NBER WORKING PAPER SERIES

INEQUALITY, SOCIAL DISCOUNTING AND ESTATE TAXATION

Emmanuel Farhi
Ivan Werning

Working Paper 11408
http://www.nber.org/papers/w11408

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
June 2005

Inequality, Social Discounting and Estate Taxation
Emmanuel Farhi and Ivan Werning
NBER Working Paper No. 11408
June 2005
JEL No. C61, D30, D63, H21, H23, H43

ABSTRACT

To what degree should societies allow inequality to be inherited? What role should estate taxation play in shaping the intergenerational transmission of welfare? We explore these questions by modeling altruistically-linked individuals who experience privately observed taste or productivity shocks. Our positive economy is identical to models with infinite-lived individuals where efficiency requires immiseration: inequality grows without bound and everyone's consumption converges to zero. However, under an intergenerational interpretation, previous work only characterizes a particular set of Pareto-efficient allocations: those that value only the initial generation's welfare. We study other efficient allocations where the social welfare criterion values future generations directly, placing a positive weight on their welfare so that the effective social discount rate is lower than the private one. For any such difference in social and private discounting we find that consumption exhibits mean-reversion and that a steady-state, cross-sectional distribution for consumption and welfare exists, where no one is trapped at misery. The optimal allocation can then be implemented by a combination of income and estate taxation. We find that the optimal estate tax is progressive: fortunate parents face higher average marginal tax rates on their bequests.

Emmanuel Farhi
Department of Economics
MIT, E52-251
50 Memorial Drive
Cambridge, MA 02142
efarhi@mit.edu

Ivan Werning
Department of Economics
MIT, E52-251
50 Memorial Drive
Cambridge, MA 02142
and NBER
iwerning@mit.edu

Inequality, Social Discounting and Estate Taxation*

Emmanuel Farhi
MIT
efarhi@mit.edu

Iván Werning
MIT, NBER and UTDT
iwerning@mit.edu

First Draft: June 2004
This Version: April 2005

Abstract

To what degree should societies allow inequality to be inherited? What role should estate taxation play in shaping the intergenerational transmission of welfare? We explore these questions by modeling altruistically-linked individuals who experience privately observed taste or productivity shocks. Our positive economy is identical to models with infinite-lived individuals where efficiency requires immiseration: inequality grows without bound and everyone's consumption converges to zero. However, under an intergenerational interpretation, previous work only characterizes a particular set of Pareto-efficient allocations: those that value only the initial generation's welfare. We study other efficient allocations where the social welfare criterion values future generations directly, placing a positive weight on their welfare so that the effective social discount rate is lower than the private one. For any such difference in social and private discounting we find that consumption exhibits mean-reversion and that a steady-state, cross-sectional distribution for consumption and welfare exists, where no one is trapped at misery. The optimal allocation can then be implemented by a combination of income and estate taxation. We find that the optimal estate tax is progressive: fortunate parents face higher average marginal tax rates on their bequests.

1 Introduction

Societies inevitably *choose* the inheritability of inequality. Some balance between equality of opportunity for newborns and incentives for altruistic parents is struck. We explore how this balancing act plays out to determine long-run inequality and draw some novel implications for optimal estate taxation.

*For useful discussions and comments we thank Daron Acemoglu, Fernando Alvarez, George-Marios Angeletos, Abhijit Banerjee, Gary Becker, Olivier Blanchard, Ricardo Caballero, Dean Corbae, Mikhail Golosov, Bengt Holmstrom, Narayana Kocherlakota, Robert Lucas, Casey Mulligan, Roger Myerson, Chris Phelan, Gilles Saint-Paul, Nancy Stokey, Jean Tirole and seminar and conference participants at Chicago, Minnesota, MIT and the Texas Monetary Conference held at the University of Austin in honor of the late Scott Freeman. This work begun motivated by a seminar presentation of Chris Phelan at MIT in May 2004. We also have gained significant insight from a manuscript by Scott Freeman and Michael Sadler—we thank Dean Corbae for bringing it to our attention.

Existing normative models of inequality reach an extreme conclusion: inequality should be perfectly inheritable and rise steadily without bound, with everyone converging to absolute misery and a vanishing lucky fraction to bliss. This *immiseration* result is robust; requires very weak assumptions on preferences (Phelan, 1998); and obtains invariably in partial equilibrium (Green, 1987, Thomas and Worrall, 1990), in general equilibrium (Atkeson and Lucas, 1992), and across environments with moral-hazard regarding work effort or with private information regarding preferences or productivity (Aiyagari and Alvarez, 1995).[1]

We depart minimally from these contributions, adopting the same positive economic models, but a slightly different normative criterion. In a generational context, previous work with infinite-lived agents characterizes 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 (2005) proposes a social planner with equal weights on all future generations. Our interest here is in exploring a class of Pareto-efficient allocations that take into account the current population along with unborn future generations. We place a positive and vanishing Pareto weight on the expected utility of future generations, this leads effectively to a social discount rate that is lower than the private one.

This relatively small change produces a drastically different result: long-run inequality remains bounded, a steady-state, cross-sectional distribution exists for consumption and welfare, social mobility is possible and everyone avoids misery. Indeed, welfare typically remains above an endogenous lower bound that is strictly better than misery. This outcome holds however small the difference between social and private discounting, and regardless of whether the source of asymmetric information is privately observed preferences or productivity shocks.

We begin by modeling a positive economy that is identical to the taste-shock setup developed by Atkeson and Lucas (1992). 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 only privately observed—thus ruling out first-best allocations. Feasible allocations must be incentive compatible and must satisfy the aggregate resource constraint in all periods.

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 from the desire to smooth the dynastic consumption path: rewards and punishments, required for incentives, are best delivered permanently. As a result, the consumption process inherits a random-walk component that leads cross-sectional inequality to grow endlessly without bound. Infinite spreading of the distribution is consistent with a constant aggregate endowment only if everyone's consump-

[1]Many find the immiseration result perplexing and some even find it morally questionable, but it is also inconvenient from a practical standpoint. Long-run steady-states often provide a natural benchmark to study dynamic economies, but such long-run analyses are not possible for private-information economies without a steady-state distribution with positive consumption. This has impaired the study of long-run implications of optimal taxation, so common in the Ramsey taxation literature.

tion eventually converges to zero. Note, that as a consequence, no steady-state, cross-sectional distribution with positive consumption exists.

Across generations, this arrangement requires a lock-step link between the welfare of parent and child. Of course, the perfect intergenerational transmission of welfare improves parental incentives—but at the expense of exposing newborns to the risk of their parent's luck. Individuals would value being insured against the uncertainty of their family's fortune—it is often recognized that one of the biggest risks one faces in life is regarding the family one is born into.

By contrast, it remains optimal to link the fortunes of parents and children in our model, 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 cross-sectional distribution for consumption and welfare, with no fraction of the population stuck at misery.

We also study a repeated Mirrleesian version of our economy and derive implications for optimal estate taxation. In this model, individuals have identical preferences with regard to consumption and work effort, but are heterogenous in the productivity of their work effort. Information about productivity and work effort is private—only the resulting output is publicly observable. We show that the analysis from the taste-shock model carries over to this setup, virtually without change. In particular, a very similar Bellman equation characterizes the solution to the social planning problem: consumption exhibits mean-reversion and has a steady-state cross-sectional distribution. This outcome highlights the fact that our results do not require any particular asymmetry of information.

More importantly, the Mirrleesian model offers new insights into estate taxation. Feasible allocations can be implemented by combining income and estate taxes. Specifically, we find that a progressive estate tax, which imposes a higher average marginal tax rate on the bequests of fortunate parents, is optimal. This result reflects the mean-reversion of consumption: more fortunate dynasties, with relatively high levels of current consumption, must have a declining consumption path induced by higher estate tax rates that lower the net rates of return across generations.

Finally, an important methodological contribution of this paper is to reformulate the social planning problem recursively. In doing so, we extend ideas introduced by Spear and Srivastava (1987) to situations where private and social preferences differ. Indeed, 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.

Related Literature. Our paper is most closely related to Phelan (2005), 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 the planning problem for any difference in social and private discounting. Unlike the case with no discounting, there is no associated static planning problem for steady-state distributions, and as a result, the methods we develop here are very different.

In overlapping-generation models without altruistic links, all market equilibria that are Pareto efficient place positive direct weight on future generations. Bernheim (1989) was the first to point out that in the dynastic extension of these models with altruism, many Pareto efficient allocations are not attainable by the market. Kaplow (1995) argued that these Pareto efficient allocations are natural social objectives and that they can be implemented by market equilibria with estate taxation policy. The estate tax is negative—it is a subsidy — so as to internalize the externality of giving on future generations.

Our work contributes to a large literature on dynamic economies with asymmetric information. In addition to the work mentioned above, this includes recent research on dynamic optimal taxation (e.g., Golosov, Kocherlakota and Tsyvinski, 2003; Albanesi and Sleet, 2004; and Kocherlakota, 2004). This application has been handicapped by the immiseration result and by the non-existence of a steady-state distribution with positive consumption, making it difficult to draw long-run conclusions for optimal taxation. Our results provide an encouraging way to overcome this problem.

Our work is also indirectly related to Sleet and Yeltekin (2004), who study an Atkeson-Lucas environment with a utilitarian planner, who lacks commitment and cares only for the current generation. In this environment, as in Phelan's, it is a foregone conclusion that immiseration will not obtain, so that the interesting question is how to solve for the best subgame-perfect equilibria. Sleet and Yeltekin derive first-order conditions from a Lagrangian and use these to numerically simulate the solution. Interestingly, it turns out that the best allocation in their no-commitment environment is asymptotically equivalent to the optimal one with commitment but featuring a more patient welfare criterion. Thus, our own approach and results provide an indirect, but effective way of characterizing the no-commitment problem and of formally establishing that a steady-state distribution with no one at misery exists.

The rest of the paper is organized as follows. 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 draw its connection to our original formulation. The resulting Bellman equation is then put to use in Section 4 to characterize the solution to the social planning problem. Here we derive our main results on mean-reversion and on the existence of a steady-state distribution for consumption. We discuss these results in Section 5 and develop intuition for them by studying some related problems and reformulations. In Section 6, we turn to the canonical optimal-taxation setup with productivity shocks and focus on its implications for estate taxes. Section 7 offers some conclusions from the analysis. All proofs omitted in the main text are contained in the Appendix.

2 A Social Insurance Problem

The backbone of our model requires a tradeoff between insurance and incentives. This tradeoff can be due to private information regarding either productivity or preferences. For purposes of comparison, we first adopt the Atkeson-Lucas taste-shock specification. In Section 6, we adapt our arguments to a repeated Mirrleesian model with privately observed productivity shocks. Similar arguments could be applied to moral-hazard situations with unobservable effort choices.

Our positive economy is identical to that of Atkeson-Lucas—the differences are only normative. An infinite-lived agent can be interpreted as a dynasty of individuals who have finite lives but are altruistically linked. Under this interpretation, Atkeson-Lucas and others focus on a particular set of efficient allocations: those that only directly consider the welfare of the initial generation.[2] In contrast, our interest here lies with efficient allocations that directly weigh the welfare of all future generations. This approach is asymptotically equivalent to postulating preferences for an infinitely-lived social planner who is more patient than individuals.

Demography, Preferences and Technology. At any point in time, our 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 utility v_t satisfies

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

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 utility v_{t+1}. The utility function $u(c)$ is assumed continuous and concave, with a continuous derivative for all $c > 0$ satisfying the Inada conditions $\lim_{c\to 0} u'(c) = \infty$ and $\lim_{c\to\infty} u'(c) = 0$. The taste shock $\theta \in \Theta$ is distributed identically and independently across individuals and time.

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 $e > 0$. These specifications choices preferences and technology are precisely those adopted by Atkeson-Lucas.

Define $U \equiv u(\mathbb{R}_+)$ to be the set of all possible utility values. Note that we allow utility to be unbounded so that the extremes $\underline{u} \equiv u(0)$ and $\bar{u} \equiv \lim_{c\to\infty} u(c)$ may be finite or infinite. The cost function $c(u)$ is defined on U as the inverse of the utility function $c \equiv u^{-1}$. For simplicity, we assume Θ contains a finite number of shocks $\underline{\theta} \equiv \theta_1 < \theta_2 < \cdots < \theta_N \equiv \bar{\theta}$. We denote the density by $p(\theta)$ and adopt the normalization that $\mathbb{E}[\theta] = \sum_{n=1}^{N} \theta_n \, p(\theta_n) = 1$. The level of dynastic utility v_t then always belongs to the set $V \equiv u(\mathbb{R}_+)/(1-\beta)$ with extremes $\underline{v} \equiv \underline{u}/(1-\beta)$ and $\bar{v} \equiv \bar{u}/(1-\beta)$.

Social Welfare. We depart from Atkeson-Lucas by assuming that the social welfare criterion can be represented by preferences given by the utility function

$$\sum_{t=0}^{\infty} \hat{\beta}^t \, \mathbb{E}_{-1}[\,\theta_t u(c_t)\,], \tag{2}$$

[2] The final paragraph in Atkeson Lucas (1992) discusses the possible importance of relaxing this assumption.

with $\hat{\beta} > \beta$. Thus, social preferences are identical to the individual preferences given by (1), except for the discount factor.

This setup puts weight on the welfare of future generations directly. 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. To see this, consider the utilitarian welfare criterion

$$\sum_{t=0}^{\infty} \alpha^t \, \mathbb{E}_{-1} v_t = \sum_{t=0}^{\infty} \delta_t \, \mathbb{E}_{-1} \left[\theta_t u\left(c_t\right) \right], \tag{3}$$

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

$$\frac{\delta_{t+1}}{\delta_t} = \beta + \frac{\alpha^{t+1}}{\delta_t} > \beta$$

and social preferences are more patient. In the limit $\delta_{t+1}/\delta_t \to \max\{\beta, \alpha\}$, so the welfare criterion (3) approaches (2) with $\hat{\beta} = \max\{\beta, \alpha\}$.[3,4]

Atkeson-Lucas' analysis applies to the case with $\beta = \hat{\beta}$, so we focus on the case where enough weight is placed on future generations to ensure that the long-run social discount factor remains strictly higher than the private one, $\hat{\beta} > \beta$. Although we adopt the preference in (2) directly for the rest of the paper, it is straightforward to adapt the arguments to the welfare criterion (3). The two specifications are slightly different for any finite horizon but are identical for the long-run, which is our primary concern.

Information and Incentives. Taste shock realizations are privately observed by individuals and their descendants. The revelation principle then allows us to restrict our 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 $\theta^t \equiv (\theta_0, \theta_1, \ldots, \theta_t)$. 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 $\sigma^t : \Theta^t \to \Theta^t$. We use σ^* to denote the truth-telling strategy with $\sigma_t^*(\theta^t) = \theta_t$ for all $\theta^t \in \Theta^t$.

Given an allocation $\{c_t\}$, the utility obtained from any reporting strategy σ is

$$U\big(\{c_t\}, \sigma; \beta\big) \equiv \sum_{t=0}^{\infty} \sum_{\theta^t \in \Theta^{t+1}} \beta^t \theta_t u\big(c_t\big(\sigma^t(\theta^t)\big)\big) \Pr(\theta^t).$$

[3] Bernheim (1989) performs similar intergenerational discount factor calculations in his welfare analysis of a deterministic dynastic saving model. Caplin and Leahy (2005) 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.

[4] One can also adopt the more general welfare criterion $\sum_{t=0}^{\infty} \alpha_t \, \mathbb{E}_{-1} v_t$ for some sequence of positive Pareto weights $\{\alpha_t\}$. In particular, the sequence $\alpha_0 = (1-\hat{\beta})/(1-\beta)$ and $\alpha_t = \alpha_0 \hat{\beta}^t$ for $t \geq 1$ delivers $\delta_{t+1}/\delta_t = \hat{\beta}$ for all $t = 0, 1, \ldots$

An allocation $\{c_t\}$ is *incentive compatible* if truth-telling is optimal:

$$U\big(\{c_t\}, \sigma^*; \beta\big) \geq U\big(\{c_t\}, \sigma; \beta\big). \tag{4}$$

Social Planning Problem. Following Atkeson-Lucas, we identify each dynasty with a number v, which we interpret as its initial entitlement to expected, discounted utility, $v_0 = v$. We assume that all dynasties with the same entitlement v receive the same treatment. We then let ψ denote a distribution of utilities v across the population of dynasties: $\psi(A)$ is the fraction of dynasties who will receive expected discounted utility in the set $A \subset \mathbb{R}$.

An allocation is a sequence of functions $\{c_t^v\}$ for each v, where $c_t^v(\theta^t)$ represents the consumption that a dynasty with initial entitlement v gets at date t after reporting the sequence of shocks θ^t. For any given initial distribution of entitlements ψ and resources e, we say that an allocation $\{c_t^v\}$ is *feasible* if: (i) it is incentive compatible for all dynasties; (ii) it delivers expected utility of at least v to all initial dynasties entitled to v; and (iii) average consumption in the population does not exceed the fixed endowment e in all periods. We let $e^*(\psi)$ denote the lowest resource level e such that there exists a feasible allocation that delivers the distribution of utility entitlements ψ—the efficiency problem studied by Atkeson and Lucas (1992) which is relevant for $\beta = \hat{\beta}$.

A social optimum maximizes the average social welfare function (2), weighed by ψ, over all feasible allocations. That is, the *social planning problem* given an initial distribution of entitlements ψ and an endowment level e is to maximize

$$\int U\big(\{c_t^v\}, \sigma^*, \hat{\beta}\big)\, d\psi(v)$$

subject to $v = U(\{c_t^v\}, \sigma^*; \beta) \geq U(\{c_t^v\}, \sigma; \beta)$ for all v, and

$$\int \sum_{\theta^t} c_t^v(\theta^t) \Pr(\theta^t)\, d\psi(v) \leq e \qquad t = 0, 1, \ldots \tag{5}$$

Our social planning problem is well defined, with a non-empty constraint set, for all $e \geq e^*(\psi)$; we are interested in situations with $\hat{\beta} > \beta$ and where $e > e^*(\psi)$.

Steady States. Our focus is on distributions of utility 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. We also require the cross-sectional distribution of consumption to replicate itself over time. We term any initial distribution of entitlements with these properties a *steady state* and denote them by ψ^*. As we shall demonstrate below, continuation utility constitutes a state variable that follows a Markov process, and steady states are then invariant distributions of this process.

Note that 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 $e^*(\psi)$ the sequence of distributions converges weakly to the distribution having full mass at misery, with zero consumption for everyone. We seek non-trivial steady states ψ^* that exhaust a strictly positive aggregate endowment e in all periods.

3 A Bellman Equation

In this section we study a relaxed version of the social planning problem whose solution coincides with that of the original problem at steady states. The relaxed problem has two important advantages. First, the relaxed problem can be solved by studying a set of subproblems—one for each dynasty with entitlement v—which avoids the need to keep track of the entire population. Second, each of these subproblems admits a simple recursive formulation, which can be characterized quite sharply. We believe that the general approach we develop here may be useful in other contexts.

Consider the *relaxed planning problem* where the sequence of resource constraints (5) is replaced by the single intertemporal condition

$$\int \sum_{t=0}^{\infty} Q_t \sum_{\theta^t} c_t^v(\theta^t) \Pr(\theta^t)\, d\psi(v) \leq e \sum_{t=0}^{\infty} Q_t, \tag{6}$$

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 relaxed and original versions of the planning problem are related in that any solution to the former which happens to satisfy the resource constraints in (5) must also be a solution to the latter. A Lagrangian argument establishes the converse: there must exist some positive sequence $\{Q_t\}$ such that the solution to the original planning problem also solves the relaxed problem. Most importantly, any steady-state solution to the relaxed problem is a steady-state solution to the original one.

Our focus on steady states leads naturally to $Q_t = q^t$ for some $q > 0$. Indeed, steady states are only compatible with $q = \hat{\beta}$, so we adopt this value for the relaxed problem from this point forward. Attaching a multiplier $\hat{\lambda} > 0$ to the intertemporal resource constraint (6), we can form the Lagrangian $L \equiv \int L^v \, d\psi(v)$ where

$$L^v \equiv \sum_{t=0}^{\infty} \sum_{\theta^t} \hat{\beta}^t \big(\theta_t u(c_t^v) - \hat{\lambda} c_t^v(\theta^t)\big) \Pr(\theta^t)$$

and study the optimization of L subject to $v = U(\{c_t^v\}, \sigma^*; \beta) \geq U(\{c_t^v\}, \sigma; \beta)$ for all v. This is equivalent to the pointwise optimization, for each v, of the subproblem: $k(v) \equiv \sup L^v$ subject to $v = U(\{c_t^v\}, \sigma^*; \beta) \geq U(\{c_t^v\}, \sigma; \beta)$. Our first result characterizes this value function and shows that it satisfies a Bellman equation.

Theorem 1 *The value function $k(v)$ is continuous, concave, and satisfies the Bellman equation*

$$k(v) = \max_{u,w} \mathbb{E}\big[\theta u(\theta) - \hat{\lambda} c\big(u(\theta)\big) + \hat{\beta} k\big(w(\theta)\big)\big] \tag{7}$$

subject to

$$v = \mathbb{E}[\theta u(\theta) + \beta w(\theta)] \tag{8}$$

$$\theta u(\theta) + \beta w(\theta) \geq \theta u(\theta') + \beta w(\theta') \qquad \textit{for all} \quad \theta, \theta' \in \Theta. \tag{9}$$

This recursive formulation imposes a promise-keeping constraint (8) and an incentive constraint (9). 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 (4). Intuitively, the rest is implicitly taken care of in (7) by evaluating the value function at the continuation utility: 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. Then $k(w(\theta))$ represents the value to the planner of this continuation allocation. Taken together, a pair $u(\theta)$ and $w(\theta)$ that satisfies (8)–(9) pasted with the corresponding continuation allocations for each $w(\theta)$, yields an allocation that satisfies the full incentive-compatibility (4). The objective function in (7) then captures the relevant value of allocations constructed in this way.

Among other things, Theorem 1 shows that the maximum in the Bellman equation (7) is attained. We let the policy functions $g^u(\theta, v)$ and $g^w(\theta, v)$ denote the unique solutions for u and w, respectively. For any initial utility entitlement v_0, an allocation $\{u_t\}$ can then be generated from the policy functions (g^u, g^w) by setting $u_0(\theta_0) = g^u(\theta_0, v_0)$ initially and defining $u_t(\theta^t)$ and $v_{t+1}(\theta^t)$ inductively for $t \geq 1$ by $u_t(\theta^t) = g^u(\theta_t, v_t(\theta^{t-1}))$ and $v_{t+1}(\theta^t) = g^w(\theta_t, v_t(\theta^{t-1}))$.

Our next result elucidates the connection between allocations generated from the policy functions in this way and solutions to the planning problem.

Theorem 2 ***(a)*** *An allocation $\{u_t\}$ is optimal for the relaxed problem, given v_0, if and only if it is generated by the policy functions (g^u, g^w) starting at v_0, is incentive compatible, and delivers a lifetime utility of 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 utility 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) of Theorem 2 implies that either the solution to the relaxed planning problem is generated by the policy functions of the Bellman equation, or there is no solution at all. Parts (b) and (c) of the theorem show that the first case is guaranteed if we can verify the limit condition in part (c). The latter is automatically satisfied for all utility functions that are bounded below and can be verified in many other cases of interest.[5]

[5]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

The case with $\beta = \hat{\beta}$ can be studied by the same approach. Recall that the efficiency problem studied by Atkeson and Lucas (1992) minimizes resources e subject to the sequence of resource constraints (5) and $v = U(\{c_t^v\}, \sigma^*; \beta) \geq U(\{c_t^v\}, \sigma; \beta)$ for all v. Consider the relaxed version of this problem that replaces the sequence of resource constraints with the single intertemporal constraint (6) for some sequence $\{Q_t\}$. Then if the solution to this problem satisfies the resource constraints (5) it is also a solution to the original problem. Although no steady state exists in this case, with constant relative risk aversion utility functions, the relaxed problem with $Q_t = q_{AL}^t$ characterize the original one, for an appropriately chosen value of $q_{AL} > 0$, not necessarily equal to β.

Since the constraint (6) binds, we can take the objective function for the relaxed problem as the left hand side of this inequality. This minimization can then be done pointwise: for each v let $K_{AL}(v) \equiv \inf \sum_{t=0}^{\infty} q_{AL}^t c_t(\theta^t)$ subject to $v = U(\{c_t\}, \sigma^*; \beta) \geq U(\{c_t\}, \sigma; \beta)$. The associated Bellman equation for this problem is then

$$K_{AL}(v) = \min_{u,w} \mathbb{E}\big[c\big(u(\theta)\big) + q_{AL} K_{AL}\big(w(\theta)\big)\big] \tag{10}$$

subject to (8) and (9). This problem can be thought of as the limiting version of the $\hat{\beta} > \beta$ case as $\hat{\lambda} \to \infty$ and where the discount factor in the objective q_{AL} is not necessarily $\hat{\beta}$. Theorems 1 and 2 also apply to this problem and its Bellman equation.

4 Optimal Inequality

In this section we exploit the connection between the Bellman equation and the planning problem. We characterize the solution and derive a key equation that illustrates mean-reverting forces in the dynamics of consumption. The main result of the section is to establish that these forces are strong enough to imply the existence of an invariant distribution with no misery. Finally, we provide sufficient conditions to verify part (c) of Theorem 2, and ensure that a solution to the planning problem exists.

Mean Reversion

We are now in a position to study the Bellman equation's optimization problem. To begin, we justify the use of first-order conditions with the following lemma:

Lemma 1 *The value function $k(v)$ is strictly concave and 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*

problem with state variable v_t. Conditions (8) and (9) 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 part (c) of Theorem 2.

$\lim_{v \to \underline{v}} k'(v) = \infty$.

Let $\lambda = k'(v)$ be the multiplier on the left-hand side of the promise-keeping constraint (8) and let $\mu(\theta, \theta')$ be the multipliers on the incentive constraints (9). The first-order condition for $u(\theta)$ is

$$\Big(1 - \hat{\lambda} c'\big(u(\theta)\big)\Big) p(\theta) - \theta \lambda p(\theta) + \sum_{\theta'} \theta \mu(\theta, \theta') - \sum_{\theta'} \theta' \mu(\theta, \theta') \leq 0,$$

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

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

Using the envelope condition $k'(v) = \lambda$ and adding up across θ, this becomes

$$\sum_{\theta \in \Theta} k'\big(g^w(\theta, v)\big) p(\theta) = \frac{\beta}{\hat{\beta}} k'(v). \tag{11}$$

This key equation can be represented in sequential notation as

$$\mathbb{E}_{t-1}\big[k'\big(v_{t+1}(\theta^t)\big)\big] = \frac{\beta}{\hat{\beta}} k'\big(v_t(\theta^{t-1})\big) \tag{12}$$

where $\{v_t\}$ is generated by the policy function g^w. Thus, $\{k'(v_t)\}$ is a Conditional Linear Auto Regressive (hereafter: CLAR) Markov process. Note that we can translate anything about the process $\{k'(v_t)\}$ into implications for the process $\{v_t\}$, since the derivative $k'(v)$ is continuous and strictly decreasing. Likewise, using the policy function $g^u(\theta, v)$, conclusions about the process $\{v_t\}$ provide information about the process for consumption.

The conditional expectation in (12) illustrates that $\beta/\hat{\beta} < 1$ creates a force for mean reversion for the process $\{k'(v_t)\}$ toward zero. Lemma 1 implies that the value function $k(v)$ has an interior maximum at $v^* > \underline{v}$ with $k'(v^*) = 0$, so reversion occurs towards this interior utility level—away from misery. This feature is key to our results on the existence of invariant distributions.

Economically, the mean-reversion equation itself embodies an interesting form of social mobility. We can divide the population into two social hierarchies, with mobility ensured between them. Descendants of individuals with current welfare above v^* will eventually fall below it. Similarly dynasties initially entitled to welfare below v^* are guaranteed to access levels above it. This rise and fall of families illustrates a strong intergenerational mobility in the model.

In deriving this result, it is important to stress the role played by the non-monotonicity of the value function $k(v)$. Although mean reversion stems from $\hat{\beta}/\beta < 1$ in equation (12), it is non-monotonicity of $k(v)$ that ensures that reversion is not toward misery. By contrast, in the Atkeson-Lucas case a CLAR equation similar to (12) may hold, but the value function in this case is monotone and reversion then occurs toward misery. Indeed, the envelope and first-order conditions for the Bellman equation (10) yield

$$\sum_{\theta} K'_{AL}\big(g^w_{AL}(\theta, v)\big)\, p(\theta) = \frac{\beta}{q_{AL}} K'_{AL}(v),$$

which is similar to condition (12) when $q_{AL} > \beta$.[6] Crucially, unlike the case with $\hat{\beta} > \beta$, here the value function $K_{AL}(v)$ is strictly increasing, so that $K'_{AL}(v) \geq 0$. Thus, $\{K'_{AL}(v_t)\}$ is a non-negative process, which implies by the Martingale Convergence Theorem that it must converge almost surely (a.s.) to some finite value. Since incentives must be provided using continuation utilities $g^w(\bar{\theta}, v) \neq g^w(\underline{\theta}, v)$, this rules out anything other than $K'_{AL}(v_t) \to 0$ a.s. Immiseration then follows, $v_t \to \underline{v}$ and $c_t \to 0$ a.s. This highlights the importance of the non-monotonicity of the value function $k(v)$ for our results in the case of $\hat{\beta} > \beta$.

Our next result pushes the characterization of reversion past the average behavior of the $\{k'_t\}$ process by deriving bounds for its evolution. These bounds are critical for guaranteeing the existence of an invariant distribution with no mass at misery.

Proposition 1 *The policy function $g^w(\theta, v)$ satisfies the CLAR equation (11). In addition:*

(a) *if utility is unbounded below, then*

$$\underline{\gamma}\big(1 - k'(v)\big) + \left(1 - \frac{\beta}{\hat{\beta}}\right) \leq 1 - k'\big(g^w(\theta, v)\big) \leq \bar{\gamma}\big(1 - k'(v)\big) + \left(1 - \frac{\beta}{\hat{\beta}}\right) \tag{13}$$

for all $\theta \in \Theta$, where the constants are given by $\bar{\gamma} \equiv (\beta/\hat{\beta}) \max_{1 \leq n \leq N} \{(1 + \theta_n - \mathbb{E}[\theta \leq \theta_n])/\theta_n\}$ and $\underline{\gamma} \equiv (\beta/\hat{\beta}) \min_{2 \leq n \leq N} \{1 + \theta_{n-1} - \mathbb{E}[\theta \mid \theta \geq \theta_n]/\theta_{n-1}\}$.

(b) *if utility is bounded below, then for low enough values of v such that $k'(v) > 1$, we have*

$$\begin{aligned} u(\theta) &= \underline{u} \\ w(\theta) &> v \\ k'\big(w(\theta)\big) &= (\hat{\beta}/\beta) k'(v) \end{aligned}$$

for all $\theta \in \Theta$. For values of v such that $k'(v) \leq 1$, the lower bound in (13) holds; the upper bound in (13) holds for sufficiently high v.

Proposition 1 illustrates a powerful tendency away from misery. For example, with utility unbounded below, continuation utility $g^w(\theta, v)$ remains bounded even as $v \to -\infty$. Thus, no matter how much a parent is supposed to be punished, his child is always somewhat spared.

[6] With logarithmic utility $q_{AL} = \beta$ yields a solution with constant average consumption. With $u(c) = c^{1-\sigma}/(1-\sigma)$ and $\sigma < 1$ the appropriate value of q_{AL}, that yields constant consumption, is strictly above β.

Main Result: Existence of an Invariant Distribution with No Misery

We now state the main result of this section: if a solution to the relaxed planning problem exists, then it admits an invariant distribution with no misery. The proof of this result relies on the conditional-expectation equation (12) and the bounds in Proposition 1. Thus, it makes use of both mean-reversion properties discussed in the previous subsection.[7]

Proposition 2 *The existence of an invariant distribution ψ^* with no mass at misery, $\psi^*(\{\underline{v}\}) = 0$, for the Markov process $\{v_t\}$ implied by g^w is guaranteed if either: utility is unbounded below, utility is bounded above, or $\hat{\gamma} < 1$.*

Proposition 2, when combined with part (a) of Theorem 2, leaves open only two possibilities: (i) the relaxed problem admits a steady-state invariant distribution with no misery; or (ii) no solution exists. This situation contrasts strongly with the Atkeson-Lucas case, with $\beta = \hat{\beta}$, where a solution exists but does not admit a steady state, and everyone ends up at misery. Towards the end of this section we show that a solution to the planning problem can be guaranteed so that case (i) holds.

Our Bellman equation also provides an efficient method for explicitly 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 $u(c) = 4c^{1/2}$ for $c \geq 0$ 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 relaxed problem where we ignore the non-negativity constraints on u and w,

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

subject to (8) and (9). This is a linear-quadratic dynamic programming problem, so it follows that the value function is a quadratic and the policy functions are linear in v:

$$\begin{aligned} g^u(\theta, v) &= \gamma_1^u(\theta) v + \gamma_0^u(\theta) \\ g^w(\theta, v) &= \gamma_1^w(\theta) v + \gamma_0^w(\theta) \end{aligned}$$

For taste shocks with sufficiently small amplitude we can guarantee, by continuity with the deterministic case $\underline{\theta} = \bar{\theta}$, that $\gamma^w(\underline{\theta}) < 1$ and $\gamma^w(\bar{\theta}) > 0$, implying a unique bounded ergodic set for utility $[v_L, v_H]$ with $v_L > 0$. Moreover, $g^u(\theta, v) > 0$ for $v \in [v_L, v_H]$. Hence, since the planning problem is convex and utility turns out to be strictly positive at the steady state, this solution does solves the original problem with non-negativity constraints on u.

[7] When utility is bounded below, we either require that utility be bounded above, or that $\hat{\gamma} < 1$, which is ensured for a small dispersion of the shocks, as a simple way of ensuring that the ergodic set is bounded away from misery. It seems very plausible, however, that these conditions could be dispensed with.

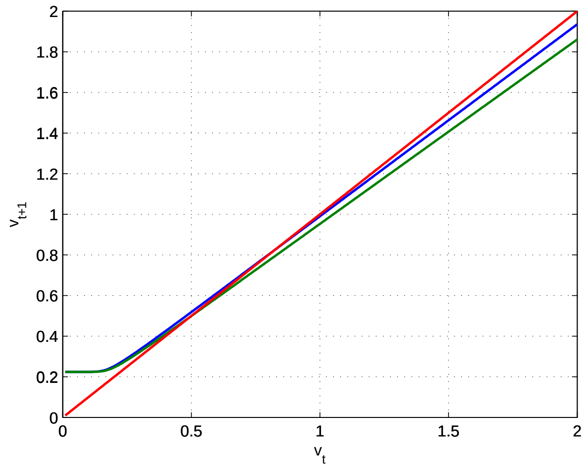

Figure 1: Policy functions $g^w(\theta, v)$

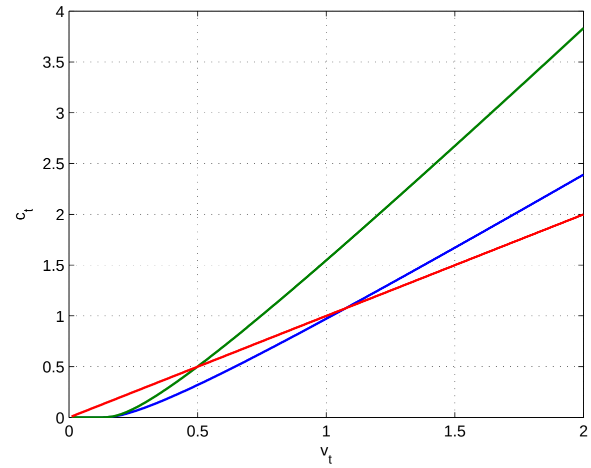

Figure 2: Policy function $g^c(\theta, w)$

Example 2. To illustrate the numerical value of our recursive formulation, we compute the solution for the logarithmic case with $\beta = 0.9$, $\hat{\beta} = 0.975$, $e = \hat{\lambda}^{-1} = 0.6$, $\theta_h = 1.2$, $\theta_l = 0.75$ and $p = 0.5$. We iterated on the Bellman equation for $k(v)$ until convergence.[8]

Figure 1 plots the policy function for continuation utility in consumption-equivalent units, $c(v(1-\beta))$ against $c(g^w(\theta, v)(1-\beta))$, while Figure 2 does the same for the policy function for consumption, $c(v(1-\beta))$ against $g^c(\theta, v)$. Both policy functions are monotonic and smooth. Figure 1 reveals a unique, bounded ergodic set for v. Note that both policy functions become nearly flat for low values of v. This illustrates the result, discussed immediately after Proposition 1, that utility is kept above some endogenous bound.

Figure 3 displays the steady-state, cross-sectional distribution of dynastic utility measured in consumption-equivalent units, $c(v(1-\beta))$ implied by the solution to the planning problem.[9] The long-run distribution has a smooth bell-curve shape—a feature that 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. The figure also shows the invariant distributions for various values of

[8] The details of this numerical exercise where as follows: we solved for $u(\theta)$ as a function of $w(\theta)$ using the incentive and promise-keeping constraints. We then maximized over $w(\theta)$. We employed a grid for v defined in terms of equally spaced consumption-equivalent units $c((1-\beta)v) = \{0.01, \ldots, 2\}$. Results with a grid size of 100 and 300 were similar; we report the latter. We used Matlab's splines package to interpolate the value function and used `fmincon.m` as our optimization routine over $w(\theta)$. Our iterations were initialized with the value function corresponding to the feasible plan that features constant consumption:

$$k_0(v) = \frac{v(1-\beta) - \hat{\lambda} c(v(1-\beta))}{1-\hat{\beta}}$$

We stopped the iterations when $\|k_n(v) - k_{n-1}(v)\| < 10^{-10}$ and verified that the policy functions had also converged. Note that $g^w(\theta, v)$ is well within the interior of $[.01, 2]$, so that the arbitrary upper and lower bounds from our grid choice were not found to be binding.

[9] The invariant distribution was approximated by generated a Monte Carlo simulation for the dynamics of the $\{v_t\}$ process generated by g^w, with an arbitrary initial value of v_0. Since this process converges to a unique invariant distribution ψ^*, starting from any initial value of v_0, the frequencies in a long time-series sample approach the frequencies of ψ^*. To create the figure we used Matlab's Wavelet Toolbox to approximate the density from the simulated Monte-Carlo sample.

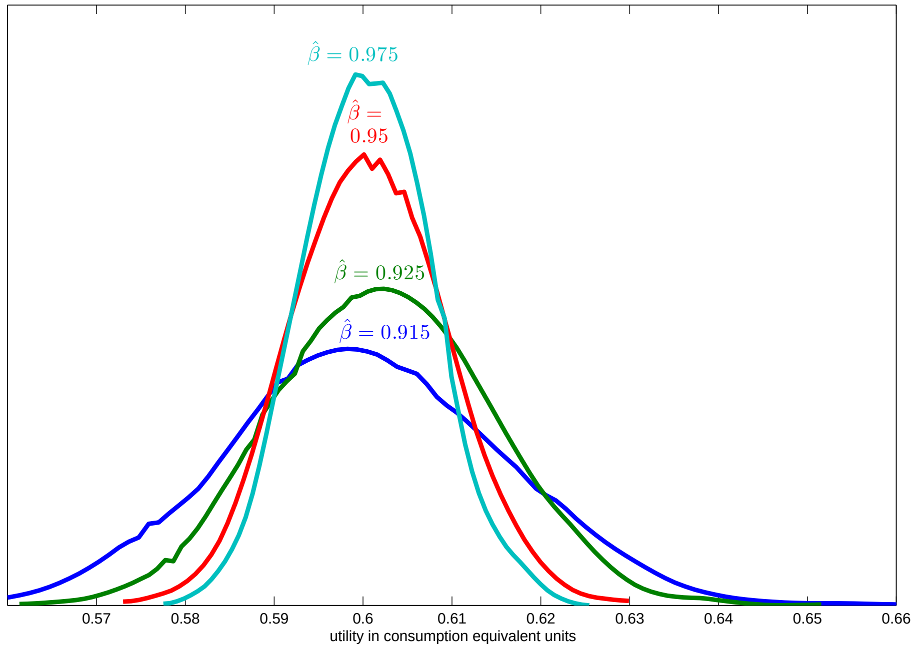

Figure 3: Steady-State Distributions of Dynastic Utility

$\hat{\beta}$. The degree of inequality appears to decrease with higher values of $\hat{\beta}$. This outcome is suggested intuitively by the coefficient on the CLAR equation (12) and the discussion in Section 5 on features of the impulse response to shocks. These simulations also support the natural conjecture that as we approach the Atkeson-Lucas case, $\hat{\beta} \to \beta$, the resulting sequence of invariant distributions blows up, since no steady state with positive consumption exists when $\hat{\beta} = \beta$.

We now turn briefly to issues of uniqueness and stability for the invariant distribution guaranteed by Proposition 2. This question is of economic interest because it represents an even stronger notion of social mobility than that implied by the mean-reversion condition (12) discussed in the previous subsection. That is, if convergence toward the distribution ψ^* occurs starting from any initial utility level v_0, then the fortunes of distant descendants—the distribution of their welfare—is independent of the individual's present condition. At the optimum, the past always exerts some influence on the present, but its influence is bounded and dies out over time, so that the advantages or disadvantages of distant ancestors are eventually wiped out.

Indeed, under some conditions we can guarantee that the social optimum in our model does display this strong notion of social mobility. To see this, suppose the ergodic set for the $\{k'_t\}$ process is compact. This is guaranteed, for example, by applying Proposition 1 when $\bar{\gamma} < 1$. Then, if the policy function $g^w(\theta, v)$ is monotone in v, the invariant distribution ψ^* is unique and stable in the sense that, starting from any initial distribution ψ_0, the sequence of distributions $\{\psi_t\}$, generated by g^w, converges weakly to ψ^*. This follows since the conditional-expectation equation (12) ensures enough mixing to apply Hopenhayn-Prescott's Theorem.[10] The monotonicity of the

[10]See pg. 382-383 in Stokey and Lucas with Prescott (1989).

policy functions for continuation utility w seems intuitive and plausible, as illustrated by Examples 1 and 2.[11]

Another approach suggests uniqueness and convergence without relying on monotonicity. Grunwald, Hyndman, Tedesco and Tweedie (1999) show that one-dimensional, irreducible Markov processes with the Feller property that are bounded below and satisfy a CLAR condition, such as (12), have a unique and stable invariant distribution. Moreover, convergence to this distribution from any initial distribution is fast, in the sense that it occurs at the geometric rate $\beta/\hat{\beta}$. All the requirements of their theorem have been verified already for our model, except for the technical condition of irreducibility, which is likely to hold if we were to assume that the taste shock has a continuous distribution. We do not pursue this formally other than to note that the forces for reversion in (12) could be further exploited to establish uniqueness and convergence.

Our focus on steady states, where the distribution of utility entitlements replicates itself over time, has exploited the fact that the relaxed and original planning problems must coincide. However, for the logarithmic utility case we can do more and characterize transitional dynamics.

Proposition 3 *If utility is logarithmic, then for any initial distribution of utility entitlements ψ there exists an endowment level $e^*(\psi)$ such that the solution to the original social planning problem is generated by the policy functions (g^u, g^w) from the relaxed problem with $Q_t = \hat{\beta}^t$. The function e^* is monotone increasing, in that if $\psi^a \prec \psi^b$ in the sense of first-order stochastic dominance then $e^*(\psi^a) < e^*(\psi^b)$.*

One interesting application of this result is to the situation where the planning problem is modified to select the best initial distribution ψ, instead of taking one as given. Then all initial dynasties are treated identically and started with identical utility level v^* solving $k'(v^*) = 0$. The optimal allocation then evolves according to the dynamics implied by the policy function $g^w(\theta, v)$. The cross-sectional distribution of welfare will spread out from its initially egalitarian condition as dynasties experience varying luck in the realization of their shocks.

By applying Proposition 3, convergence to a unique invariant distribution ψ^* of the Markov process $\{v_t\}$ implied by the policy function $g^w(\theta, v)$ takes on additional economic meaning. It implies the stability of the cross-sectional distributions of welfare and consumption in the population. That is, if the Markov process $\{v_t\}$ generated by g^w is stable, then the cross-sectional distributions of welfare and consumption eventually settle down to the steady state.

As mentioned in Section 3, for any utility function specification one can characterize the solution for any (ψ, e) as the solution to a relaxed problem with some sequence of prices $\{Q_t\}$, that are not necessarily exponential. Proposition 3 identifies the distributions and endowment pairs (ψ, e) that lead to exponential prices in the logarithmic case. More generally, with logarithmic utility for *any* pair (ψ, e), we can show that $Q_t = \hat{\beta}^t + \Lambda\beta^t$ for some constant Λ. The entire optimal allocation can then be characterized by the policy functions from a non-stationary Bellman equation. Since

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

prices are asymptotically exponential, in that $\lim_{t\to\infty} \hat{\beta}^{-t} Q_t = 1$, it follows that long-run dynamics are always dominated by the policy functions (g^u, g^w) from the relaxed problem with exponential prices $Q_t = \hat{\beta}^t$ that we have characterized.

Sufficient Conditions for Verification

We conclude this section by describing sufficient conditions for a solution to the planning problem to exist at the steady state ψ^* identified by the policy functions in Proposition 2. This involves two steps. First, we establish that allocations generated by the policy functions are indeed incentive compatible by verifying the condition in part (c) of Theorem 2. Second, we verify that average consumption is finite under the invariant distribution ψ^*.

Lemma 2 *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 $\underline{\gamma} > 0$.*

We now find sufficient conditions that guarantee average consumption is finite under the invariant distribution ψ^*. If the ergodic set for utility v is bounded away from the extremes, then consumption is bounded and average consumption is trivially finite. Even when a bounded ergodic set for utility v cannot be ensured, finite average consumption can be guaranteed for a large class of utility functions.

Lemma 3 *Average consumption is finite under the invariant distribution ψ^**

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

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.

Note that a bounded ergodic set is guaranteed by $\bar{\gamma} < 1$, which is ensured for taste shocks with sufficiently small amplitude; and condition (b) holds, for example, for all constant relative risk aversion utility functions with $\sigma \geq 1$.

The value of average consumption depends on the value of $\hat{\lambda}$. For instance, in the case of constant relative risk aversion utility, average steady state consumption is a power function of $\hat{\lambda}$, 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.

5 Discussion: Mean-Reversion

This section develops an intuitive understanding of the key mean-reversion property discussed previously. We first derive the impulse response of consumption to a one-time taste shock. We then revisit the full problem with an alternative Bellman equation that is useful as a source of intuition.

Impulse Response

Consider a version of our model where only the first generation faces uncertainty. In the first period, there are two possible values for the taste shock $\theta_0 \in \{\theta_L, \theta_H\}$, but thereafter the economy is deterministic: $\theta_t = 1$ for $t \geq 1$. We compare this to the case with no uncertainty in the first period. This allows us to trace out the consumption response to the taste shock over time. To simplify, we adopt logarithmic utility.

We begin by studying a subproblem of the deterministic planning problem from the second generation onward, that is for $t = 1, 2, \ldots$. For a given, promised continuation utility v_1, the planning problem is

$$k_{\text{det}}(v_1) \equiv \max_{\{c_t\}} \sum_{t=1}^{\infty} \hat{\beta}^{t-1} \left(\log c_t - \hat{\lambda} c_t \right),$$

subject to

$$v_1 = \sum_{t=1}^{\infty} \beta^{t-1} \log c_t.$$

The associated Bellman equation is

$$k_{\text{det}}(v_t) = \max_{c_t, v_{t+1}} \left(\log(c_t) - \hat{\lambda} c_t + \hat{\beta} k_{\text{det}}(v_{t+1}) \right)$$

subject to $v_t = \log c_t + \beta v_{t+1}$. The first-order and envelope conditions imply that

$$k'_{\text{det}}(v_{t+1}) = \frac{\beta}{\hat{\beta}} k'_{\text{det}}(v_t), \tag{14}$$

$$c_t = \hat{\lambda}^{-1}(1 - k'_{\text{det}}(v_t)). \tag{15}$$

Condition (14) shows that $\{k'_{\text{det}}(v_t)\}$ reverts geometrically towards zero at the rate $\beta/\hat{\beta}$. This is a deterministic version of the conditional-expectation equation (12). In the logarithmic case, it translates directly into consumption by the first-order condition (15). Thus, consumption reverts back to a common steady state at the same rate; deviations from the steady-state level of consumption have a half-life of $(\log_2(\hat{\beta}/\beta))^{-1}$. Note that in the Atkeson-Lucas case when social and private discounting coincide, so that $\beta = \hat{\beta}$, consumption remains perfectly constant after the shock at its new level $c_t = c(v_1/(1-\beta))$.

Turning to the first generation at $t = 0$, the planning problem solves

$$\max_{c_0, v_1} \mathbb{E}\left[\theta_0 \log\left(c_0(\theta)\right) - \hat{\lambda}_0 c_0(\theta) + \hat{\beta} k_{\text{det}}\left(v_1(\theta)\right)\right]$$

subject to

$$\theta_L \log\left(c_0(\theta_L)\right) + \beta v_1(\theta_L) \geq \theta_L \log\left(c_0(\theta_H)\right) + \beta v_1(\theta_H).$$

where we omit the other incentive constraint since it is not binding at the optimum, and the problem is convex. At the optimum, $c_0(\theta_H) > c_0(\theta_L)$, $v_1(\theta_H) < v_1(\theta_L)$, and

$$\mathbb{E}\left[k'_{\text{det}}\left(v_1(\theta)\right)\right] = 0,$$

implying that $v_1(\theta_H) < v^* < v_1(\theta_L)$ where $k'(v^*) = 0$. Note that average consumption is constant and equal to $\hat{\lambda}^{-1}$ in all periods.

Figure 4 shows the consumption response to a taste shock in the first period, for subsequent periods. That is, we use (14) and (15) for $t \geq 1$ starting at $v_1(\theta_L)$, v^* and $v_1(\theta_H)$. The effect on consumption from the shock dies out over time and consumption returns to a common steady-state level. Again, this illustrates that the influence of past fortunes eventually vanishes for distant descendants. We also plot the Atkeson-Lucas case with $\beta = \hat{\beta}$, where the luck of the first generation has a permanent impact on the consumption of all descendants.

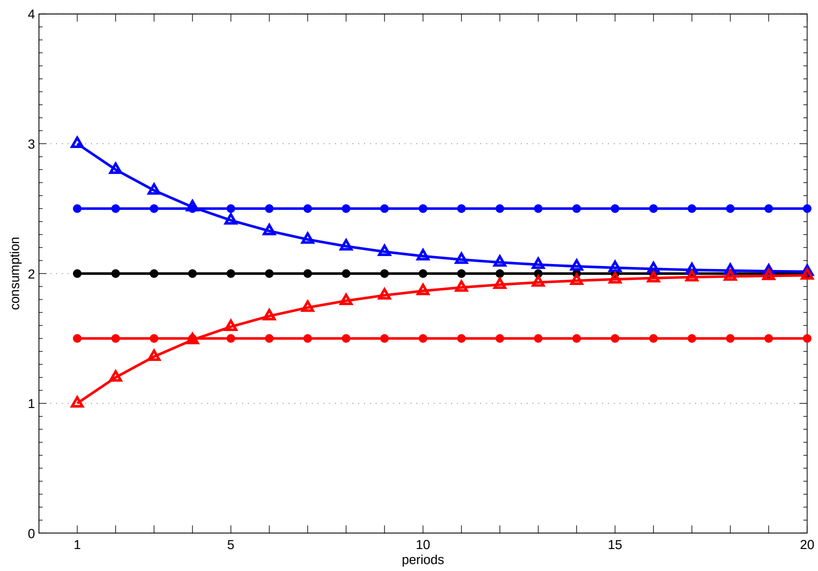

Figure 4: Consumption path for $t \geq 1$ in response to taste shock at $t = 0$

To provide incentives for the first generation, society rewards the descendants of an individual reporting a low taste shock. Rewards can take two forms and society makes use of both. The first is standard and involves increased consumption spending, in present-value terms. The second is more subtle and exploits differences in preferences: it allows an adjustment in the pattern of consumption, for a given present value, in the direction preferred by individuals.[12]

Since individuals are more impatient than the planner, this latter form of reward is delivered by tilting the consumption profile toward the present. Similarly, punishments involve tilting the

[12]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.

consumption path toward the future. In both cases, earlier consumption dates are used more intensively to provide incentives—rewards and punishments are front-loaded. Indeed, consumption returns to a common steady-state level in the long-run regardless of the initial shock because affecting the consumption of very distant descendants is not an efficient way for society to provide incentives to the first generation.

Another Bellman Equation

Here we develop another Bellman equation that holds for any value of q not necessarily equal to $\hat{\beta}$. This alternative formulation is useful, both as a source of intuition and to motivate our focus on $q = \hat{\beta}$.

Consider the following cost minimization problem

$$K(v, \hat{v}) \equiv \min_{\{c_t\}} \sum_{t=0}^{\infty} q^t \sum_{\theta^t} c_t(\theta^t) \Pr(\theta^t),$$

subject to the incentive compatibility constraint $U(\{c_t\}, \sigma^*; \beta) \geq U(\{c_t\}, \sigma; \beta)$ and

$$\begin{aligned} \hat{v} &= \sum_{t=0}^{\infty} \hat{\beta}^t \, \mathbb{E}_{-1}\big[\theta_t u(c_t)\big] \\ v &= \sum_{t=0}^{\infty} \beta^t \, \mathbb{E}_{-1}\big[\theta_t u(c_t)\big], \end{aligned}$$

that is, delivering utility $\hat{v}$ and v for the planner and individual, respectively. Then the value function must satisfy the Bellman equation

$$K(v, \hat{v}) = \max_{u,w,\hat{w}} \mathbb{E}\big[\, c\big(u(\theta)\big) + qK\big(w\,(\theta)\,, \hat{w}\,(\theta)\,\big)\,\big],$$

subject to

$$\begin{aligned} \hat{v} &= \mathbb{E}\big[\theta u(\theta) + \hat{\beta}\hat{w}(\theta)\big] \\ v &= \mathbb{E}\big[\theta u(\theta) + \beta w(\theta)\big] \end{aligned}$$

and

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

This formulation could be used to derive all of our results, although the lower-dimensional Bellman equation (7) is slightly more convenient for that purpose. The advantage of this cost-minimization formulation, however, is that it lends itself naturally to economic interpretations.

The following story provides a useful reinterpretation and source for intuition. Consider an infinite-lived household with two members, husband and wife, and assume that consumption is a public good—there is no intra-period resource allocation problem. However, husband and wife disagree on how to discount the future. Suppose the wife is more patient, but only the husband can observe and report taste-shock realizations.

Then this cost-minimization problem characterizes the constrained Pareto problem for this household, in the sense that the isocost curve $K(v, \hat{v}) = K_0$ represents, given resources K_0, the Pareto frontier between husband and wife. The Pareto frontier is non-standard in that it is not everywhere decreasing and does not represent the usual transfer of private goods between two agents. Instead, it arises from differences in preferences that generate a disagreement about the optimal consumption path for the only public good. Since disagreement on preferences is bounded, the Pareto frontier is non-monotone and the highest possible utility for the wife is attained for an interior utility level for the husband, where $K_1(v^*, \hat{v}^*) = 0$. Reductions in the husband's utility to the left of this point must also decrease utility for the wife, for a given level of resources.

The first-order conditions can be rearranged to deliver

$$\hat{\beta} K_2(v, \hat{v}) = q K_2\big(w(\theta), \hat{w}(\theta)\big) \tag{16}$$

$$-\frac{K_1(v, \hat{v})}{K_2(v, \hat{v})} = \frac{\beta}{\hat{\beta}} \mathbb{E}\left[-\frac{K_1\big(w(\theta), \hat{w}(\theta)\big)}{K_2\big(w(\theta), \hat{w}(\theta)\big)}\right]. \tag{17}$$

Condition (16) can then be used to argue that a steady-state requires $q = \hat{\beta}$. Indeed, if $q < \hat{\beta}$, then $\{K_{2t}\}$ would increase without bound; likewise, if $q > \hat{\beta}$, then $\{K_{2t}\}$ decreases toward zero. Both situations clearly do not lend themselves to the existence of an invariant distribution for $(v, \hat{v})$. On the other hand, if $q = \hat{\beta}$ then $K_2(v_t, \hat{v}_t)$ is constant along the optimal path and an invariant distribution is possible.

When $q = \hat{\beta}$, the state $(v_t, \hat{v}_t)$ moves along a one-dimensional locus given by $K_2(v, \hat{v}) = K_2(v_0, \hat{v}_0)$. Intuitively, since no incentives are required for the wife, she is perfectly insured in the sense that the marginal cost of delivering welfare to her is held constant across time.

Figure 5 shows that the curve $K_2(v, \hat{v}) = K_2(v_0, \hat{v}_0)$ for continuation utilities cuts the isocost curves from below, and cuts $K_1(v, \hat{v}) = 0$ from above. Intuitively, incentives require foregoing perfect insurance for the husband and accepting fluctuations in v as rewards and punishments. Starting from $(v^*, \hat{v}^*)$, rewards can be delivered in two ways. The optimum makes use of both forms of rewards, explaining the shape of the schedule for continuation utilities.

The first form of rewarding involves increasing resources K, and can be seen as an upward movement along the diagonal $K_1(v, \hat{v}) = 0$. However, the husband is also rewarded by allowing an allocation of these resources that is more to his liking, which can be represented as lateral movements along the Pareto frontier, which at $(v^*, \hat{v}^*)$ is horizontally flat. The solution combines both forms of rewards and, as a result, $(v, \hat{v})$ travels along $K_2(v, \hat{v}) = K_2(v_0, \hat{v}_0)$ to the right of $K_1(v, \hat{v}) = 0$ and above the initial isocost curve. Note that punishments will push the agent on the upward sloping section of the Pareto frontier. Thus, ex-ante efficiency demands ex-post inefficiency.[13]

[13] Returning to the analogy in footnote 12: parents often complain that the punishments they choose to inflict on their children hurt them more than they do their kids.

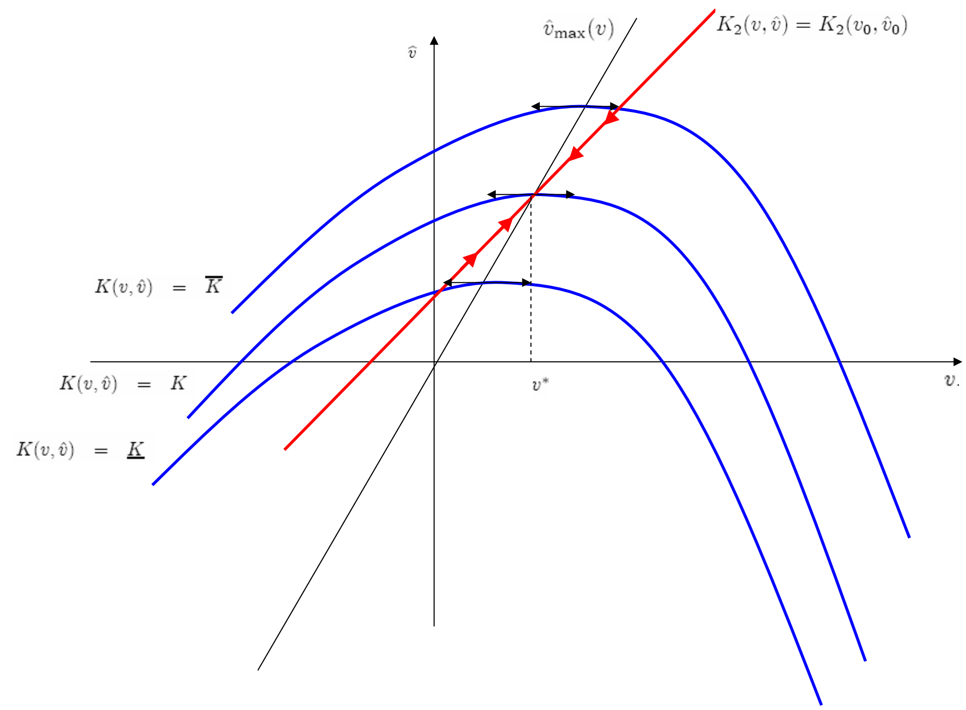

Figure 5: Isocost curves of $K(v, \hat{v})$

Condition (17) is the analog of the conditional-expectation equation (12) obtained from the one-dimensional Bellman equation. Here, it implies that the slope of the isocost curve (Pareto frontier), $-K_1/K_2$ reverts geometrically toward 0. Thus, $(v_t, \hat{v}_t)$ moves along $K_2(v, \hat{v}) = K_2(v_0, \hat{v}_0)$ and eventually reverts toward $(v^*, \hat{v}^*)$. Intuitively, the solution deviates from $(v^*, \hat{v}^*)$ to provide the husband with incentives, but it is efficient to revert back to this point of maximum efficiency for the wife: Patience ensures that the wife has her way in the long-run.

6 Estate Taxation

We now turn to a repeated Mirrleesian economy and study optimal taxation. In this version of our model, individuals have identical preferences over consumption and work effort but are heterogenous regarding their labor productivity, which is privately observed by the individual and independently distributed across generations and dynasties. We continue to focus on the case where the social welfare criterion discounts the future at a lower rate than individuals.

Unlike the taste-shock model, here even if we were to restrict allocations to feature no link between parent and child, there would still be a non-trivial planning problem. Indeed, in each period the situation would then be identical to the static, nonlinear income tax problem originally studied by Mirrlees (1971). Moreover, in the absence of altruism, so that $\beta = 0$, the social optimum actually coincides with this static solution. With altruism, however, we shall see below that it is always optimal to link welfare across generations within a dynasty to enhance incentives for parents.

Despite differences between the Mirrleesian economy and our taste-shock model, our previous analysis can be adapted virtually without change. In particular, a recursive representation can be derived, and the Bellman equation can be used to characterize the solution and to establish that a steady-state, invariant distribution exists. This highlights the fact that our model requires asymmetric information, but not any particular form of it.

We focus on an implementation of the allocation that uses income and estate taxes, and derive some interesting results for the latter. We find that estate taxation should be progressive: more fortunate parents should face a higher average marginal tax rate on their bequests. This result reflects the mean reversion in consumption explained in the previous section. A higher estate tax ensures that the fortunate face a lower net rate of return across generations, and that consequently their consumption path decreases over time toward the mean.

Repeated Mirrlees: Productivity Shocks

Each period of this economy is identical to the canonical optimal taxation setup in Mirrlees (1971). Utility depends on the level of consumption c and work effort n. We assume that individuals in generation t have identical preferences that satisfy

$$V_t = \mathbb{E}_{t-1}[u(c_t) - h(n_t) + \beta V_{t+1}],$$

but differ regarding their productivity in translating work effort into output. An individual with productivity w, exerting work effort n, produces output $y = wn$. We assume that productivity w is independently and identically distributed across dynasties and generations. Thus, the productivity talents of parent and child are unrelated—innate skills are assumed nonheritable. Given this assumption, if the optimum features intergenerational transmission of welfare, then it represents a social decision to provide altruistic parents with incentives in this way, and not a mechanical result originating from the assumed physical environment.[14]

For convenience, we adopt the power disutility function $h(n) = n^\gamma/\gamma$ so that, defining $\theta \equiv w^{-\gamma}$, we can write total utility over consumption and output as being subject to taste shocks

$$\sum_{t=0}^{\infty} \beta^t \, \mathbb{E}_{-1}\big[u(c_t) - \theta_t h(y_t)\big] = \sum_{t=1}^{\infty} \beta^t \, \mathbb{E}_{-1}\big[\beta^{-1} u(c_{t-1}) - \theta_t h(y_t)\big] - \mathbb{E}_{-1}\big[\theta_0 h(y_0)\big]$$

The right-hand side of this equation leads to a convenient recursive representation of the planning problem in the continuation utility defined by $v_t = \sum_{s=0}^{\infty} \beta^s \, \mathbb{E}_{t-1}[u(c_{t+s-1}) - \theta_{t+s} h(y_{t+s})]$ (where we are abusing notation slightly by folding the β^{-1} into the definition of the utility function $u(c)$).

The resource constraint requires total consumption not to exceed total output plus some fixed constant endowment

$$\int \sum_{\theta^t} c_t^v(\theta^t) \, \Pr(\theta^t) \, d\psi(v) \leq \int \sum_{\theta^t} y_t^v(\theta^t) \, \Pr(\theta^t) \, d\psi(v) + e \qquad t = 0, 1, \ldots$$

[14] As in the taste shock model, here the case with $\beta = \hat{\beta}$ leads to immiseration. This case has been studied by Albanesi and Sleet (2004), who impose an exogenous lower bound on dynastic welfare to circumvent immiseration.

where individuals are indexed by their initial utility entitlement v, with distribution ψ in the population.

We continue to assume social discounting is lower than private discounting: $\hat{\beta} > \beta$. The planning problem is to choose an allocation $\{c_t^v(\theta^t), y_t^v(\theta^t)\}$ to maximize average social welfare subject to the incentive-compatibility constraints and the resource constraints.

Using the last expression for the utility function and applying similar reasoning as in the taste-shock model yields the Bellman equation for the associated relaxed problem

$$k(v) = \max_{u_-,h,w} \mathbb{E}\big[u_- - \hat{\lambda}c(u_-) - \theta h(\theta) + \hat{\lambda}y\big(h(\theta)\big) + \hat{\beta}k\big(w(\theta)\big)\big]$$

$$v = \mathbb{E}\big[u_- - \theta h(\theta) + \beta w(\theta)\big]$$

$$-\theta h(\theta) + \beta w(\theta) \geq -\theta h(\theta') + \beta w(\theta'),$$

where the function $y(h)$ represents the inverse of the disutility function, $y = h^{-1}$. The arguments that justify the study of this Bellman equation, are similar to those that underlie Theorems 1 and 2 in the context of the taste-shock model. The results regarding steady states parallel those obtained previously, and imply that an invariant distribution exists with no immiseration, as in Proposition 2.

Implementation with Income and Estate Taxation

Any allocation that is incentive compatible and feasible can be implemented by a combination of taxes on labor income and estates. Here we first describe this implementation, and explore some features of the optimal estate tax in the next subsection.

For any incentive-compatible and feasible allocation $\{c_t^v(\theta^t),\, y_t^v(\theta^t)\}$ we propose an implementation along the lines of Kocherlakota (2004). In each period, conditional on the history of their dynasty's reports $\hat{\theta}^{t-1}$ and any inherited wealth, individuals report their current shock $\hat{\theta}_t$, produce, consume, pay taxes and bequeath wealth subject to the following set of budget constraints

$$c_t + b_t \leq y_t\big(\hat{\theta}^t\big) - T_t\big(\hat{\theta}^t\big) + \big(1 - \tau_t\big(\hat{\theta}^t\big)\big)R_{t-1,\,t}b_{t-1} \qquad t = 0, 1, \ldots \tag{18}$$

where $R_{t-1,t}$ is the before-tax interest rate across generations, and initially $b_{-1} = 0$. Individuals are subject to two forms of taxation: a labor income tax $T_t(\hat{\theta}^t)$, and a proportional tax on inherited wealth $R_{t-1,\,t}b_{t-1}$ at rate $\tau_t(\hat{\theta}^t)$.[15]

Given a tax policy $\{T_t^v(\theta^t),\, \tau_t^v(\theta^t),\, y_t^v(\theta^t)\}$, an equilibrium consists of a sequence of interest rates $\{R_{t,\,t+1}\}$; an allocation for consumption, labor income and bequests $\{c_t^v(\theta^t),\, b_t^v(\theta^t)\}$; and a reporting strategy $\{\sigma_t^v(\theta^t)\}$ such that: (i) $\{c_t,\, b_t,\, \sigma_t\}$ maximize dynastic utility subject to (18), taking the

[15]In this formulation, taxes are a function of the entire history of reports, and labor income y_t is mandated given this history. However, if the labor income histories $y^t\colon \Theta^t \to \mathbb{R}^t$ being implemented are invertible, then by the taxation principle we can rewrite T and τ as functions of this history of labor income and avoid having to mandate labor income. Under this arrangement, individuals do not make reports on their shocks, but instead simply choose a budget-feasible allocation of consumption and labor income, taking as given prices and the tax system.

sequence of interest rates $\{R_{t,t+1}\}$ and the tax policy $\{T_t, \tau_t, y_t\}$ as given; and (ii) the asset market clears so that $\int \mathbb{E}_{-1}[b_t^v(\theta^t)]\,d\phi(v) = 0$ for all $t = 0, 1, \ldots$ We say that a competitive equilibrium is incentive compatible if, in addition, it induces truth telling.

For any feasible, incentive-compatible allocation $\{c_t^v, y_t^v\}$, we construct an incentive-compatible competitive equilibrium with no bequests by setting $T_t^v(\theta^t) = y_t(\theta^t) - c_t(\theta^t)$ and

$$\tau_t^v(\theta^t) = 1 - \frac{1}{\beta R_{t-1,t}} \frac{u'(c_{t-1}^v(\theta^{t-1}))}{u'(c_t^v(\theta^t))} \tag{19}$$

for any sequence of interest rates $\{R_{t-1,t}\}$. These choices work because the estate tax ensures that for any reporting strategy σ, the resulting consumption allocation $\{c_t^v(\sigma^t(\theta^t))\}$ with no bequests $b_t^v(\theta^t) = 0$ satisfies the consumption Euler equation

$$u'\big(c_t^v\big(\sigma^t(\theta^t)\big)\big) = \beta R_{t,t+1} \sum_{\theta_{t+1}} u'\big(c_{t+1}^v\big(\sigma^{t+1}(\theta^t, \theta_{t+1})\big)\big)\big(1 - \tau_{t+1}^v\big(\sigma^{t+1}(\theta^t, \theta_{t+1})\big)\big) \Pr(\theta_{t+1}).$$

The labor income tax is such that the budget constraints are satisfied with this consumption allocation and no bequests. Thus, this no-bequest choice is optimal for the individual regardless of the reporting strategy followed. Since the resulting allocation is incentive compatible, by hypothesis, it follows that truth telling is optimal. The resource constraints together with the budget constraints then ensure that the asset market clears.[16]

As noted above, in our economy without capital only the after-tax interest rate matters so the implementation allows any equilibrium before-tax interest rate $\{R_{t-1,t}\}$. In the next subsection, we set the interest rate to the reciprocal of the social discount factor, $R_{t-1,t} = \hat{\beta}^{-1}$. This choice is natural because it represents the interest rate that would prevail at the steady state in a version of our economy with capital.

Optimal Progressive Estate Taxation

In this subsection we derive an important intertemporal condition that must be satisfied by the optimal allocation. This condition has interesting implications for the optimal estate tax, computed using (19) at the optimal allocation.

Let λ be the multiplier on the promise-keeping constraint and let $\mu(\theta, \theta')$ represent the multipliers on the incentive constraints. Then the first-order conditions for u_- and $w(\theta)$ are

$$\begin{aligned} c'(u_-) - \hat{\lambda} - \lambda &= 0 \\ \hat{\beta} k'\big(w(\theta)\big)p(\theta) - \beta\lambda p(\theta) - \sum_{\theta'} \mu(\theta, \theta') + \sum_{\theta'} \mu(\theta', \theta) &= 0 \end{aligned}$$

[16]Since the consumption Euler equation holds with equality, the same estate tax can be used to implement allocations with any other bequest plan with income taxes that are consistent with the budget constraints.

and the envelope condition is $k'(v) = \lambda$. Together these imply

$$\sum_{\theta} k'\big(w(\theta)\big)p(\theta) = \frac{\beta}{\hat{\beta}}\, k'(v),$$

Using $c'(u_{-}) = 1/u'(c_{-}) = \hat{\lambda} + k'(v)$ we arrive at the *Modified Inverse Euler equation*

$$\frac{1}{u'(c_{-})} \;=\; \frac{\hat{\beta}}{\beta} \sum_{\theta} \frac{1}{u'\big(c(\theta)\big)}\; p(\theta) \;-\; \hat{\lambda}\left(\frac{\hat{\beta}}{\beta} - 1\right). \tag{20}$$

The left-hand side together with the first term on the right-hand side is the standard inverse Euler equation. The second term on the right-hand side is novel, since it is zero when $\beta = \hat{\beta}$ and is strictly negative when $\hat{\beta} > \beta$.[17]

In our environment, the relevant past history is encoded in the continuation utility so the estate tax $\tau(\theta^{t-1}, \theta_t)$ can actually be reexpressed as a function of $v_t(\theta^{t-1})$ and θ_t. Abusing notation we then denote the estate tax by $\tau_t(v, \theta_t)$. Since we focus on the steady-state, invariant distribution, we also drop the time subscripts and write $\tau(v, \theta)$.

The average estate tax rate $\bar{\tau}(v)$ is then defined by

$$1 - \bar{\tau}(v) \equiv \sum_{\theta} \big(1 - \tau(v, \theta)\big)p(\theta) \tag{21}$$

Using the modified inverse Euler equation (20) we obtain

$$\bar{\tau}(v) \;=\; -\hat{\lambda}\, u'\big(c_{-}(v)\big)\left(\frac{\hat{\beta}}{\beta} - 1\right)$$

In particular, this implies that the average estate tax rate is negative, $\bar{\tau}(v) < 0$, so that bequests are subsidized. However, recall that before-tax interest rates are not uniquely determined in our implementation. As a consequence, neither are the estate taxes computed by (19). With our particular choice for the before-tax interest rate, however, the tax rates are pinned down and acquires a corrective, Pigouvian role. Differences in discounting can be interpreted as a form of externalities from future consumption, and the negative average tax can then be seen as a way of countering these externalities as prescribed by Pigou. In our setup without capital, this result depends on the choice of the before-tax interest rate. However, the negative tax on estates would be a robust steady-state outcome in a version of our economy with capital.

In our model it is more interesting to understand how the average tax varies with the history of past shocks encoded in the promised continuation utility v. The average tax is an increasing function of consumption, which, in turn, is an increasing function of v. Thus, estate taxation is progressive: the average tax on transfers for more fortunate parents is higher.

[17]This equation can also be derived from an elementary variation argument. This is done in Farhi, Kocherlakota and Werning (2005), who also show that this equation, and its implications for estate taxation, generalize to an economy with capital and an arbitrary process for skills.

Proposition 4 *In the repeated Mirrlees economy, the optimal allocation can be implemented by a combination of income and estate taxes. At a steady-state, invariant distribution ψ^*, the optimal average estate tax $\bar{\tau}(v)$ defined by (19) and (21) is increasing in promised continuation utility v.*

The progressivity of the estate tax reflects the mean-reversion in consumption. The fortunate must face lower net rates of return so that their consumption path decreases towards the mean.[18]

7 Conclusions

Should privately-felt parental altruism affect the social contract? If so, what are the long-run implications for inequality? To address these questions, we modeled a central tension in society: the tradeoff between ensuring equality of opportunity for newborns and providing incentives for altruistic parents.

Our model's answer is that society should indeed exploit altruism to motivate parents, linking the welfare of children to that of their parents. However, we also find that if we value the welfare of future generations directly, the inheritability of good or bad fortune should be tempered. This produces a steady-state outcome in which welfare and consumption are mean-reverting, long-run inequality is bounded, social mobility is possible and misery is avoided by everyone.

What instruments should society use to implement such allocations? For a Mirrleesian version of our model we find an important role for the estate tax. The optimal tax on inheritances is progressive: more fortunate parents should face a higher average marginal tax rate on their bequests. This result illustrates an interesting way in which the conflict between corrective and redistributive taxation is optimally resolved. Further examination of other situations with similar conflicts remains an interesting direction for future work.[19]

Appendix

Proof of Theorem 1

Weak concavity of the value function $k(v)$ follows because the relaxed 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

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

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

[18] Farhi, Kocherlakota and Werning (2005) explore more general versions of this result and discuss other intuitions.
[19] Some progress along these lines can be found in Amador, Angeletos and Werning (2005).

continuous.

The constraint (6) with $q = \hat{\beta}$ implies that utility and continuation utility 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 $\lim_{T\to\infty} \sum_{t=0}^{T} \hat{\beta}^t \, \mathbb{E}_{-1} c_t = \infty$. 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 (6). Thus, for both the relaxed and unrelaxed problems utility and continuation utility 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:

$$\sup_{u,w} \mathbb{E}\big[\theta u(\theta) - \hat{\lambda} c\big(u(\theta)\big) + \hat{\beta} k\big(w(\theta)\big)\big] \tag{22}$$

$$v = \mathbb{E}\big[\theta u(\theta) + \beta w(\theta)\big] \tag{23}$$

$$\theta u\left(\theta\right) + \beta w(\theta) \geq \theta u\left(\theta'\right) + \beta w(\theta') \text{ for all } \theta, \theta' \in \Theta \tag{24}$$

Define $m \equiv \max_{c\geq 0, \theta\in\Theta}(\theta u(c) - \hat{\lambda} c)$ and $\hat{k}\left(v\right) \equiv k\left(v\right) - m/(1-\hat{\beta}) \leq 0$. The problem in (22) is equivalent to the following optimization with non-positive objective:

$$\sup_{u,w} \mathbb{E}\big[\theta u(\theta) - \hat{\lambda} c\big(u(\theta)\big) - m + \hat{\beta} \hat{k}\big(w(\theta)\big)\big] \tag{25}$$

subject to (23) and (24).

Lemma A.1 *The supremum in (22), or equivalently (25), 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

$$\lim_{v\to\infty} \hat{k}(v) = \lim_{v\to-\infty} \hat{k}(v) = -\infty \tag{26}$$

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

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

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 (26) follow. Since $\theta u - \hat{\lambda} c(u) - m \leq 0$ and $\beta < \hat{\beta}$ it follows that

$$h(v,\hat{\beta}) \leq h(v,\beta) = v - \hat{\lambda} C(v,\beta) - \frac{m}{1-\beta}, \tag{27}$$

where

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

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

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

so that $\lim_{v\to-\infty} h(v,\beta) = -\infty$ and $\lim_{v\to\infty} h(v,\beta) = -\infty$. Using the inequality (27) 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 (26).

Fix a v. Take any allocation that verifies the constraints (23) and (24) and let $\bar{k} < \infty$ be the corresponding value of (25). 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 (26), 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 $\theta u(\theta) - \hat{\lambda} c(u(\theta)) - m \geq \bar{k}/p(\theta)$. Since $(\theta u - \hat{\lambda} c(u))$ is strictly concave, with $(\theta u - \hat{\lambda} c(u)) \to -\infty$ 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 (25) 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 (7)–(9).*

Proof. Suppose that for some v

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

where the maximization is subject to (23) and (24). Then there exists $\Delta > 0$ such that

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

for all (u, w) that satisfy (23) and (24). But then by definition

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

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

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

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 $k(v) \leq \max_{u,w} \mathbb{E}[\theta u(\theta) - \hat{\lambda} c(u(\theta)) + \hat{\beta} k(w(\theta))]$ subject to (23) and (24).

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

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

Let $(u^*(\theta), w^*(\theta)) \in \arg\max_{u,w} \mathbb{E}[\theta u(\theta) - \hat{\lambda} c(u(\theta)) + \hat{\beta} k(w(\theta))]$. 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

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

Since $\varepsilon > 0$ was arbitrary it follows that $k(v) \geq \max_{u,w} \mathbb{E}[\theta u(\theta) - \hat{\lambda} c(u(\theta)) + \hat{\beta} k(w(\theta))]$ subject to (23) and (24).

Finally, together both inequalities imply $k(v) = \max_{u,w} \mathbb{E}[\theta u(\theta) - \hat{\lambda} c(u(\theta)) + \hat{\beta} k(w(\theta))]$ subject to (23) and (24). ■

Proof of Theorem 2

Part (a). Suppose the allocation $\{u_t\}$ is generated by the policy functions starting from v_0, is incentive compatible and delivers lifetime utility v_0. After repeated substitutions of the Bellman equation (7), we arrive at

$$k(v_0) = \sum_{t=0}^{T} \hat{\beta}^t \, \mathbb{E}_{-1}[\theta_t u_t(\theta^t) - \hat{\lambda} c(u_t(\theta^t))] + \hat{\beta}^T \, \mathbb{E}_{-1} k(v_T(\theta^T)). \tag{28}$$

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

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

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 utility v_0. Define the continuation utility 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 (23) and (24). The Bellman equation then implies that

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

The first inequality follows since u_0 does not maximize (7), 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 $\lim_{t\to\infty} \beta^t \, \mathbb{E}_{-1} v_t(\theta^t) = 0$. Then, after repeated substitutions of (8), we obtain

$$v = \sum_{t=0}^{T} \beta^t \, \mathbb{E}_{-1}\big[\theta_t u_t(\theta^t)\big] + \beta^T \, \mathbb{E}_{-1}\big[v_T(\theta^T)\big]. \tag{29}$$

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 lifetime utility v_0. Next, we show that for any allocation generated by (g^u, g^w), starting from finite v_0, we have $\lim_{t\to\infty} \beta^t \, \mathbb{E}_{-1} v_t(\theta^t) = 0$.

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

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

and then (28) implies that $k(v_0) = -\infty$, a contradiction since there are feasible plans that yield finite values. We conclude that $\limsup_{t\to\infty} \beta^t \, \mathbb{E}_{-1} v_t(\theta^t) \leq 0$.

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

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

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

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

Part (c). Suppose $\limsup_{t\to\infty} \beta^t \, \mathbb{E}_{-1} v_t(\sigma^t(\theta^t)) \geq 0$ for every reporting strategy σ. Then after repeated substitutions of (9),

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

implying

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

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

Proof of Lemma 1

Part (a) (Strict Concavity) Let $\{u_t(\theta^t, v_0), v_t(\theta^t, v_0)\}$ be the plans generated from the policy functions starting at v_0 (note: no claim of incentive compatibility is required). Take two initial utility values v_a and v_b, with $v_a \neq v_b$. Define the average utilities

$$\begin{aligned} u_t^{\alpha}(\theta^t) &\equiv \alpha u_t(\theta^t, v_a) + (1-\alpha) u_t(\theta^t, v_b) \\ v_t^{\alpha}(\theta^t) &\equiv \alpha v_t(\theta^t; v_a) + (1-\alpha) v_t(\theta^t; v_b) \end{aligned}$$

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 utility $v^{\alpha} \equiv \alpha v_a + (1-\alpha) v_b$. It also implies that there exists a finite time T such that

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

so that

$$u_t\left(\theta^t; v_a\right) \neq u_t\left(\theta^t; v_b\right), \tag{30}$$

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

$$\begin{aligned} k(v_a) &= \sum_{t=0}^{T} \hat{\beta}^t \, \mathbb{E}_{-1}\big[\theta_t u_t(\theta^t; v_a) - c\big(u_t(\theta^t; v_a)\big)\big] + \hat{\beta}^T \, \mathbb{E}_{-1} k\big(v_T(\theta^T; v_a)\big) \\ k(v_b) &= \sum_{t=0}^{T} \hat{\beta}^t \, \mathbb{E}_{-1}\big[\theta_t u_t(\theta^t; v_b) - c\big(u_t(\theta^t; v_b)\big)\big] + \hat{\beta}^T \, \mathbb{E}_{-1} k\big(v_T(\theta^T; v_b)\big), \end{aligned}$$

and averaging we obtain

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

where the strict inequality follows from the strict concavity of the cost function $c(u)$, the fact that we have the inequality (30), 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 Bellman equations constraints 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—that is, there is at least one sub-gradient 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 utility. For a neighborhood around v_0 define the test function

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

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, Lucas and Prescott, 1989), that $k'(v_0)$ also exists and

$$k'(v_0) = W'(v_0) = 1 - \hat{\lambda} \, \mathbb{E}\big[c'\big(u^*(\theta)\big)\big]. \tag{31}$$

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

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

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

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

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.

Proof of Proposition 1

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

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

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

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

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. We study all of them. The optimal allocation of our problem must solve one of these programs, although not necessarily the one that yields the highest value, since this one may not be feasible if the monotonicity condition is violated.

The first-order conditions are

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

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

where, by the Envelope theorem, $\lambda = k'(v)$.

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

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

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

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

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

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

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

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

we get

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

For any n, $w_K \leq w_n \leq w_1$, we have for every n

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

After rearranging, we obtain

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

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 $1 - k'(v) \geq 0$.

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 $u(\theta) = \underline{u}$ and $w(\theta) = \beta^{-1}v > v$. 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 $u(\theta) > \underline{u}$ for all θ, which must be true for high enough v, i.e. for low enough $k'(v)$.

Proof of Proposition 2

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

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

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 each $x < 1$ there exists a subsequence $\{T_{Q,\phi(n)}(\delta_x)\}$ that converges weakly, i.e. in distribution, to an invariant distribution on $(-\infty, 1)$ under Q.*

Proof. The bounds (13) derived in Proposition 1 imply that for all $\theta \in \Theta$

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

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 $T_{\hat{Q}}^n(\delta_x)$ is simply $\mathbb{E}_{-1}[k'(v_t(\theta^{t-1}))]$ with $x = k'(v_0) < 1$. Recall that $\mathbb{E}_{-1}[k'(v_t(\theta^{t-1}))\big] = (\beta/\hat{\beta})^t k'(v_0) \to 0$. This implies that

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

for all $A > 0$. Rearranging,

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

which implies that $\{T_{\hat{Q}}^n(\delta_x)\}$, 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

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

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 on $(-\infty, 1)$ under Q. ■

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 $v_H < \infty$ for the ergodic set for v. Define the utility 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 $Q(x,\theta) \le k'(v_L) < \infty$. 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).

Proof of Proposition 3

Consider indexing the relaxed planning problem by e by setting $\hat{\lambda} = e^{-1}$, 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 problem and original problem 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 $k'(v_t; e) = \mathbb{E}_{t-1}\left[1 - \hat{\lambda} c'\left(u_t^v(\theta^t)\right)\right]$. Applying the law of iterated expectations to (12) then yields

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

With logarithmic utility $c'(u) = c(u)$, so that $\int k'(v; e)\, d\psi(v) = 0$ implies $\int \mathbb{E}_{-1}[c_t]\, d\psi = \hat{\lambda}^{-1} = e$ for all $t = 0, 1, \ldots$ The allocation is incentive compatible by Lemma 2 below, 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 $\hat{\lambda} = e^{-1}$ 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 e^* for any initial distribution ψ. The monotonicity of $e^*(\psi)$ then follows immediately by using the fact that $k'(\cdot; 1)$ is a strictly decreasing function.

Proof of Lemma 2

(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 1 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 1, 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 (13) implies that $k'\left(g^w\left(\bar{\theta},v\right)\right) \geq 1 - \beta/\hat{\beta}$ 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 Lemma 3

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

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

Applying Jensen's inequality we obtain

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

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

References

[1] **Aiyagari, Rao, and Fernando Alvarez (1995)**, "Stationary Efficient Distributions with Private Information: A Tale of Kings and Slaves,"mimeo.

[2] **Albanesi, Stefania, and Christoper Sleet (2004)**, "Dynamic Optimal Taxation with Private Information,"mimeo.

[3] **Amador, Manuel, and George-Marios Angletos, and Iván Werning (2005)**, "Corrective vs. Redistributive Taxation,"work in progress, mimeo.

[4] **Atkeson, Andrew, and Robert E. Lucas Jr. (1992)** "On Efficient Distribution with Private Information,"*Review of Economic Studies,* v59, n3 (July): 427–53.

[5] **Bernheim, B. Douglas (1989)** "Intergenerational Altruism, Dynastic Equilibria and Social Welfare,"*Review of Economic Studies,* v56, n1 (Jan), pp. 119–128.

[6] **Caplin, Andrew, and John Leahy (2004)** "The Social Discount Rate,"*Journal of Political Economy*, v112, n6 (December): 1257–68.

[7] **Farhi, Emmanuel, Narayana Kocherlakota and Iván Werning (2005)** work in progress.

[8] **Green, Edward J. (1987)** "Lending and the Smoothing of Uninsurable Income,"in *Contractual Arrangements for Intertemporal Trade*: 3–25.

[9] **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.

[10] **Golosov, Mikhail, Narayana Kocherlakota, and Aleh Tsyvinski (2003)** "Optimal Indirect and Capital Taxation,"*Review of Economic Studies*, v70, n3 (July): 569–87.

[11] **Kaplow, Louis (1995)** "A Note on Subsidizing Gifts,"*Journal of Public Economics,* v58, n3 (Nov): 469–77.

[12] **Kocherlakota, Narayana (2004)** "Zero Expected Wealth Taxes: A Mirrlees Approach to Dynamic Optimal Taxation,"mimeo.

[13] **Mirrlees, James (1971)** "An Exploration in the Theory of Optimum Income Taxation,"*Review of Economic Studies,* v38, n2: 175–208.

[14] **Phelan, Christopher (1998)** "On the Long Run Implications of Repeated Moral Hazard,"*Journal of Economic Theory,* v79, n2 (April 1998): 174–91.

[15] **Phelan, Christopher (2005)** "Opportunity and Social Mobility,"Federal Reserve Bank of Minneapolis, Research Department Staff Report 323 (March).

[16] **Sleet, Christopher, and Sevin Yeltekin (2004)** "Credible Social Insurance,"mimeo.

[17] **Spear, Stephen E., and Sanjay Srivastava (1987)** "On Repeated Moral Hazard with Discounting,"*Review of Economic Studies,* v54, n4 (October): 599–617.

[18] **Stokey, Nancy L., and Robert E. Lucas, Jr., with Edward C. Prescott (1989)** "Recursive Methods in Economic Dynamics,"Harvard Press.

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