (21) and (22). The Bellman equation then implies that

$$\begin{aligned} k(v_0) &= \mathbb{E}\big[g^u(\theta; v_0) - \eta c\big(g^u(\theta; v_0)\big) + \beta k\big(g^w(\theta; v_0)\big)\big] \\ &> \mathbb{E}\big[u_0(\theta) - \eta c\big(u_0(\theta)\big) + \beta k\big(w_0(\theta)\big)\big] \\ &\geq \mathbb{E}_{-1}\big[u_0(\theta_0) - \eta c\big(u_0(\theta_0)\big)\big] + \sum_{t=1}^{\infty} \beta^t\,\mathbb{E}_{-1}\big[u_t(\theta^t) - \eta c\big(u_t(\theta^t)\big)\big]. \end{aligned}$$

The first inequality follows since u_0 does not maximize (13), while the second inequality follows the definition of $k(w_0(\theta))$. Thus, the allocation $\{u_t\}$ cannot be optimal, a contradiction. A similar argument applies if the plan is not generated by the policy functions after some history θ^t and $t \geq 1$. We conclude that an optimal allocation must be generated from the policy functions.

Part (b). First, suppose an allocation $\{u_t, v_t\}$ generated by the policy functions (g^u, g^w) starting at v_0 satisfies $\lim_{t\to\infty} \beta^t \mathbb{E}_{-1} v_t(\theta^{t-1}) = 0$. Then, after repeated substitutions of (14), we obtain

$$v_0 = \sum_{t=0}^{T} \beta^t \mathbb{E}_{-1}\left[\theta_t u_t(\theta^t)\right] + \beta^{T+1} \mathbb{E}_{-1}\left[v_{T+1}(\theta^T)\right]. \tag{27}$$

Taking the limit we get $v_0 = \sum_{t=0}^{\infty} \beta^t \mathbb{E}_{-1}[\theta_t u_t(\theta^t)]$ so that the allocation $\{u_t\}$ delivers welfare v_0. Next, we show that for any allocation generated by (g^u, g^w), starting from finite v_0, we have $\lim_{t\to\infty} \beta^t \mathbb{E}_{-1} v_t(\theta^{t-1}) = 0$.

Suppose utility is unbounded above and $\limsup_{t\to\infty} \beta^t \mathbb{E}_{-1} v_t(\theta^{t-1}) > 0$. Then $\hat\beta > \beta$ implies that $\limsup_{t\to\infty} \hat\beta^t \mathbb{E}_{-1} v_t(\theta^{t-1}) = \infty$. Since the value function $k(v)$ is non-constant, concave and reaches an interior maximum, we can bound the value function so that $k(v) \le av + b$, with $a < 0$. Thus,

$$\liminf_{t\to\infty} \hat\beta^t \mathbb{E}_{-1} k\big(v_t(\theta^{t-1})\big) \le a \limsup_{t\to\infty} \hat\beta^t \mathbb{E}_{-1} v_t(\theta^{t-1}) + b = -\infty$$

and then (26) implies that $k(v_0) = -\infty$, a contradiction since there are feasible plans that yield finite values. We conclude that $\limsup_{t\to\infty} \beta^t \mathbb{E}_{-1} v_t(\theta^{t-1}) \le 0$.

Similarly, suppose utility is unbounded below and that $\liminf_{t\to\infty} \beta^t \mathbb{E}_{-1} v_t(\theta^{t-1}) < 0$. Then $\liminf_{t\to\infty} \hat\beta^t \mathbb{E}_{-1} v_t(\theta^{t-1}) = -\infty$. Using $k(v) \le av + b$, with $a > 0$, we conclude that

$$\liminf_{t\to\infty} \hat\beta^t \mathbb{E}_{-1} k\big(v_t(\theta^{t-1})\big) = -\infty$$

implying $k(v_0) = -\infty$, a contradiction. Thus, we must have $\liminf_{t\to\infty} \beta^t \, \mathbb{E}_{-1} v_t(\theta^{t-1}) \geq 0$.

The two established inequalities imply $\lim_{t\to\infty} \beta^t \, \mathbb{E}_{-1} v_t(\theta^{t-1}) = 0$.

Part (c). Suppose $\limsup_{t\to\infty} \beta^t \, \mathbb{E}_{-1} v_t(\sigma^{t-1}(\theta^{t-1})) \geq 0$ for every reporting strategy σ. Then after repeated substitutions of (15),

$$v_0 \geq \sum_{t=0}^{T} \beta^t \, \mathbb{E}_{-1} \left[\theta_t u_t\big(\sigma^t(\theta^t)\big) \right] + \beta^{T+1} \, \mathbb{E}_{-1} \left[v_{T+1}\big(\sigma^T(\theta^T)\big) \right].$$

implying

$$v_0 \geq \liminf_{T\to\infty} \sum_{t=0}^{T} \beta^t \, \mathbb{E}_{-1} \left[\theta_t u_t\big(\sigma^t(\theta^t)\big) \right].$$

Therefore, $\{u_t\}$ is incentive compatible, since v_0 is attainable with truth telling from part (b).

Proof of Proposition 1

Part (a) (Strict Concavity) Let $\{u_t(\theta^t, v_0), v_t(\theta^{t-1}, v_0)\}$ be generated from the policy functions starting at v_0 (note: no claim of incentive compatibility is required). Take two initial welfare values

v_a and v_b, with $v_a \neq v_b$. Define the average utilities

$$u_t^\alpha(\theta^t) \equiv \alpha u_t(\theta^t, v_a) + (1-\alpha) u_t(\theta^t, v_b)$$
$$v_t^\alpha(\theta^t) \equiv \alpha v_t(\theta^t; v_a) + (1-\alpha) v_t(\theta^t; v_b)$$

Theorem 2 part (b) implies that $\{u_t(\theta^t, v_a)\}$ and $\{u_t(\theta^t, v_b)\}$ deliver v_a and v_b, respectively. This immediately implies that $\{u_t^\alpha(\theta^t)\}$ delivers initial welfare $v^\alpha \equiv \alpha v_a + (1-\alpha) v_b$. It also implies that there exists a finite time T such that

$$\sum_{t=0}^{T} \beta^t \, \mathbb{E}_{-1}\left[\theta_t u_t(\theta^t; v_a)\right] \neq \sum_{t=0}^{T} \beta^t \, \mathbb{E}_{-1}\left[\theta_t u_t(\theta^t; v_b)\right],$$

so that

$$u_t\left(\theta^t; v_a\right) \neq u_t\left(\theta^t; v_b\right), \tag{28}$$

for some history $\theta^t \in \Theta^{t+1}$. Consider iterating T times on the Bellman equations starting from v_a and v_b:

$$k(v_a) = \sum_{t=0}^{T} \hat{\beta}^t \, \mathbb{E}_{-1} \left[\theta_t u_t(\theta^t; v_a) - \eta c\big(u_t(\theta^t; v_a)\big) \right] + \hat{\beta}^{T+1} \, \mathbb{E}_{-1} k\big(v_{T+1}(\theta^T; v_a)\big)$$

$$k(v_b) = \sum_{t=0}^{T} \hat{\beta}^t \, \mathbb{E}_{-1} \left[\theta_t u_t(\theta^t; v_b) - \eta c\big(u_t(\theta^t; v_b)\big) \right] + \hat{\beta}^{T+1} \, \mathbb{E}_{-1} k\big(v_{T+1}(\theta^T; v_b)\big),$$

and averaging we obtain

$$\begin{aligned} \alpha k(v_a) + (1-\alpha) k(v_b) &= \sum_{t=0}^{T} \hat{\beta}^t \, \mathbb{E}_{-1} \left[\theta_t u_t^\alpha(\theta^t) - \eta \left[\alpha c\big(u_t(\theta^t; v_a)\big) + (1-\alpha) c\big(u_t(\theta^t; v_b)\big) \right] \right] \\ &\quad + \hat{\beta}^{T+1} \, \mathbb{E}_{-1} \left[\alpha k(v_{T+1}(\theta^T; v_a)) + (1-\alpha) k\big(v_{T+1}(\theta^T; v_b)\big) \right] \\ &\quad < \sum_{t=0}^{T} \hat{\beta}^t \, \mathbb{E}_{-1} \left[\theta_t u_t^\alpha(\theta^t) - \eta c\big(u_t^\alpha(\theta^t)\big) \right] + \hat{\beta}^{T+1} \, \mathbb{E}_{-1} k\big(v_{T+1}^\alpha(\theta^T)\big) \le k(v^\alpha), \end{aligned}$$

where the strict inequality follows from the strict concavity of the cost function $c(u)$, the fact that we have the inequality (28), and the weak concavity of the value function k. The last weak inequality follows from iterating on the Bellman equation for v^α since the average plan (u^α, v^α)

satisfies the constraints of the Bellman equation at every step. This proves that the value function $k(v)$ is strictly concave.

(b) (Differentiability) Since the value function $k(v)$ is concave, it is sub-differentiable: there is at least one subgradient at every point v. Differentiability can then be established by the following variational envelope arguments.

Suppose first that utility is unbounded below. Fix an interior value v_0 for initial welfare. For a neighborhood around v_0 define the test function

$$W(v) \equiv \mathbb{E}\big[\theta\big(g^u(\theta, v_0) + (v - v_0)\big) - \eta c\big(g^u(\theta, v_0) + (v - v_0)\big) + \hat{\beta} k\big(g^w(\theta, v_0)\big)\big].$$

Since $W(v)$ is the value of a feasible allocation in the neighborhood of v_0 it follows that $W(v) \leq k(v)$, with equality at v_0. Since $W'(v_0)$ exists it follows, by application of the Benveniste-Scheinkman Theorem (see Theorem 4.10, in Stokey and Lucas, 1989), that $k'(v_0)$ also exists and

$$k'(v_0) = W'(v_0) = 1 - \eta\, \mathbb{E}\big[c'\big(u^*(\theta)\big)\big]. \tag{29}$$

Finally, since $c'(u) \geq 0$ this shows that $k'(v) \leq 1$. The limit $\lim_{v \to -\infty} k'(v) = 1$ is inherited by the upper bound $k(v) \leq h(v, \beta) + m/(1 - \hat{\beta})$ introduced in the proof of Theorem 1, since $\lim_{v \to -\infty} \frac{\partial}{\partial v} h(v, \beta) = 1$.

The limit $\lim_{v \to \bar{v}} k'(v) = -\infty$ follows immediately from $\lim_{v \to \bar{v}} k(v) = -\infty$, if $\bar{v} < \infty$. Otherwise it is inherited by the upper bound $k(v) \leq h(v, \beta) + m/(1-\hat{\beta})$ introduced in the proof of Theorem 1, since $\lim_{v\to\infty} \frac{\partial}{\partial v} h(v, \beta) = -\infty$.

Next, suppose utility is bounded below but unbounded above, and without loss in generality suppose that the utility of zero consumption is zero. Then the argument above establishes differentiability at a point v_0 as long as $g^u(\theta, v_0) > 0$, for all $\theta \in \Theta$. However, corner solutions with $g^u(\theta, v_0) = 0$ are possible here even with Inada assumption on the utility function, so a different envelope argument is required. We provide one that exploits the homogeneity of the constraint set.

If utility is bounded below, then $\limsup_{t\to\infty} \mathbb{E}_{-1}\beta^t v_t(\sigma^{t-1}(\theta^{t-1})) \geq 0$ for all reporting strategies σ so that, applying Theorem 2, a solution $\{u_t\}$ to the planner's sequence problem is ensured. Then, for any interior v_0, the plan $\{(v/v_0)u_t\}$ is incentive compatible and attains value v for the agent. In addition the test function

$$W(v) \equiv \sum_{t=0}^{\infty} \hat{\beta}^t \, \mathbb{E}_{-1}\left[\theta_t \frac{v}{v_0} u_t(\theta^t) - \eta c\left(\frac{v}{v_0} u_t(\theta^t)\right)\right]$$

satisfies $W(v) \leq k(v)$, $W(v_0) = k(v_0)$ and is differentiable. It follows from the Benveniste-Scheinkman Theorem, that $k'(v_0)$ exists and equals $W'(v_0)$.

The proof of $\lim_{v\to\bar{v}} k'(v) = -\infty$ is the same as in the case with utility unbounded below. Finally, we show that $\lim_{v \to \underline{v}} k'(v) = \infty$. Consider the deterministic planning problem

$$\underline{k}(v) \equiv \max_u \sum_{t=0}^{\infty} \hat{\beta}^t \left(u_t - \eta c(u_t)\right)$$

subject to $v = \sum_{t=0}^{\infty} \beta^t u_t$. Note that $\underline{k}(v)$ is differentiable with $\lim_{v \to \underline{v}} \underline{k}'(v) = \infty$. Since deterministic plans are trivially incentive compatible, it follows that $\underline{k}(v) \leq k(v)$, with equality at $\underline{v}$. Then we must have $\lim_{v \to \underline{v}} k'(v) = \infty$ to avoid a contradiction.

If utility is bounded above and unbounded below then a symmetric argument, normalizing utility of infinite consumption to zero, also works. If utility is bounded above and below we can generate a test function that combines both arguments, one for the $v < v_0$ and another for $v \geq v_0$.

Proof of Proposition 2

The CLAR equation was shown in the main text, so we focus here on the bounds. Consider the program

$$\max_{u,w} \sum_n \bar{p}_n \{\bar{\theta}_n u_n - \eta c(u_n) + \hat{\beta} k(w_n)\}$$

$$v = \sum_n \bar{p}_n (\bar{\theta}_n u_n + \beta w_n)$$

$$\theta_n u_n + \beta w_n \geq \theta_n u_{n+1} + \beta w_{n+1} \text{ for } n = 1, 2, \ldots, K-1,$$

This problem and its notation require some discussion. We do not incorporate the monotonicity constraint on u. But this notation allows us to consider bunching in the following way. If any set of neighboring agents is bunched, then we group these agents under a single index and let $\bar{p}_n$ be the total probability of this group. Likewise let $\bar{\theta}_n$ represent the conditional average of θ within this group, which is what is relevant for the promise-keeping constraint and the objective function. Let θ_n be the taste shock of the highest agent in the group. The incentive constraint must rule the highest agent in each group from deviating and taking the allocation of the group above him.

Of course, every combination of bunched agents leads to a different program. The optimal allocation of our problem must solve one of these programs with a strictly monotone allocation—since bunching can be characterized by regrouping agents. Thus, below we characterize solutions to these programs with strict monotonicity of the solution.

The first-order conditions are

$$\bar{p}_n\{\bar{\theta}_n - \eta c'(u_n) - \lambda\bar{\theta}_n\} + \theta_n\mu_n - \theta_{n-1}\mu_{n-1} \leq 0$$

$$\bar{p}_n\{\hat{\beta}k'(w_n) - \beta\lambda\} + \beta(\mu_n - \mu_{n-1}) = 0$$

where, by the envelope condition $\lambda = k'(v)$.

Consider first case with utility unbounded below, so that the first order condition for consumption holds with equality. Summing the first-order conditions for consumption, we get

$$\eta\,\mathbb{E}\big[c'\big(u(\theta)\big)\big] = 1 - k'(v)$$

The first-order conditions for $n = 1$ imply

$$(1-\lambda) + \frac{\theta_1}{\bar{\theta}_1}\,\frac{\mu_1}{\bar{p}_1} = \frac{\eta c'(u_1)}{\bar{\theta}_1} \leq \frac{\eta\,\mathbb{E}\big[c'(u_\theta)\big]}{\bar{\theta}_1} = \frac{1-\lambda}{\bar{\theta}_1}.$$

This implies

$$\frac{\mu_1}{\bar{p}_1} \leq \frac{1-\lambda}{\theta_1} - (1-\lambda)\frac{\bar{\theta}_1}{\theta_1}.$$

Using

$$k'(w_1) = \frac{\beta}{\hat{\beta}}\lambda - \frac{\beta}{\hat{\beta}}\frac{\mu_1}{\bar{p}_1},$$

we get

$$k'(w_1) \geq \frac{\beta}{\hat{\beta}}\left[\lambda - \frac{1-\lambda}{\theta_1} + (1-\lambda)\frac{\bar{\theta}_1}{\theta_1}\right] = \frac{\beta}{\hat{\beta}}\left[1 + \frac{1}{\theta_1} - \frac{\bar{\theta}_1}{\theta_1}\right]k'(v) + \frac{\beta}{\hat{\beta}}\left[\frac{\bar{\theta}_1}{\theta_1} - \frac{1}{\theta_1}\right].$$

Similarly, writing the first-order conditions for $n = K$, we get

$$(1-\lambda) - \frac{\theta_{K-1}}{\bar{\theta}_K}\frac{\mu_{K-1}}{\bar{p}_K} = \frac{\eta c'\left(u_K\right)}{\bar{\theta}_K} \geq \frac{\eta\,\mathbb{E}[c'(u_\theta)]}{\bar{\theta}_K} = \frac{1-\lambda}{\bar{\theta}_K}.$$

This implies

$$-\frac{\mu_{K-1}}{\bar{p}_K} \geq \frac{1-\lambda}{\theta_{K-1}} - (1-\lambda)\frac{\bar{\theta}_K}{\theta_{K-1}}.$$

Using

$$k'\left(w_K\right) = \frac{\beta}{\hat{\beta}}\lambda + \frac{\beta}{\hat{\beta}}\frac{\mu_{K-1}}{\bar{p}_K},$$

we get

$$k'\left(w_K\right) \leq \frac{\beta}{\hat{\beta}}\left[\lambda - \frac{1-\lambda}{\theta_{K-1}} + (1-\lambda)\frac{\bar{\theta}_K}{\theta_{K-1}}\right] = \frac{\beta}{\hat{\beta}}\left[1 + \frac{1}{\theta_{K-1}} - \frac{\bar{\theta}_K}{\theta_{K-1}}\right]k'(v) + \frac{\beta}{\hat{\beta}}\left[\frac{\bar{\theta}_K}{\theta_{K-1}} - \frac{1}{\theta_{K-1}}\right].$$

For any n, $w_K \leq w_n \leq w_1$,

$$\frac{\beta}{\hat{\beta}}\left[1 + \frac{1}{\theta_1} - \frac{\bar{\theta}_1}{\theta_1}\right]k'(v) + \frac{\beta}{\hat{\beta}}\left[\frac{\bar{\theta}_1}{\theta_1} - \frac{1}{\theta_1}\right] \leq k'\left(w_n\right)$$
$$\leq \frac{\beta}{\hat{\beta}}\left[1 + \frac{1}{\theta_{K-1}} - \frac{\bar{\theta}_K}{\theta_{K-1}}\right]k'(v) + \frac{\beta}{\hat{\beta}}\left[\frac{\bar{\theta}_K}{\theta_{K-1}} - \frac{1}{\theta_{K-1}}\right].$$

After rearranging, we obtain

$$\frac{\beta}{\hat{\beta}}\left[1+\frac{1}{\theta_1}-\frac{\bar{\theta}_1}{\theta_1}\right](1-k'(v))+1-\frac{\beta}{\hat{\beta}}\geq 1-k'\left(g^w\left(\theta,v\right)\right)$$
$$\geq \frac{\beta}{\hat{\beta}}\left[1+\frac{1}{\theta_{K-1}}-\frac{\bar{\theta}_K}{\theta_{K-1}}\right](1-k'(v))+1-\frac{\beta}{\hat{\beta}}.$$

To arrive at the expression in the text we take the worst case scenario: we choose the subproblem that is most unfavorable to each bound, noting that $1-k'(v)\geq 0$. Namely we define

$$\bar{\gamma}\equiv(\beta/\hat{\beta})\max_{1\leq n\leq N}\{(1+\theta_n-\mathbb{E}[\theta\mid\theta\leq\theta_n])/\theta_n\}$$
$$\underline{\gamma}\equiv(\beta/\hat{\beta})\min_{2\leq n\leq N}\{1+\theta_{n-1}-\mathbb{E}[\theta\mid\theta\geq\theta_n]/\theta_{n-1}\}.$$

Turning to the bounded utility case, note that all the first-order conditions and constraints are satisfied when $\lambda\geq 1$ with $\mu_n=0$ and $u\left(\theta\right)=U(0)$ and $w\left(\theta\right)=\beta^{-1}v>v$. The first-order condition for w implies $k'(w(\theta))=k'(\beta^{-1}v)=(\hat{\beta}/\beta)k'(v)$. Since the problem is strictly convex, this represents the unique solution. Recall that in the arguments above establishing the lower bound involved no assumption on interior solutions for u, so this holds for all v. The upper bound, on the other hand, did require $u(\theta)>U(0)$ for all θ, which must be true for high enough v, i.e. for low enough $k'(v)$.

Proof of Proposition 3

Consider first the case with utility unbounded below. Since the derivative $k'(v)$ is continuous and strictly decreasing, we can define the transition function

$$Q(x,\theta) = k'\left(g^w\left((k')^{-1}(x),\theta\right)\right)$$

for all $x < 1$ if utility is unbounded below. For any probability distribution μ, let $T_Q(\mu)$ be the probability distribution defined by

$$T_Q(\mu)(A) = \int \mathbf{1}_{\{Q(x,\theta)\in A\}}\, d\mu(x)\, dp\,(\theta)$$

for any Borel set A. Define

$$T_{Q,n} \equiv \frac{T_Q + T_Q^2 + \cdots + T_Q^n}{n}$$

For example, $T_{Q,n}(\delta_x)$ is the empirical average of $\{k'(v_t)\}_{t=1}^n$ over all histories of length n starting with $k'(v_0) = x$. The following lemma establishes the existence of an invariant distribution by considering the limits of $\{T_{Q,n}\}$.

Lemma A.3 *If utility is unbounded below, then for any $x < 1$ there exists a subsequence of distributions $\{T_{Q,\phi(n)}(\delta_x)\}$ that converges weakly (i.e. in distribution) to an invariant distribution under*

Q on $(-\infty, 1)$.

Proof. The bounds (18) derived in Proposition 2 imply that for all $\theta \in \Theta$

$$\lim_{x \uparrow 1} Q\left(x, \theta\right) = \lim_{v \to -\infty} k'\big(g^w(\theta, v)\big) = \beta/\hat{\beta} < 1.$$

We first extend the continuous transition function $Q(x, \theta)\colon (-\infty, 1) \times \Theta \to (-\infty, 1)$ to a continuous transition function $\hat{Q}(x, \theta)\colon\ (-\infty, 1] \times \Theta \to (-\infty, 1)$, with $\hat{Q}(1, \theta) = \beta/\hat{\beta}$ and $\hat{Q}(x, \theta) = Q(x, \theta)$, for all $x \in (-\infty, 1)$. It follows that $T_{\hat{Q}}$ maps probability distributions over $(-\infty, 1]$ to probability distributions over $(-\infty, 1)$, and $T_Q(\delta_x) = T_{\hat{Q}}(\delta_x)$, for all $x \in (-\infty, 1)$.

We next show that the sequence $\{T_{\hat{Q},n}(\delta_x)\}$ is tight, in that for any $\varepsilon > 0$ there exists a compact set A_ε such that $T_{\hat{Q},n}(\delta_x)(A_\varepsilon) \geq 1 - \varepsilon$, for all n. The expected value of the distribution $T^n_{\hat{Q}}(\delta_x)$ is simply $\mathbb{E}_{-1}[k'(v_t(\theta^{t-1}))]$ with $x = k'(v_0) < 1$. Recall that $\mathbb{E}_{-1}[k'(v_t(\theta^{t-1}))] = (\beta/\hat{\beta})^t k'(v_0) \to 0$. This implies that

$$\begin{aligned}
\min\{0, k'(v_0)\} &\leq \mathbb{E}_{-1}\big[k'\big(v_t(\theta^{t-1})\big)\big] \\
&\leq T^n_{\hat{Q}}(\delta_x)(-\infty, -A)(-A) + \big(1 - T^n_{\hat{Q}}(\delta_x)(-\infty, -A)\big)1
\end{aligned}$$

for all $A > 0$. Rearranging,

$$T^n_{\hat{Q}}(\delta_x)(-\infty, -A) \leq \frac{1 - \min\{0, x\}}{A + 1}$$

which implies that $\{T^n_{\hat{Q}}(\delta_x)\}$, and therefore $\{T_{\hat{Q},n}(\delta_x)\}$, is tight.

Tightness implies that there exists a subsequence $T_{\hat{Q},\phi(n)}(\delta_x)$ that converges weakly, i.e. in distribution, to some distribution π. Since $\hat{Q}(x,\theta)$ is continuous in x, then $T_{\hat{Q}}(T_{\hat{Q},\phi(n)}(\delta_x))$ converges weakly to $T_{\hat{Q}}(\pi)$. But the linearity of $T_{\hat{Q}}$ implies that

$$T_{\hat{Q}}\big(T_{\hat{Q},\phi(n)}(\delta_x)\big) = \frac{T^{\phi(n)+1}_{\hat{Q}}(\delta_x) - T_{\hat{Q}}(\delta_x)}{\phi(n)} + T_{\hat{Q},\phi(n)}(\delta_x)$$

and since $\phi(n) \to \infty$ we must have $T_{\hat{Q}}(\pi) = \pi$.

Recall that $T_{\hat{Q}}$ maps probability distributions over $(-\infty, 1]$ to probability distributions over $(-\infty, 1)$. This implies that $\pi = T_{\hat{Q}}(\pi)$ has no probability mass at $\{1\}$. Since T_Q and $T_{\hat{Q}}$ coincide for such distributions, it follows that $\pi = T_Q(\pi)$, so that π is an invariant distribution under Q on $(-\infty, 1)$. ■

The argument for the case with utility bounded below is very similar. Define the transition function $Q(x,\theta)$ as above, but for all $x \in \mathbb{R}$, since now $k'(v)$ can take on any real value. If utility is unbounded above but $\bar{\gamma} < 1$, then there exists an upper bound $v_H < \bar{v}$ for the ergodic set for v. Define the welfare level $v_0 > \underline{v}$ by $k'(v_0) = 1$. Next, define v_L to be the minimum of the policy function g^w over $v \in [v_0, v_H]$, which is defined since g^w is continuous over this compact set. If utility is bounded above then let v_L by the minimum of g^w over $v \in [v_0, \bar{v})$, which is defined since

$\lim_{v \to \bar{v}} g^w(\theta, v) = \bar{v}$. In both cases, since $g^w > \underline{v}$ we must have that this minimum is above misery: $v_L > \underline{v}$. Finally, the transition function is continuous with $Q(x, \theta) \leq k'(v_L) < \infty$. The rest of the argument is then a simple modification of the one above for utility unbounded below, with $k'(v_L)$ playing the role of 1 (things are actually slightly simpler here, since no continuous extension of Q is required).

If $\underline{\gamma} > 0$ then the bound in (18) implies that $k'\left(g^w\left(\bar{\theta}, v\right)\right) \leq 1 - \beta/\hat{\beta}$ and the result follows immediately.

Proof of Proposition 4

Consider indexing the relaxed planning problem by e and setting $\eta = e^{-1}$ for the associated component planning problem, with associated value function $k(v; e)$. We first show that if an initial distribution ψ satisfies the condition $\int k'(v; e)\, d\psi(v) = 0$, then the solution to the relaxed and original planning problems coincide. We then show that for any initial distribution there exists a value for e that satisfies this condition.

Since utility is unbounded below, we have $k'(v_t; e) = \mathbb{E}_{t-1}\left[1 - \eta c'\left(u_t^v(\theta^t)\right)\right]$. Applying the law of iterated expectations to (17) then yields

$$\mathbb{E}_{-1}\left[1 - \eta c'\left(u_t^v(\theta^t)\right)\right] = \left(\frac{\beta}{\hat{\beta}}\right)^t k'\left(v; e\right).$$

With logarithmic utility $c'(u) = c(u)$, so that $\int k'(v;e)\,d\psi(v) = 0$ implies $\int \mathbb{E}_{-1}[c_t]\,d\psi = \eta^{-1} = e$ for all $t = 0, 1, \ldots$ Then, Proposition 5 implies that the allocation is incentive compatible, and applying part (c) of Theorem 2, it follows that it must solve the original planning problem.

Now consider any initial distribution ψ. We argue that we can find a value of $\eta = e^{-1}$ such that $\int k'(v;e)\,d\psi(v) = 0$. The homogeneity of the sequential problem implies that

$$k(v;e) = \frac{1}{1-\hat{\beta}}\log(e) + k\left(v - \frac{1}{1-\beta}\log(e); 1\right)$$

Note that $k'(v - \frac{1}{1-\beta}\log(e); 1)$ is strictly increasing in e and limits to 1 as $e \to \infty$, and to $-\infty$ as $e \to -\infty$. It follows that

$$\int k'(v;e)\,d\psi(v) = \int k'\left(v - \frac{1}{1-\beta}\log(e); 1\right)\,d\psi(v) = 0$$

defines a unique value of $\hat{e}$ for any initial distribution ψ. The monotonicity of $\hat{e}(\psi)$ then follows immediately by using the fact that $k'(\cdot; 1)$ is a strictly decreasing function.

Proof of Proposition 5

(a) If utility is also bounded below, then the result follows from part (b). So suppose utility is unbounded below, but bounded above. Then $k'(g^w(\bar{\theta},\cdot))$ is continuous and Proposition 2 implies that $\lim_{v\to-\infty} k'(g^w(\bar{\theta},v)) = 1$. It follows that $\max_v k'(g^w(\bar{\theta},v))$ is attained, so there exists a $v_L > -\infty$ such that $g^w(\bar{\theta},v) > v_L$.

(b) If utility is bounded below, the result follows immediately from part (c) of Theorem 2.

(c) Using the first-order conditions from the proof of Proposition 2, one can show that:

$$\frac{c'\left(u\left(\bar{\theta}\right)\right)}{c'\left(u\left(\underline{\theta}\right)\right)} \leq \frac{\bar{\theta}}{\underline{\theta}}.$$

With logarithmic utility this implies that $g^u(\bar{\theta},v) - g^u(\underline{\theta},v) \leq \log\left(\bar{\theta}/\underline{\theta}\right)$. The incentive constraint then implies that $g^w(\underline{\theta},v) - g^w(\bar{\theta},v) \leq (\bar{\theta}/\beta)\log(\bar{\theta}/\underline{\theta}) \equiv A$. It follows that $v_t(\hat{\theta}^{t-1}) \geq v_t(\theta^{t-1}) - tA$ for all pairs of histories θ^{t-1} and $\hat{\theta}^{t-1}$. Then

$$\beta^t\,\mathbb{E}_{-1}\left[v_t\left(\sigma^{t-1}(\theta^{t-1})\right)\right] \geq \beta^t\,\mathbb{E}_{-1}\left[v_t(\theta^{t-1})\right] - \beta^t tA.$$

From part (b) of Theorem 2 we have $\lim_{t\to\infty}\beta^t\,\mathbb{E}_{-1}[v_t(\theta^{t-1})] = 0$. Since $\lim_{t\to\infty}\beta^t tA = 0$, it follows that $\limsup_{t\to\infty}\beta^t\,\mathbb{E}_{-1}[v_t(\sigma^{t-1}(\theta^{t-1})] \geq 0$.

(d) If $\underline{\gamma} > 0$ then the bound in (18) implies that $k'\left(g^w\left(\bar{\theta},v\right)\right) \leq 1 - \beta/\hat{\beta}$ and the result follows immediately. If $\bar{\gamma} < 1$, then we can define $\kappa = 1 - (1 - \beta/\hat{\beta})/(1 - \bar{\gamma})$, and define v_H by $k'(v_H) = \kappa$. Then for all $v \leq v_H$ we have $g^w(\theta, v) \leq v$. It follows that the unique ergodic set is bounded above by v_H. We can now apply the argument in (a) so there exists a $v_L > -\infty$ such that $g^w\left(\bar{\theta}, v\right) > v_L$.

Proof of Proposition 6

Part (a) is immediate since by continuity of the policy functions, consumption is bounded. For part (b), recall that $\int k'(v)\, d\psi^*(v) = 0$ under the invariant distribution ψ^*. If utility is unbounded below then all solutions for consumption are interior. If utility is bounded below, then corner solutions with $g^c(\theta, v) = 0$ for some θ can only occur for low enough levels of v, so that $g^c(\theta, v)$ is bounded, for all θ in this compact set. Recall that for interior solutions

$$1 - k'(v) = \eta\, \mathbb{E}\left[\, c'\big(g^u(\theta, v)\big)\,\right] = \eta\, \mathbb{E}\left[\, c'\Big(u\big(c\big(g^u(\theta, v)\big)\big)\Big)\,\right]$$

Applying Jensen's inequality we obtain

$$c'\left(u\left(\int \mathbb{E}\big[c\big(g^u(\theta, v)\big)\big]\, d\psi^*(v)\right)\right) \leq \int \mathbb{E}\Big[c'\Big(u\big(c\big(g^u(\theta, v)\big)\big)\Big)\Big]\, d\psi^*(v) = 1.$$

The result then follows since $c'(U(c))$ is an increasing function of c.

Proof of Proposition 7

Let $S^R(\psi; e)$ denote the maximum value attained for the relaxed planning problem with $Q_t = \hat{\beta}^t$, distribution ψ and endowment e. Note that

$$S(\psi; e) \leq S^R(\psi; e),$$

with equality at steady state distributions ψ^*. Since $\int k(v; \eta) d\psi(v) + \eta e/(1 - \hat{\beta})$ represents the full Lagrangian function (which now includes the omitted term due to e in (11)), by duality (see Luenberger, 1969, Chapter 8.6)

$$S^R(\psi; e) - \int v \, d\psi(v) \leq \int k(v; \eta) \, d\psi(v) + \eta \frac{e}{1 - \hat{\beta}} - \int v \, d\psi(v),$$

with equality whenever ψ, η and e are such that at the allocation that attains $k(v, \eta)$ the intertemporal resource constraint (10) holds, which is true at the constructed steady state. Integrating the right-hand side by parts gives

$$S^R(\psi; e) - \int v \, d\psi(v) \leq \int (1 - \psi(v))(k'(v; \eta) - 1) \, dv + \eta \frac{e}{1 - \hat{\beta}}.$$

It follows that

$$S(\psi;e) - \int v\,d\psi(v) - \left(S(\psi^*;e) - \int v\,d\psi^*(v)\right)$$
$$\leq S^R(\psi;e) - \int v\,d\psi(v) - \left(S^R(\psi^*;e) - \int v\,d\psi^*(v)\right)$$
$$\leq \int \big(\psi^*(v) - \psi(v)\big)\big(k'(v;\eta) - 1\big)\,dv.$$

If $k'(v;\eta) < 1$ on the support of ψ^* then the last term is strictly negative for all ψ with $\psi(v) < \psi^*(v)$ for all v. Thus, ψ^* maximizes $S(\psi;e) - \int v\,d\psi(v)$ subject to $\psi(v) \leq \psi^*(v)$ for all v. Proposition 1 implies that $k'(v;\eta) < 1$ if utility is unbounded below, in which case consumption is strictly positive. Proposition 2 implies that when utility is bounded $k'(v;\eta) < 1$ if and only if consumption is strictly positive.

References

Aiyagari, Rao, and Fernando Alvarez, 1995. "Stationary Efficient Distributions with Private Information: A Tale of Kings and Slaves." mimeo.

Atkeson, Andrew, and Robert E. Lucas, Jr., 1992. "On Efficient Distribution with Private Information." *Review of Economic Studies*. v59 (3): 427–453.

Atkeson, Andrew, and Robert E. Lucas, Jr., 1995. "Efficiency and Equality in a Simple Model of Efficiency Unemployment Insurance." *Journal of Economic Theory*. 66: 64–88.

Bernheim, B. Douglas, 1989. "Intergenerational Altruism, Dynastic Equilibria and Social Welfare." *Review of Economic Studies*. 56 (1): 119–128.

Caplin, Andrew, and John Leahy, 2004. "The Social Discount Rate." *Journal of Political Economy*. 112 (6): 1257–68.

Farhi, Emmanuel, Narayana Kocherlakota, and Iván Werning, 2006. mimeo.

Farhi, Emmanuel, and Iván Werning, 2006. "Progressive Estate Taxation." Nber working papers, National Bureau of Economic Research.

Green, Edward J., 1987. "Lending and the Smoothing of Uninsurable Income." in Edward C. Prescott, and Neil Wallace (ed.), *Contractual Arrangements for Intertemporal Trade* . pp. 3–25, University of Minnesota Press, Minneapolis, MN.

Grunwald, G.K., R.J. Hyndman, L.M. Tedesco, and R.L. Tweedie, 2000. "Non-Gaussian Conditional Linear AR(1) Models." *Australian and New Zealand Journal of Statistics*. 42 (4): 479–495.

Luenberger, David G., 1969. *Optimization by Vector Space Methods*, John Wiley & Sons, Inc.

Phelan, Christopher, 1998. "On the Long Run Implications of Repeated Moral Hazard." *Journal of Economic Theory*. 79 (2): 174–91.

Phelan, Christopher, 2006. "Opportunity and Social Mobility." *Review of Economic Studies*. 73 (2): 487–505.

Sleet, Christopher, and Sevin Yeltekin, 2004. "Credible Social Insurance." mimeo.

Spear, Stephen, and Sanjay Srivastava, 1987. "On Repeated Moral Hazard with Discounting." *Review of Economic Studies*. 54 (4): 599–617.

Stokey, Nancy L., and Robert E. Lucas, Jr., 1989. *Recursive Methods in Economic Dynamics*, Harvard University Press.

Thomas, Jonathan, and Tim Worrall, 1990. "Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem." *Journal of Economic Theory*. 51 (2): 367–390.

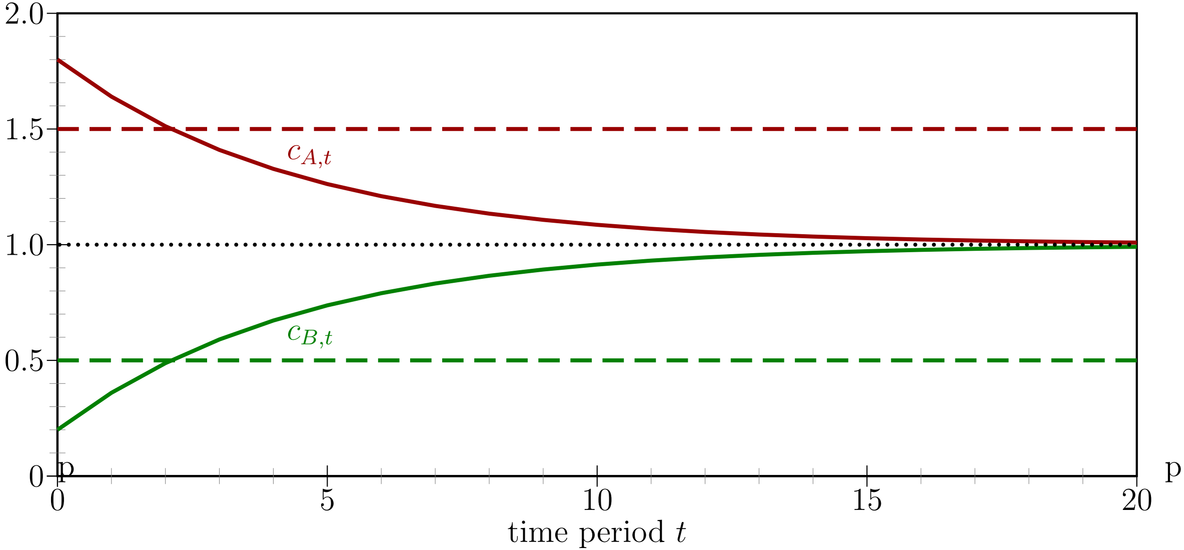

Figure 1: Consumption paths for groups A and B. Solid lines represent the case with $\hat{\beta} > \beta$; the dotted line at $c = e = 1$ is the steady state. The horizontal dashed lines represent the Atkeson-Lucas case with $\hat{\beta} = \beta$.

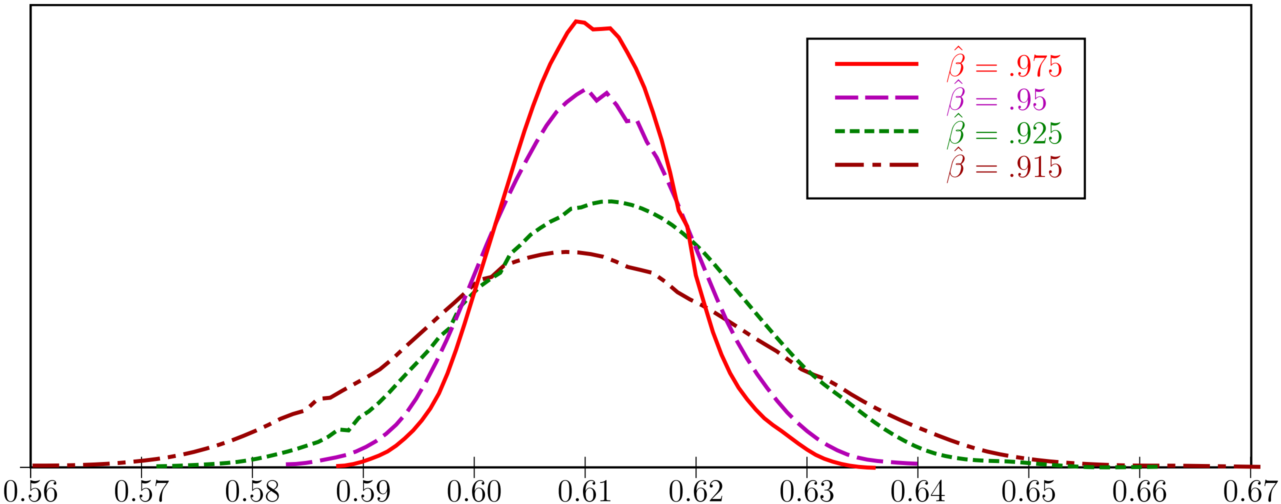

Figure 2: Invariant distributions for welfare, measured in consumption-equivalent units $c((1-\beta)v)$, for various values of $\hat{\beta}$.