Inequality and Social Discounting

Emmanuel Farhi

Harvard University

Iván Werning

Massachusetts Institute of Technology

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

I. Introduction

Societies inevitably *choose* the inheritability of welfare. Some balance between equality of opportunity for newborns and incentives for altru-

istic parents is struck. In this paper, we explore how this balancing act plays out to determine long-run inequality.

An important backdrop to this question is provided by Atkeson and Lucas (1992). They study a model populated by infinitely lived agents subject to idiosyncratic shocks that are private information. They reach an extreme conclusion by proving an *immiseration* result: consumption and welfare inequality should be perfectly inheritable and rise steadily without bound, with everyone converging to absolute misery and a vanishing lucky fraction to bliss.[1] We depart minimally from this framework by adopting the same positive economic model, but using a slightly different normative criterion. In a generational context, efficient allocations for infinitely lived agents characterize the instance in which future generations are not considered *directly*, but only *indirectly* through the altruism of earlier ones. On the opposite side of the spectrum, Phelan (2006) proposes a social planner with equal weights on all generations that avoids the immiseration result because any allocation that leads everyone to misery actually minimizes the welfare criterion. Our interest here is in exploring a large class of Pareto-efficient allocations that also value future generations, but not equally. We place a positive and vanishing Pareto weight on the expected welfare of future generations, which allows us to remain arbitrarily close to Atkeson-Lucas. As this weight varies, we trace the Pareto frontier between Atkeson-Lucas and Phelan.[2]

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

Our positive economy is identical to Atkeson-Lucas's taste-shock setup.

---

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

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

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

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

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

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

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

A methodological contribution of this paper is to reformulate the social planning problem recursively in a way that extends the ideas introduced by Spear and Srivastava (1987) to a general-equilibrium sit-

uation in which private and social objectives potentially differ. We are able to reduce the dynamic program to a one-dimensional state variable, and our analysis and results heavily exploit the resulting Bellman equation.

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

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

II.   Social Discounting and Mean Reversion

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

discussion also provides a novel intuition for the immiseration result in Atkeson-Lucas.

Social Discounting

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

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

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

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

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

A Planning Problem

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

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

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

subject to

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

The first-order conditions for an interior optimum are

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

where λ is the Lagrange multipliers on the constraint.

Imperfect Inheritability

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

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

In this simple deterministic example, initial inequality Δ was taken as exogenously given. However, in the model with taste shocks, inequality is continuously generated in order to provide incentives. The dynamics after a shock are similar to those illustrated here, so that figure 1 can be loosely interpreted as an impulse response function. If $\hat{\beta} = \beta$, shocks have a permanent effect on inequality and consumption inherits a random-walk component. If $\hat{\beta} > \beta$, the impact of shocks decays over time and consumption is mean-reverting.

As long as $\hat{\beta} > \beta$, inequality vanishes in the long run in this deterministic example. With ongoing taste shocks, inequality remains positive in the long run and the mean-reverting force ensures that inequality remains bounded. By contrast, when $\hat{\beta} = \beta$ as in Atkeson and Lucas (1992), there is no mean reversion, so that shocks accumulate indefinitely and inequality increases without bound.
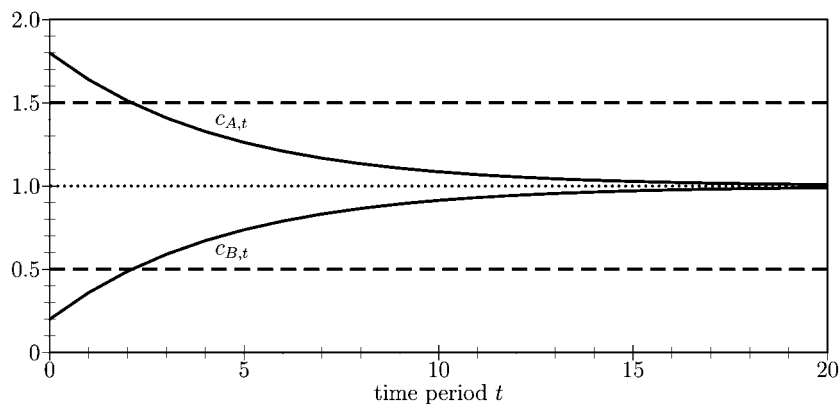
Fig. 1.—Consumption paths for groups A and B. Solid lines represent the case with $\hat{\beta} > \beta$; the dotted line at $c = e = 1$ is the steady state. The horizontal dashed lines represent the Atkeson-Lucas case with $\hat{\beta} = \beta$.

III. An Intergenerational Insurance Problem

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

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

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

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

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

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

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

The following notation and conventions will be used. We refer to $U(c_t)$ as utility and the discounted, expected utility v_t as welfare. Let $p(\theta)$ denote the probability of $\theta \in \Theta$; we adopt the normalization that

$\mathbb{E}[\theta] = \sum_{\theta \in \Theta} \theta p(\theta) = 1$ and define $\overline{\theta} = \max \Theta$ and $\underline{\theta} = \min \Theta$. The cost function $C(u)$ is defined as the inverse of the utility function: $C \equiv U^{-1}$. Dynastic welfare v_t belongs to the set $U(\mathbb{R}_+)/(1 - \beta)$ with extremes $\underline{v} \equiv U(0)/(1 - \beta)$ and $\bar{v} \equiv \lim_{c \to \infty} U(c)/(1 - \beta)$, which may be finite or infinite.

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

We identify each dynasty with its founder's welfare entitlement v. We assume, without loss of generality, that all dynasties with the same entitlement v receive the same treatment. We then let ψ denote a cumulative distribution function for v across dynasties. An allocation is a sequence of functions $\{u_t^v\}$ for each v. For any given initial distribution of entitlements ψ and resources e, we say that an allocation $\{u_t^v\}$ is *feasible* if (i) it delivers expected utility of v to all initial dynasties entitled to v:

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

(ii) it is incentive compatible for all v:

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

whenever this sum converges; and (iii) total consumption does not exceed the fixed endowment e in all periods:

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

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

Social Discounting

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

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

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

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

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

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

so that social preferences are more patient. Future generations are already indirectly valued through the altruism of the current generation. If, in addition, they are also *directly* included in the welfare function, the social discount factor must be higher than β.[3]

In particular, weighing future generations $t = 1, 2, \ldots$ with geometric Pareto weights $\alpha_t = \hat{\beta}^t$ gives

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

for $\hat{\beta} > \beta$. The first term is identical to the expression in (5); the second is a constant when initial welfare promises v_0 for the founding generation are given, as they are in the social planning problem defined below.

Planning Problem

Define the social optimum as a feasible allocation that maximizes the integral of (5) with respect to distribution ψ. That is, the *social planning*

[3] Bernheim (1989) performs similar intergenerational discount factor calculations in his welfare analysis of a deterministic dynastic saving model. Caplin and Leahy (2004) argue that these ideas also apply to intrapersonal 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.

problem given an initial distribution of welfare entitlements ψ and an endowment level e is

$$S(\psi; \ e) \equiv \sup_{\{u_t^v\}} \int \ \sum_{t=0}^{\infty} \sum_{\theta^t \in \Theta^{t+1}} \hat{\beta}^t \theta_t u_t^v(\theta^t) \Pr(\theta^t) d\psi(v) \tag{8}$$

subject to (2), (3), and (4). The constraint set is nonempty as long as $e \geq e^*(\psi)$. If $e = e^*(\psi)$, then the only feasible allocation is the one characterized by Atkeson-Lucas; we study cases with $e > e^*(\psi)$. This problem is convex: the objective is linear, constraints (2) and (3) are linear, and the resource constraints (4) are strictly convex.

The way we have defined the social planning problem imposes that initial welfare entitlements v be delivered exactly, in the sense that the promise-keeping constraints (2) are equalities instead of inequalities. Alternatively, suppose that the founder of each dynasty is indexed by some minimum welfare entitlement $\tilde{v}$, with distribution $\tilde{\psi}$. The *Pareto problem* maximizes the integral of the welfare criterion (7) subject to delivering $\tilde{v}$ or more to the founders and incentive compatibility. The two problems are related: the solution to the Pareto problem solves the social planning problem for some distribution ψ that first-order stochastically dominates $\tilde{\psi}$ (so that $\psi(v) \leq \tilde{\psi}(v)$ for all v). In particular, comparing the terms in (7) with (5) implies that the Pareto problem chooses ψ to maximize

$$S(\psi; \ e) - \int \ v d\psi(v) \tag{9}$$

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

Steady States

The social planning problem takes the initial distribution of welfare entitlements ψ as given. In later periods the current cross-sectional distribution of continuation welfare ψ_t is a sufficient statistic for the remaining social planning problem: the problem is recursive with state variable ψ_t. It follows that the solution to the social planning problem from any period t onward, $\{u_{t+s}^v\}_{s=0}^{\infty}$, is a time-independent function of

the current distribution ψ_t, which evolves according to a stationary recursion $\psi_{t+1} = \Psi\psi_t$, for some fixed mapping Ψ.

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

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

Using the entire distribution ψ_t as a state variable is one way to approach the social planning problem. Indeed, this is the method adopted by Atkeson and Lucas (1992). They were able to keep the analysis manageable, despite the large dimensionality of the state variable, by exploiting the homogeneity of the problem with constant relative risk aversion (CRRA) preferences. In contrast, even in the CRRA case, our model lacks this homogeneity, making such a direct approach intractable.[4] Consequently, in the next section, we attack the problem differently, using a dynamic program with a one-dimensional state variable.[5] The idea is that the continuation welfare v_t of each dynasty follows a Markov process and that steady states are invariant distributions of this process.

IV. A Bellman Equation

In this section we approach the social planning problem by studying a relaxed version of it, whose solution coincides with that of the original problem at steady states. The relaxed problem has two important advantages. First, it can be solved by studying a set of subproblems, one for each dynasty, thereby avoiding the need to keep track of the entire distribution ψ_t in the population. Second, each subproblem admits a

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

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

one-dimensional recursive formulation, which we are able to characterize quite sharply. We believe that the general route we develop here may be useful in other contexts.

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

$$\int \sum_{t=0}^{\infty} Q_t \sum_{\theta^t \in \Theta^{t+1}} C(u_t^v(\theta^t)) \Pr(\theta^t) d\psi(v) \le e \sum_{t=0}^{\infty} Q_t \qquad (10)$$

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 $\{Q_t\}$.[6] The original and relaxed versions of the social planning problem are related in that any solution to the latter that happens to satisfy the resource constraints (4) is also a solution to the former.[7] It follows that any steady-state solution to the relaxed problem is a steady-state solution to the original one, since at a steady state the intertemporal constraint (10) implies the resource constraints (4). A steady state requires $Q_t = \hat{\beta}^t$.

Consider then the intertemporal resource constraint (10) with $Q_t = \hat{\beta}^t$:

$$\int \sum_{t=0}^{\infty} \hat{\beta}^t \sum_{\theta^t \in \Theta^{t+1}} C(u_t^v(\theta^t)) \Pr(\theta^t) d\psi(v) \le e \sum_{t=0}^{\infty} \hat{\beta}^t. \qquad (11)$$

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

$$\mathcal{L} \equiv \int \mathcal{L}^v d\psi(v), \qquad (12)$$

where

$$\mathcal{L}^v \equiv \sum_{t=0}^{\infty} \sum_{\theta^t \in \Theta^{t+1}} \hat{\beta}^t [\theta_t u_t^v(\theta^t) - \eta C(u_t^v(\theta^t))] \Pr(\theta^t).$$

---

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

[7] Since the problem is convex, a Lagrangian argument establishes the converse: there must exist some positive sequence $\{Q_t\}$ such that the solution to the original social planning problem also solves the relaxed problem. This is analogous to the second theorem of welfare economics for our environment. However, we will not require this converse result to construct a steady-state solution.

We study the maximization of $\mathcal{L}$ subject to (2) and (3). Maximizing $\mathcal{L}$ is equivalent to the pointwise optimization of $\mathcal{L}^v$ for each v:

$$k(v) \equiv \sup_{\{u_t^v\}} \mathcal{L}^v \tag{13}$$

subject to (2) and (3). We call this subproblem, for a given v and η, the *component planning problem*. Its connection with the relaxed problem is that for any e there exists a positive multiplier η such that an allocation $\{u_t^v\}$ solves the relaxed planning problem with endowment e if and only if for each v the allocation $\{u_t^v\}$ solves the component planning problem given v and η (Luenberger 1969, chap. 8).

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

THEOREM 1. The value function of the component planning problem $k(v)$ defined by equation (13) is continuous, is concave, and satisfies the Bellman equation

$$k(v) = \max_{u,w} \mathbb{E}[\theta u(\theta) - \eta C(u(\theta)) + \hat{\beta}k(w(\theta))] \tag{14}$$

subject to

$$v = \mathbb{E}[\theta u(\theta) + \beta w(\theta)] \tag{15}$$

and

$$\theta u(\theta) + \beta w(\theta) \geq \theta u(\theta') + \beta w(\theta') \quad \text{for all } \theta, \theta' \in \Theta. \tag{16}$$

This recursive formulation imposes a promise-keeping constraint (15) and an incentive constraint (16). 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, which is necessary for full incentive compatibility (3). Full incentive compatibility (3) is taken care of in (14) by evaluating the value function defined from the sequence problem at the continuation welfare: in the next period, envision the planner as solving the remaining sequence problem, selecting an allocation for each $w(\theta)$ that is incentive compatible for $t = 1, 2, \ldots$. Then any pair $(u(\theta), w(\theta))$ that satisfies (15) and (16), pasted with the corresponding continuation allocations for each $w(\theta)$, describes an allocation that satisfies (2) and (3).

Among other things, theorem 1 shows that the maximization on the right-hand side of the Bellman equation is uniquely attained by some continuous policy functions $g^u(\theta, v)$ and $g^w(\theta, v)$ for u and w, respectively. We emphasize that these policy functions solve the maximization in the Bellman equation (14) using the value function $k(v)$ defined from the

sequence problem (13).[8] For any initial welfare entitlement v, an allocation $\{u_t\}$ can then be generated from the policy functions (g^u, g^w) by $u_t(\theta^t) = g^u(\theta_t, v_t(\theta^{t-1}))$, with $v_0 = v$ and $v_{t+1}(\theta^t) = g^w(\theta_t, v_t(\theta^{t-1}))$. Our next result provides a connection between allocations generated this way and solutions to the component planning problem (13).

THEOREM 2.   For any (v, η), if an allocation $\{u_t\}$ attains the maximum in the component planning problem (13), then it is generated by (g^u, g^w). Conversely, if an allocation $\{u_t\}$ generated by (g^u, g^w) is such that for each σ

$$\limsup_{t \to \infty} \mathbb{E}_{-1} \beta^t v_t(\sigma^{t-1}(\theta^{t-1})) \geq 0, \tag{17}$$

then it attains the maximum in the component planning problem (13).

The first part of theorem 2 implies either that the solution to the relaxed planning problem is generated by the policy functions or that there is no solution at all. From the second part of theorem 2, a solution is guaranteed if the limiting condition (17) can be verified. The proof proceeds by showing that the allocation generated by the policy functions is optimal if it satisfies the incentive compatibility constraint (3); the role of the limiting condition (17) is to ensure the latter.[9] Conditions (15) and (16) ensure that finite deviations from truth-telling are not optimal; condition (17) then rules out infinite deviations. Condition (17) is trivially satisfied for utility functions that are bounded below; proposition 5 below verifies this condition when utility is unbounded below.

V.   Mean Reversion

In this section we use the Bellman equation to characterize the solution to the planning problem and to show that it displays mean reversion. Our first result establishes that $k(v)$ is differentiable and strictly concave with an interior peak.

---

[8] If one assumes bounded utility, then when the contraction mapping theorem is applied, the Bellman equation is guaranteed to have a unique solution, which must then coincide with $k(v)$ defined from the sequence problem. However, we do not assume bounded utility and, for our purposes, find it unnecessary to solve fixed points of the Bellman equation or prove that it has a unique solution. Instead, we work directly with $k(v)$ defined from the sequence problem (13) and simply exploit the fact that this function satisfies the Bellman equation.

[9] The proof of theorem 2 applies versions of the principle of optimality to verify incentive compatibility. In particular, for any (g^u, g^w) and an initial v_0, a dynasty faces a recursive dynamic programming problem with state variable v_t and with the report $\hat{\theta}_t$ as the control. Conditions (15) and (16) then amount to guessing and verifying a solution to the Bellman equation of the agent's problem, i.e., that the identity function satisfies the Bellman equation with truth-telling. The limiting condition (17) then verifies that this represents the dynasty's value from the sequential problem.

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

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

Let λ be the multiplier on the promise-keeping constraint (15) and let $\mu(\theta, \theta')$ be the multipliers on the incentive constraints (16). The first-order condition for $u(\theta)$ is

$$[\theta - \eta C'(g^u(\theta, v))]p(\theta) - \theta \lambda p(\theta) + \sum_{\theta' \in \Theta} \theta \mu(\theta, \theta') - \sum_{\theta' \in \Theta} \theta' \mu(\theta', \theta) \leq 0,$$

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

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

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

$$\sum_{\theta \in \Theta} k'(g^w(\theta, v))p(\theta) = \frac{\beta}{\hat{\beta}} k'(v). \tag{18}$$

In sequential notation, this condition is

$$\mathbb{E}_{t-1}[k'(v_{t+1})] = \frac{\beta}{\hat{\beta}} k'(v_t), \tag{19}$$

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

To provide incentives, the planner rewards the descendants of an individual reporting a low taste shock. Rewards can take two forms, and it is optimal to make use of both. The first is standard and involves spending more on a dynasty in present-value terms. The second is more subtle and exploits differences in preferences: it is to allow an adjustment in the pattern of consumption, for a given present value, in the direction preferred by individuals relative to the planner.[10] Since indi-

---

[10] 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 television time. In these instances, the child values some goods more than the parent wishes, and the parent uses them to provide incentives.

viduals are more impatient than the planner, this form of reward is delivered by tilting the consumption profile toward the present. Earlier consumption dates are used more intensively to provide incentives: rewards and punishments are front-loaded.

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

It is convenient to reexpress equation (19) as

$$\mathbb{E}_{t-1}[1 - k'(v_{t+1})] = \frac{\beta}{\hat{\beta}}[1 - k'(v_t)] + 1 - \frac{\beta}{\hat{\beta}}, \tag{20}$$

so that the stochastic process $\{1 - k'(v_t)\}$ reverts toward one. Our next result derives upper and lower bounds for the evolution of this process.

PROPOSITION 2.   For $1 - k'(v) \geq 0$,

$$\underline{\gamma}[1 - k'(v)] + \left(1 - \frac{\beta}{\hat{\beta}}\right) \leq 1 - k'(g^w(\theta, v)) \leq \overline{\gamma}[1 - k'(v)] + \left(1 - \frac{\beta}{\hat{\beta}}\right) \tag{21}$$

for some constants $\underline{\gamma}$ and $\overline{\gamma}$ with $\underline{\gamma} \leq \beta/\hat{\beta} \leq \overline{\gamma}$ and $\underline{\gamma}, \overline{\gamma} \to \beta/\hat{\beta}$ as $\overline{\theta}/\underline{\theta} \to 1$. Moreover, consumption $C(g^u(\theta, v))$ is zero if and only if $1 - k'(v) \leq 0$.

The bounds in (21) are instrumental in proving that a steady-state distribution exists, but they also illustrate a powerful force away from misery. Proposition 2 can be seen as providing a corridor around the expected value $(\beta/\hat{\beta})[1 - k'(v)] + 1 - (\beta/\hat{\beta})$ for the realization $1 - k'(g^w(\theta, v))$. This corridor becomes narrower as $1 - k'(v)$ is decreased and shrinks to zero as $1 - k'(v) \to 0$. This implies that welfare must rise, for all realized shocks, if current welfare is low enough. Indeed, if utility is unbounded below, then next period's welfare $g^w(\theta, v)$ remains bounded even as $v \to -\infty$. No matter how badly a parent is to be punished, the child is always somewhat spared.

When the solution for $u(\theta)$ is interior,

$$\eta \sum_{\theta \in \Theta} C'(g^u(\theta, v))p(\theta) = 1 - k'(v).$$

Equation (20) then implies

$$\mathbb{E}_{t-1}\left[\frac{1}{U'(c_{t+1})}\right] = \frac{\beta}{\hat{\beta}}\mathbb{E}_{t-1}\left[\frac{1}{U'(c_t)}\right] + \left(1 - \frac{\beta}{\hat{\beta}}\right)\frac{1}{\eta}.$$

For example, in the logarithmic utility case, $1/U'(c) = c$, implying that

the one-step-ahead forecast for consumption $x_t = \mathbb{E}_{t-1}[c_t]$ mean-reverts at rate $\beta/\hat{\beta}$ toward its mean, η^{-1}. That is,

$$\mathbb{E}_{t-1}[x_t] = \frac{\beta}{\hat{\beta}} x_{t-1} + \left(1 - \frac{\beta}{\hat{\beta}}\right)\frac{1}{\eta}.$$

Moreover, if the amplitude of taste shocks is not too wide, we can guarantee that $\overline{\gamma} < 1$ and $\underline{\chi} > 0$. If $\overline{\gamma} < 1$, then the ergodic set for $1 - k'(v)$ is bounded above, implying that welfare v_t is bounded away from $\bar{v}$; whereas if $\underline{\chi} > 0$, then the ergodic set for $1 - k'(v)$ is bounded away from zero and consumption is bounded away from zero. In this way, one can guarantee that inequality of welfare and consumption remains bounded.

VI. Existence of a Steady State

In this section we show that a steady-state invariant distribution exists. The proof relies on the mean reversion in equation (19) and the bounds in proposition 2.

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

When any of the conditions of proposition 3 are satisfied, theorem 2 leaves open only two possibilities. Either the social planning problem admits a steady-state invariant distribution or no solution exists. This contrasts with the Atkeson-Lucas case, with $\beta = \hat{\beta}$, where a solution exists but does not admit a steady state. Later in this section we verify the second part of theorem 2 to confirm that a solution to the social planning problem can be guaranteed and a steady state exists.

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

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

The Bellman equation for $\hat{\beta} > \beta$ is

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

subject to (15) and (16). If we ignore the nonnegativity constraints on u and w, this is a linear-quadratic dynamic programming problem, so
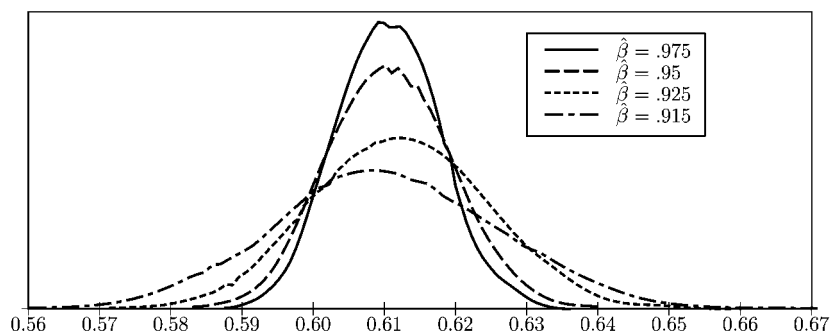
Fɪɢ. 2.—Invariant distributions for welfare, measured in consumption-equivalent units $C((1 - \beta)v)$, for various values of $\hat{\beta}$.

the value function $k(v)$ is a quadratic function and the policy functions are linear in v:

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

and

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

For taste shocks that are not too wide, we can guarantee strictly positive consumption and a bounded ergodic set for welfare. The nonnegativity constraints are then satisfied, so that this solves the problem that imposes them.

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

We now briefly discuss issues of uniqueness and stability of the invariant distribution guaranteed by proposition 3. This question is of economic interest because it represents an even stronger notion of social mobility than that implied by the mean reversion condition (19) discussed in the previous section. Suppose that the economy finds itself at a steady state ψ^*. Then convergence from any initial v_0 toward the distribution ψ^* means that the distribution of welfare for distant de-

scendants is independent of an individual's present condition. The past exerts some influence on the present, but its influence dies out over time. The inheritability of welfare is imperfect, and the advantages or disadvantages of distant ancestors are eventually wiped out.

Indeed, the solution may display this strong notion of social mobility. To see this, suppose that the ergodic set for $\{k'(v_t)\}$ is compact, which is guaranteed if $\overline{\gamma} < 1$. Then if the policy function $g^w(\theta, v)$ is monotone in v, the invariant distribution ψ^* is unique and stable: starting from any initial distribution ψ_0, the sequence of distributions $\{\psi_t\}$, generated by g^w, converges weakly to ψ^*.[11] The required monotonicity of the policy functions was satisfied by examples 1 and 2 and seems plausible more generally.[12] Another approach suggests uniqueness and convergence without relying on monotonicity of the policy functions. Grunwald et al. (2000) show that one-dimensional, irreducible Markov processes with the Feller property that are bounded below and are conditional linear autoregressive, as implied by (19), have a unique and stable invariant distribution. All their hypotheses have been verified here except for the technical condition of irreducibility.[13] Although we do not pursue this further, our discussion illustrates how the forces for reversion in (19) might be exploited to establish uniqueness and convergence.

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

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

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

---

[11] This follows since the conditional-expectation equation (19) ensures enough mixing to apply Hopenhayn and Prescott's theorem (see Stokey and Lucas 1989, 382–83).

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

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

As mentioned in Section IV, for any utility function one can characterize the solution for any (ψ, e) as the solution to a relaxed problem with some sequence $\{Q_t\}$ that is not necessarily exponential, that is, imposing the general intertemporal constraint (10) instead of (11). Proposition 4 identifies the distributions and endowment pairs (ψ, e) that lead to exponential $\{Q_t\}$ in the logarithmic case. More generally, with logarithmic utility for *any* pair (ψ, e), we can show that $Q_t = \hat{\beta}^t + a\beta^t$ for some constant a. The entire optimal allocation can then be characterized by the policy functions from a nonstationary Bellman equation. Since $\{Q_t\}$ is asymptotically exponential (i.e., $\lim_{t\to\infty}\hat{\beta}^{-t}Q_t = 1$), the long-run dynamics are dominated by the policy functions (g^u, g^w) from the problem with $Q_t = \hat{\beta}^t$ that we have characterized.

We have shown the existence of a steady state ψ^* generated by the policy function g^w. We now provide sufficient conditions to ensure that ψ^* is also a steady state for the social planning problem. This involves two things. We first establish that allocations generated by the policy functions are indeed incentive compatible by verifying the limiting condition (17) in theorem 2; this guarantees that, given ψ^* and η, the allocation maximizes the Lagrangian (12). Second, we verify that average consumption is finite under ψ^*, so that there exists some endowment e for which the resource constraints (4) and (10) hold. It follows that the allocations generated by (g^u, g^w) solve the social planning problem, given e and ψ^*.

PROPOSITION 5.   The allocation generated by the policy functions (g^u, g^w) starting at any v_0 solves the component planning problem in any of the following cases: (*a*) utility is bounded above, (*b*) utility is bounded below, (*c*) utility is logarithmic, or (*d*) $\overline{\gamma} < 1$ or $\chi > 0$.

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

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

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

if either (*a*) the function $C'(U(c)) = 1/U'(c)$ is convex over $c \geq \bar{c}$ for some $\bar{c} < \infty$ or (*b*) $\chi > 0$ or $\overline{\gamma} < 1$.

It is worth remarking that the hypotheses of all these propositions are met for a wide range of primitives. In particular, they hold for any utility function as long as the amplitude of taste shocks is not too wide,

so that we can ensure that $\chi > 0$ or $\overline{\gamma} < 1$. They also hold for any arbitrary amplitude of taste shocks when utility is logarithmic or when utility is bounded above and $1/U'(c)$ is asymptotically convex. For example, Atkeson and Lucas (1992) focused on the CRRA specification $U(c) = c^{1-\sigma}/(1 - \sigma)$ with $\sigma > 0$. All our results apply in this specification: for $\sigma \in [1, \infty)$ with any shock distribution and for $\sigma \in (0, 1)$ with shocks that are not too wide.

The steady-state distribution ψ and the implied value of total consumption will generally vary with η. Thus different values η translate into different required endowments e. For the CRRA case, we can say that steady-state consumption is a power function of η and thus has full range. In fact, in this case the entire solution for consumption is homogeneous 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.

VII.   Pareto Problem

We now return to the Pareto problem (9) and its relation to the social planning problem. Recall that the former is exactly as the latter except that the promise-keeping constraints are inequalities instead of equalities. The next result establishes that these inequality constraints bind for steady-state distributions ψ^* with strictly positive consumption. Thus the solutions to the Pareto and social planning problems coincide. The proof relies on the fact that a marginal increase in v contributes $k'(v) - 1$ to the welfare criterion (9) (see also [7]) and that $k'(v) < 1$, unless consumption is zero for some agents. Recall that $S(\psi; e)$ was defined in (8).

PROPOSITION 7.   Let ψ^* denote a steady state for the social planning problem. Suppose that consumption is strictly positive for all agents, so that $C(g^u(\theta, v)) > 0$ for all v in the support of ψ^*. Then $\psi = \psi^*$ solves the Pareto problem

$$\max_{\psi} \left[S(\psi; e) - \int v d\psi(v) \right]$$

subject to

$$\psi(v) \leq \psi^*(v) \quad \text{for all } v.$$

Thus steady states for the social planning problem coincide with steady states for the Pareto problem as long as consumption is positive. That is, if the Pareto problem is started with the distribution $\tilde{\psi} = \psi^*$, then it is replicated over time and $\psi_t = \psi^*$ for all $t = 0, 1,$ .... By implication, the Pareto optimum is then time-consistent: the

initial solution at $t = 0$ also solves the Pareto problem at any future period t. In other words, this Pareto-efficient allocation is ex post Pareto efficient. Note that the condition that consumption be strictly positive is guaranteed if utility is unbounded below or if the amplitude of the taste shocks is not too wide so that $\overline{\gamma} > 0$ (see proposition 2).

In the Pareto problem (9), the welfare of future generations was aggregated using geometric Pareto weights. We now map these weights into their welfare implications and discuss a planning problem that is cast directly in terms of welfare, without Pareto weights.

Starting from the steady state ψ^*, the optimum for the Pareto problem with weights $\alpha_t = \hat{\beta}^t$ for $t \geq 1$ has $\psi_t = \psi^*$ and delivers a constant expected welfare level $V^* \equiv \int v d\psi^*(v)$ to all future generations. Consider then the problem of maximizing the expected welfare $\mathbb{E}_{-1}[v_s] = \int v_s d\psi_s(v_s)$ of any particular generation $s \geq 1$ subject to delivering expected welfare of at least V^* for all other future generations,

$$\int v_t d\psi_t(v_t) \geq V^*, \tag{22}$$

while delivering $\psi_0 \leq \tilde{\psi}$ to the first generation in an incentive-compatible way for some given initial distribution $\tilde{\psi}$. Thus this planning problem solves for an efficient point on the frontier of attainable welfare for founding members of each dynasty, v_0, and expected welfare for future generations, $\mathbb{E}_{-1}[v_t]$ for $t = 1, 2, \ldots$. A steady state ψ^* for the Pareto problem (9) is also a steady state for this planning problem and represents a symmetric point on this frontier: given $\tilde{\psi} = \psi^*$ and $V^* = \int v d\psi^*(v)$, the solution has $\psi_t = \psi^*$. The Pareto weights $\alpha_t = \hat{\beta}^t$ represent the Lagrange multipliers on the inequality constraints (22).

VIII.   Conclusions

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

The backbone of our model requires a trade-off between insurance and incentives. The source for this trade-off is inessential. In this paper,

we adopted the Atkeson-Lucas taste-shock specification for purposes of comparison. In Farhi and Werning (2006), we study a dynamic Mirrleesian model—with productivity shocks instead of taste shocks—and find that a progressive estate tax implements efficient allocations by providing the necessary mean reversion across generations.

Appendix

Proof of Theorem 1

The value function $k(v)$ defined by (13) is weakly concave since the objective function is concave and the constraint set convex. Weak concavity implies continuity over the interior of its domain: $(\underline{v}, \bar{v})$. If utility is bounded below, then continuity at $\bar{v}$ is established as follows. Define the first-best value function

$$k^*(v) \equiv \max_{\{u_t\}} \sum_{t=0}^{\infty} \sum_{\theta^t \in \Theta^{t+1}} \hat{\beta}^t [\theta_t u_t(\theta^t) - \eta C(u_t(\theta^t))] \Pr(\theta^t)$$

subject to

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

Then $k^*(v)$ is continuous and $k(v) \leq k^*(v)$ with equality at $\underline{v}$. Since $k(v)$ is weakly convex, $\lim_{v \downarrow \underline{v}} k(v) \geq k(\underline{v})$. Toward a contradiction, suppose $\lim_{v \downarrow \underline{v}} k(v) > k(\underline{v})$. Then

$$0 \leq \lim_{v \downarrow \underline{v}} [k^*(v) - k(v)] = k^*(\underline{v}) - \lim_{v \downarrow \underline{v}} k(v) < k^*(\underline{v}) - k(\underline{v}),$$

a contradiction since $k(\underline{v}) - k^*(\underline{v}) = 0$. Thus $k(\underline{v})$ must be continuous at $\underline{v}$. This completes the proof that $k(v)$ is continuous.

We first show that the constraint (11) with $q = \hat{\beta}$ implies that utility and continuation welfare are well defined. Toward a contradiction, suppose that

$$\lim_{T \to \infty} \sum_{t=0}^{T} \beta^t \mathbb{E}_s \theta_t u_t$$

is not defined, for some $s \geq -1$. This implies that

$$\lim_{T \to \infty} \sum_{t=0}^{T} \beta^t \max\{\mathbb{E}_s \theta_t u_t, \ 0\} = \infty.$$

Since utility is concave, $\theta u \leq AC(u) + B$ (i.e., $\theta U(c) \leq Ac + B$) for some $A, B > 0$, so it follows that

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

Taking the limit yields $\lim_{T \to \infty} \sum_{t=0}^{T} \hat{\beta}^t \mathbb{E}_{-1}[C(u_t)] = \infty$. Since there are finitely many histories $\theta^s \in \Theta^{s+1}$, this implies $\lim_{T \to \infty} \sum_{t=0}^{T} \hat{\beta}^t \mathbb{E}_{-1}[C(u_t)] = \infty$. If there is a nonzero measure of such agents, this implies a contradiction of the intertemporal constraint (11) and thus of at least one resource constraint in (4). Thus, for both

the relaxed and the original social planning problems, utility and continuation welfare are well defined, which is important for the recursive formulation.

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

$$\sup_{u,w} \mathbb{E}[\theta u(\theta) - \eta C(u(\theta)) + \hat{\beta} k(w(\theta))] \tag{A1}$$

subject to (15) and (16). Define $m \equiv \max_{c \geq 0, \theta \in \Theta} [\theta U(c) - \eta c]$ and $\hat{k}(v) \equiv k(v) - m/(1-\hat{\beta}) \leq 0$. The maximization in (A1) is then equivalent, up to constants, to

$$\sup_{u,w} \mathbb{E}[\theta u(\theta) - \eta C(u(\theta)) - m + \hat{\beta}\hat{k}(w(\theta))] \tag{A2}$$

subject to (15) and (16). The objective function in (A2) is nonpositive, which simplifies the arguments below.

Lemma A1.   The supremum in (A1), or equivalently (A2), is attained.

Proof.   Suppose first that utility is unbounded below. We show that

$$\lim_{v \to \bar{v}} \hat{k}(v) = \lim_{v \to -\infty} \hat{k}(v) = -\infty \tag{A3}$$

and then use this result to restrict, without loss, the optimization within a compact set, ensuring that a maximum is attained. To establish these limits, define the function

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

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

$$h(v, \hat{\beta}) \leq h(v, \beta) = v - \eta \hat{C}(v, \beta) - \frac{m}{1-\beta}, \tag{A4}$$

where

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

subject to

$$v = \sum_{t=0}^{\infty} \beta^t \mathbb{E}_{-1}[\theta_t u(\theta^t)].$$

This is a standard allocation problem, with solution $u(\theta^t) = (C')^{-1}(\theta\gamma(v))$ 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

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

(note that here $(C')^{-1} : \mathbb{R}_+ \to U(\mathbb{R}_+)$ represents the inverse function of

$C'(u) : U(\mathbb{R}_+) \to \mathbb{R}_+$; i.e., if $y = C'(x)$, then $(C')^{-1}(y) = x$), so that $\lim_{v \to -\infty} h(v, \beta) = -\infty$ and $\lim_{v \to \infty} h(v, \beta) = -\infty$. Using the inequality (A4) 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 (A3), using the fact that $\hat{k}(v) \leq h(v, \hat{\beta})$.

Fix a v. Then $w(\theta) = v$ and $u(\theta) = (1 - \beta)v$ satisfies constraints (15) and (16), so that the maximized value must be greater than $\overline{k} = v(1 - \beta) - C(v(1 - \beta)) - m + \hat{\beta}\hat{k}(v)$. Then, since the objective is nonpositive, we can restrict the maximization over $w(\theta)$ so that $\hat{k}(w(\theta)) \geq \overline{k}/[\hat{\beta}p(\theta)]$. Since $\hat{k}(w(\theta))$ is concave with the limits (A3), this defines a closed, bounded interval for $w(\theta)$ for each $\theta \in \Theta$. Similarly, we can restrict the maximization over $u(\theta)$ so that $\theta u(\theta) - \eta C(u(\theta)) - m \geq \overline{k}/p(\theta)$. Since $\theta u - \eta C(u)$ is strictly concave, with $\theta u - \eta C(u) \to -\infty$ when either $u \to \infty$ or $u \to -\infty$, this defines a closed, bounded interval for $u(\theta)$ for each $\theta \in \Theta$. Hence, without loss of generality, we can restrict the maximization of the continuous objective function (A2) to a compact set, so that a maximum must be attained.

If utility is bounded below, then $\lim_{v \to \underline{v}} \hat{k}(v) = \underline{v} > -\infty$ and (by the same argument as above) $\lim_{v \to \overline{v}} \hat{k}(v) = -\infty$. Hence, the restrictions $\hat{k}(w(\theta)) \geq \overline{k}/[\hat{\beta}p(\theta)]$ and $\theta u(\theta) - \eta C(u(\theta)) - m \geq \overline{k}/p(\theta)$ continue to define, given v, closed intervals for $w(\theta)$ and $u(\theta)$ for each $\theta \in \Theta$. Again, restricting the maximization in (A2), without loss of generality, within these intervals ensures that the maximum is attained. QED

LEMMA A2.   The value function $k(v)$ satisfies the Bellman equation (14)–(16).

Proof.   Suppose that for some v

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

where the maximization is subject to (15) and (16). Then there exists $\Delta > 0$ such that

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

for all (u, w) that satisfy (15) and (16). But then by definition

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

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

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

for *all* incentive-compatible allocations that deliver v, a contradiction with the definition of $k(v)$; namely, there should be a plan with value arbitrarily close to $k(v_0)$. We conclude that

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

subject to (15) and (16).

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

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

Let

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

Consider the plan $u_0(\theta_0) = u^*(\theta_0)$ and $u_t(\theta^t) = \tilde{u}_{t-1}((\theta_1,\ \ldots,\ \theta_t);\ w^*(\theta_0),\ \epsilon)$ for $t \geq 1$. Then

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

$$= \mathbb{E}_{-1}\left[\theta_0 u^*(\theta_0) - \eta C(u^*(\theta_0)) + \hat{\beta} \sum_{t=0}^{\infty} \hat{\beta}^t \mathbb{E}_0[\theta_{t+1} u_{t+1}(\theta^{t+1}) - \eta C(u_{t+1}(\theta^{t+1}))]\right]$$

$$\geq \max_{u,w} \mathbb{E}[\theta u(\theta) - \eta C(u(\theta)) + \hat{\beta} k(w(\theta))] - \hat{\beta}\epsilon.$$

Since $\epsilon > 0$ was arbitrary, it follows that $k(v) \geq \max_{u,w} \mathbb{E}[\theta u(\theta) - \eta C(u(\theta)) + \hat{\beta} k(w(\theta))]$ subject to (15) and (16). QED

Proof of Theorem 2

We establish the following results from which the theorem follows: (*a*) An allocation $\{u_t\}$ is optimal for the component planning problem (13) with multiplier η, given $v_0 = v$, if and only if it is generated by the policy functions (g^u, g^w) starting at v_0, is incentive compatible, and delivers welfare v_0. (*b*) An allocation $\{u_t\}$ generated by the policy functions $(g^u,\ g^w)$, starting at v_0, has $\lim_{t\to\infty} \beta^t \mathbb{E}_{-1}[v_t(\theta^{t-1})] = 0$ and delivers welfare v_0. (*c*) An allocation $\{u_t\}$ generated by the policy functions (g^u, g^w), starting from v_0, is incentive compatible if

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

for all reporting strategies σ.

   Part *a*: Suppose that the allocation $\{u_t\}$ is generated by the policy functions starting from v_0, is incentive compatible, and delivers welfare v_0. After repeated substitutions of the Bellman equation (14), we arrive at

$$k(v_0) = \sum_{t=0}^{T} \hat{\beta}^t \mathbb{E}_{-1}[\theta_t u_t(\theta^t) - \eta C(u_t(\theta^t))] + \hat{\beta}^{T+1} \mathbb{E}_{-1} k(v_{T+1}(\theta^T)). \tag{A5}$$

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

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

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

   Conversely, suppose that an allocation $\{u_t\}$ is optimal given v_0. Then by defi-

nition it must be incentive compatible and deliver welfare v_0. Define the continuation welfare implicit in the allocation

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

and suppose that either $u_0(\theta) \neq g^u(\theta; \ v_0)$ or $w_0(\theta) \neq g^w(\theta; \ v_0)$ for some $\theta \in \Theta$. Since the original plan $\{u_t\}$ is incentive compatible, $u_0(\theta)$ and $w_0(\theta)$ satisfy (15) and (16). The Bellman equation then implies that

$$k(v_0) = \mathbb{E}[g^u(\theta; \ v_0) - \eta C(g^u(\theta; \ v_0)) + \beta k(g^w(\theta; \ v_0))]$$

$$> \mathbb{E}[u_0(\theta) - \eta C(u_0(\theta)) + \beta k(w_0(\theta))]$$

$$\geq \mathbb{E}_{-1}[u_0(\theta_0) - \eta C(u_0(\theta_0))] + \sum_{t=1}^{\infty} \beta^t \mathbb{E}_{-1}[u_t(\theta^t) - \eta C(u_t(\theta^t))].$$

The first inequality follows since u_0 does not maximize (14); 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 that an allocation $\{u_t, v_t\}$ generated by the policy functions (g^u, g^w) starting at v_0 satisfies $\lim_{t\to\infty} \beta^t \mathbb{E}_{-1} v_t(\theta^{t-1}) = 0$. Then, after repeated substitutions of (15), we obtain

$$v_0 = \sum_{t=0}^{T} \beta^t \mathbb{E}_{-1}[\theta_t u_t(\theta^t)] + \beta^{T+1} \mathbb{E}_{-1}[v_{T+1}(\theta^T)]. \tag{A6}$$

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

$$\lim_{t\to\infty} \beta^t \mathbb{E}_{-1} v_t(\theta^{t-1}) = 0.$$

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

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

and then (A5) implies that $k(v_0) = -\infty$, a contradiction since there are feasible plans that yield finite values. We conclude that $\limsup_{t\to\infty} \beta^t \mathbb{E}_{-1} v_t(\theta^{t-1}) \leq 0$.

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

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

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

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

Part c: Suppose that $\limsup_{t\to\infty}\beta^t\mathbb{E}_{-1}v_t(\sigma^{t-1}(\theta^{t-1})) \geq 0$ for every reporting strategy σ. Then after repeated substitutions of (16),

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

implying

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

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

Proof of Proposition 1

Strict concavity.—Let $\{u_t(\theta^t, v_0), v_t(\theta^{t-1}, v_0)\}$ be generated from the policy functions starting at v_0 (note that no claim of incentive compatibility is required). Take two initial welfare values v_a and v_b, with $v_a \neq v_b$. Define the average utilities

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

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

As shown in the proof of theorem 2, $\{u_t(\theta^t, v_a)\}$ and $\{u_t(\theta^t, v_b)\}$ deliver welfare v_a and v_b, respectively. This implies that $\{u_t^\alpha(\theta^t)\}$ delivers welfare $v^\alpha \equiv \alpha v_a + (1-\alpha)v_b$. It also implies that

$$u_t(\theta^t; v_a) \neq u_t(\theta^t; v_b), \tag{A7}$$

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

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

and

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

and averaging we obtain for large enough T

$$\alpha k(v_a) + (1-\alpha)k(v_b)$$

$$= \sum_{t=0}^{T} \hat{\beta}^t \mathbb{E}_{-1}[\theta_t u_t^\alpha(\theta^t) - \eta[\alpha C(u_t(\theta^t; v_a)) + (1-\alpha)C(u_t(\theta^t; v_b))]]$$

$$+ \hat{\beta}^{T+1} \mathbb{E}_{-1}[\alpha k(v_{T+1}(\theta^T; v_a)) + (1-\alpha)k(v_{T+1}(\theta^T; v_b))]$$

$$< \sum_{t=0}^{T} \hat{\beta}^t \mathbb{E}_{-1}[\theta_t u_t^\alpha(\theta^t) - \eta C(u_t^\alpha(\theta^t))] + \hat{\beta}^{T+1} \mathbb{E}_{-1}k(v_{T+1}^\alpha(\theta^T)) \le k(v^\alpha),$$

where the strict inequality follows from the strict concavity of the cost function $C(u)$, the fact that we have the inequality (A7), and the weak concavity of the value function k. The last weak inequality follows from iterating on the Bellman equation for v^α since the average plan (u^α, v^α) satisfies the constraints of the Bellman equation at every step. This proves that the value function $k(v)$ is strictly concave.

Differentiability.—Since $k(v)$ is concave, it is subdifferentiable: there is at least one subgradient at every v. We establish differentiability by proving that there is a unique subgradient by variational envelope arguments.

Suppose first that utility is unbounded below. Fix an interior value v_0. In a neighborhood of v_0 define the test function

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

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

$$k'(v_0) = W'(v_0) = 1 - \eta\mathbb{E}[C'(u^*(\theta))]. \tag{A8}$$

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

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 $k(v) \le h(v, \beta) + m/(1 - \hat{\beta})$ introduced in the proof of theorem 1, since $\lim_{v \to \infty} \partial h(v, \beta)/\partial v = -\infty$.

Next, suppose that utility is bounded below but unbounded above. Without loss of generality, we normalize so that $U(0) = 0$. Then $\limsup_{t \to \infty} \mathbb{E}_{-1}\beta^t v_t(\sigma^{t-1}(\theta^{t-1})) \ge 0$ for all reporting strategies σ so that, when we apply theorem 2, a solution $\{u_t\}$ to the planner's sequence problem is ensured. Then, for any interior v_0, the plan $\{(v/v_0)u_t\}$ is incentive compatible and attains value v for the agent. In addition, the test function

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

satisfies $W(v) \le k(v)$, $W(v_0) = k(v_0)$ 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

$$\underline{k}(v) \equiv \max_u \sum_{t=0}^{\infty} \hat{\beta}^t [u_t - \eta C(u_t)]$$

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

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

Proof of Proposition 2

We first show that we can simplify problem (14)–(16). Let $\Theta = \{\theta_1, \theta_2, \ldots, \theta_N\}$ with $\underline{\theta} = \theta_1 < \theta_2 < \cdots < \theta_N = \overline{\theta}$. The incentive compatibility constraint (16) is equivalent to the subset that considers only neighboring deviations:

$$\theta_n u(\theta_n) + \beta w(\theta_n) \geq \theta_n u(\theta_{n+1}) + \beta w(\theta_{n+1}), \quad n = 1, 2, \ldots, N-1, \qquad \text{(A9)}$$

and

$$\theta_n u(\theta_n) + \beta w(\theta_n) \geq \theta_n u(\theta_{n-1}) + \beta w(\theta_{n-1}), \quad n = 2, 3, \ldots, N. \qquad \text{(A10)}$$

These imply that $u(\theta)$ is nondecreasing and that $w(\theta)$ is nonincreasing.

Now, consider a simplified version of the problem (14)–(16) that replaces (16) with (A9) and the *monotonicity constraint* that $u(\theta)$ be nondecreasing (ignoring the constraints [A10]). We show next that at the solution to this simplified problem, all the inequalities (A9) hold with equality. This, in turn, implies that the constraints (A10) are satisfied. As a result, the solution to this simplified problem is also a solution to problem (14)–(16).

That the constraints (A9) hold with equality at a solution to the simplified problem is proved by contradiction. Suppose that, for some $n = 1, 2, \ldots, K-1$, (A9) is slack:

$$\theta_n u(\theta_n) + \beta w(\theta_n) > \theta_n u(\theta_{n+1}) + \beta w(\theta_{n+1}). \qquad \text{(A11)}$$

This implies, from the monotonicity constraint, that $u(\theta_{n+1}) > u(\theta_n)$ and $w(\theta_{n+1}) < w(\theta_n)$. Now consider the alternative allocation that changes only $w(\theta_n)$ and $w(\theta_{n+1})$: decreasing $w(\theta_n)$ and increasing $w(\theta_{n+1})$, while keeping the average $w(\theta_{n+1})p(\theta_{n+1}) + w(\theta_n)p(\theta_n)$ constant. Such a change does not affect the promise-keeping constraint (15). Increasing $w(\theta_{n+1})$ relaxes the incentive constraint (A9) for $n+1$, whereas lowering $w(\theta_n)$ is feasible because of the strict inequality (A11). Thus a small change in this direction is feasible. However, since $k(w)$ is strictly concave in w (proposition 1), this change increases the objective (14), a contradiction with optimality. Thus, at an optimum, the constraints (A9) hold with equality.

The arguments above justify focusing on the simplified problem that maximizes (14) subject to (15) (with multiplier λ), (A9) (with multiplier μ_n), and

$u(\theta_{n+1}) \leq u(\theta_n)$ (with multiplier ϑ_n) for $n = 1, \ldots, N-1$. The first-order conditions are

$$p(\theta_n)[\theta_n - \eta C'(g^u(\theta_n, v)) - \lambda\theta_n] + \theta_n\mu_n - \theta_{n-1}\mu_{n-1} - \vartheta_n + \vartheta_{n-1} \leq 0,$$

$$p(\theta_n)[\hat{\beta}k'(g^w(\theta_n, v)) - \beta\lambda] + \beta(\mu_n - \mu_{n-1}) = 0$$

for $n = 1, 2, \ldots, N$, where $\mu_0 = \mu_N = \vartheta_0 = \vartheta_N = 0$, and we adopt the convention that $\theta_0 = \theta_1$.

Consider first the case in which utility is unbounded below, so that the solution for $u(\theta)$ is interior. Summing the first-order conditions for $u(\theta)$ over all $\theta \in \Theta$, we get

$$\eta\sum_{\theta \in \Theta} C'(g^u(\theta, v))p(\theta) = 1 - k'(v)$$

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

Now take the set of all agents $n = 1, \ldots, \underline{m}$ that receive the same allocation as the agent with the lowest shock $\underline{\theta}$; then $\vartheta_{\underline{m}} = 0$. Then summing the first-order conditions for $u(\theta_n)$ over $n = 1, \ldots, \underline{m}$ implies

$$\mathbb{E}[\theta|\theta \leq \theta_{\underline{m}}][1 - k'(v)] + \frac{\theta_{\underline{m}}\mu_{\underline{m}}}{\Pr(\theta \leq \theta_{\underline{m}})} = \eta C'(g^u(\underline{\theta}, v)) \leq \eta\sum_{\theta \in \Theta} C'(g^u(\theta, v))p(\theta)$$

$$= 1 - k'(v), \tag{A12}$$

where the inequality follows by $g^u(\underline{\theta}, v) \leq g^u(\theta, v)$. Summing the first-order conditions for $w(\theta_n)$ over $n = 1, \ldots, \underline{m}$, we get

$$\frac{\mu_{\underline{m}}}{\Pr(\theta \leq \theta_{\underline{m}})} = k'(v) - \frac{\hat{\beta}}{\beta}k'(g^w(\underline{\theta}, v)) \geq k'(v) - \frac{\hat{\beta}}{\beta}k'(g^w(\theta, v)),$$

where the inequality follows by $g^w(\underline{\theta}, v) \geq g^w(\theta, v)$. Combining both inequalities gives

$$1 - k'(g^w(\theta, v)) \leq \frac{\beta}{\hat{\beta}}\frac{\beta\theta_{\underline{m}} + 1 - \mathbb{E}[\theta|\theta \leq \theta_{\underline{m}}]}{\theta_{\underline{m}}}[1 - k'(v)] + 1 - \frac{\beta}{\hat{\beta}}. \tag{A13}$$

Symmetrically, take the set of all agents $n = \overline{m}, \ldots, N$ receiving the same allocations as the agent with the highest taste shock $\overline{\theta}$; then $\vartheta_{\overline{m}} = 0$. Then adding up their first-order condition gives

$$\mathbb{E}[\theta|\theta \geq \theta_{\overline{m}}][1 - k'(v)] - \frac{\theta_{\overline{m}-1}\mu_{\overline{m}-1}}{\Pr(\theta \geq \theta_{\overline{m}})} = \eta C'(g^u(\overline{\theta}, v)) \geq \eta\sum_{\theta \in \Theta} C'(g^u(\theta, v))p(\theta)$$

$$= 1 - k'(v), \tag{A14}$$

$$-\frac{\mu_{\overline{m}-1}}{\Pr(\theta \geq \theta_{\overline{m}})} = k'(v) - \frac{\hat{\beta}}{\beta}k'(g^w(\overline{\theta}, v)) \leq k'(v) - \frac{\hat{\beta}}{\beta}k'(g^w(\theta, v)),$$

where the inequality arises from $g^u(\overline{\theta}, v) \geq g^u(\theta, v)$ and $g^w(\overline{\theta}, v) \leq g^w(\theta, v)$. Combining both inequalities gives

$$\frac{\beta}{\hat{\beta}} \frac{\theta_{\overline{m}-1} + 1 - \mathbb{E}[\theta | \theta \geq \theta_{\overline{m}}]}{\theta_{\overline{m}-1}} [1 - k'(v)] + 1 - \frac{\beta}{\hat{\beta}} \leq 1 - k'(g^w(\theta, v)). \qquad \text{(A15)}$$

To arrive at expression (21) take the $\underline{m}$ that maximizes the right-hand side of (A13) and $\overline{m}$ that minimizes the left-hand side of (A15) to define

$$\underline{\gamma} \equiv \frac{\beta}{\hat{\beta}} \min_{1 \leq n \leq N} \frac{\theta_{n-1} + 1 - \mathbb{E}[\theta | \theta \geq \theta_n]}{\theta_{n-1}}$$

and

$$\overline{\gamma} \equiv \frac{\beta}{\hat{\beta}} \max_{1 \leq n \leq N} \frac{\theta_n + 1 - \mathbb{E}[\theta | \theta \leq \theta_n]}{\theta_n}$$

(recall the convention that $\theta_0 = \theta_1$). Note that $\underline{\gamma} \leq \beta/\hat{\beta} \leq \overline{\gamma}$ and that $\underline{\gamma} \geq (\beta/\hat{\beta})(\underline{\theta} + 1 - \overline{\theta}/\underline{\theta})/\underline{\theta}$ and $\overline{\gamma} \leq (\beta/\hat{\beta})(\overline{\theta} + 1 - \underline{\theta}/\overline{\theta})/\overline{\theta}$. It follows that in the limit as $\overline{\theta}/\underline{\theta} \to 1$ (which requires both $\overline{\theta} \to 1$ and $\underline{\theta} \to 1$ given that $\mathbb{E}[\theta] = 1$), we get $\underline{\gamma} \to \beta/\hat{\beta}$ and $\overline{\gamma} \to \beta/\hat{\beta}$.

In the bounded utility case the arguments are similar, but we need to consider the possibility that consumption is zero. First, take the case in which $\lambda = k'(v) > 1$. As before, let $n = 1, \ldots, \underline{m}$ be the set of all agents who receive the same allocation as the agent with the lowest taste shock $\underline{\theta}$. Then adding up the first-order conditions for $u(\theta)$ for $n = 1, \ldots, \underline{m}$ and evaluating it at $u(\theta) = U(0)$ (so that $C'(u(\theta)) = 0$) requires

$$[1 - k'(v)] \sum_{n=1}^{\underline{m}} \theta_n p(\theta_n) + \theta_{\underline{m}} \mu_{\underline{m}} \leq 0,$$

a contradiction, since the left-hand side is strictly positive: $1 - k'(v) > 0$ and $\mu_{\underline{m}} \geq 0$. Thus, for $k'(v) < 1$ consumption is strictly positive and the same argument used in the unbounded utility case applies to derive the bounds.

Next, consider the case in which $\lambda = k'(v) \geq 1$. As before, let $n = \overline{m}, \ldots, N$ denote the set of all agents who receive the same allocation as the agent with the highest taste shock $\overline{\theta}$. Then adding up the first-order conditions for $u(\theta)$ for $n = \overline{m}, \ldots, N$, we find that an interior solution would require

$$[1 - k'(v) - \eta C'(g^u(\overline{\theta}, v))] \sum_{n=\overline{m}}^{N} \theta_n p(\theta_n) - \theta_{\underline{m}-1} \mu_{\underline{m}-1} = 0,$$

which is not possible since the left-hand side is strictly negative. Thus, for $k'(v) \geq 1$ consumption must be zero: $g^u(\theta, v) = U(0)$ and $g^w(\theta, v) = [v - U(0)]/\beta > v$ for all $\theta \in \Theta$.

Proof of Proposition 3

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

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

for all $x < 1$ if utility is unbounded below. For any probability measure μ, let $T_Q(\mu)$ be the probability measure 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 comparing the limits of $\{T_{Q,n}\}$.

LEMMA A3. If utility is unbounded below, then for any $x < 1$ there exists a subsequence of distributions $\{T_{Q,\phi(n)}(\delta_x)\}$ that converges weakly (i.e., in distribution) to an invariant distribution under Q on $(-\infty, 1)$.

Proof. The bounds (21) derived in proposition 2 imply that, for all $\theta \in \Theta$,

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

We first extend the continuous transition function $Q(x, \theta) : (-\infty, 1) \times \Theta \to (-\infty, 1)$ to a continuous transition function $\hat{Q}(x, \theta) : (-\infty, 1) \times \Theta \to (-\infty, 1)$ 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 measures over $(-\infty, 1]$ to probability measures 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 $\epsilon > 0$ there exists a compact set A_ϵ such that $T_{\hat{Q},n}(\delta_x)[A_\epsilon] \geq 1 - \epsilon$ 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 $\mathbb{E}_{-1}[k'(v_t(\theta^{t-1}))] = (\beta/\hat{\beta})^t k'(v_0) \to 0$. This implies that

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

$$+ \{1 - T_{\hat{Q}}^n(\delta_x)[(-\infty, -A]]\}1$$

for all $A > 0$. Rearranging, we get

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

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, that is, in distribution, to some probability measure π. Since $\hat{Q}(x, \theta)$ is

continuous in x, $T_{\hat{Q}}(T_{\hat{Q}\phi(n)}(\delta_x))$ converges weakly to $T_{\hat{Q}}(\pi)$. But the linearity of $T_{\hat{Q}}$ implies that

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

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

Recall that $T_{\hat{Q}}$ maps probability measures over $(-\infty, 1]$ to probability measures 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 probability measures, it follows that $\pi = T_Q(\pi)$, so that π is an invariant measure under Q on $(-\infty, 1)$. QED

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 $\overline{\gamma} < 1$, then there exists an upper bound $v_H < \bar{v}$ for the ergodic set for v. Define the welfare level $v_0 > \underline{v}$ by $k'(v_0) = 1$ and let v_L be the minimum of $g^w(\overline{\theta}, v)$ over $v \in [v_0, v_H]$; this minimum is attained since g^w is continuous. If utility is bounded above, then let v_L be the minimum of g^w over $v \in [v_0, \bar{v}]$; this minimum is attained since $\lim_{v \to \bar{v}} g^w(\theta, v) = \bar{v}$. In both cases, since $g^w > \underline{v}$, we must have that $v_L > \underline{v}$. Finally, the transition function is continuous with $Q(x, \theta) \leq k'(v_L) < \infty$. The rest of the argument is then a straightforward modification of the one spelled out for the case with utility unbounded below, except that $k'(v_L)$ plays the role of 1. Indeed, things are slightly simpler here since we do not need to construct an extension of Q.

If $\underline{\chi} > 0$, then the bound in (21) implies that $k'(g^w(\overline{\theta}, v)) \leq 1 - (\beta/\hat{\beta})$, and the result follows immediately.

Proof of Proposition 4

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

Since utility is unbounded below, we have $k'(v_t; e) = \mathbb{E}_{t-1}[1 - \eta C'(u_t^v(\theta^t))]$. Applying the law of iterated expectations to (19) then yields

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

With logarithmic utility, $C'(u) = C(u)$, so that $\int k'(v; e)d\psi(v) = 0$ implies

$$\int \mathbb{E}_{-1}[C(u_t^v)]d\psi(v) = \eta^{-1} = e \quad \text{for all } t = 0, 1, \ldots.$$

That the limiting condition (17) is satisfied is shown in proposition 5. It then follows that allocations generated this way solve the relaxed and original social planning problems with $e = \eta^{-1}$.

Now consider any initial distribution ψ. We argue that we can find a value of

$\eta = e^{-1}$ such that $\int k'(v;\ e)d\psi(v) = 0$. The homogeneity of the sequential problem implies that

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

Note that

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

is strictly increasing in e and limits to one as $e \to \infty$ and to $-\infty$ as $e \to -\infty$. It follows that

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

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

Proof of Proposition 5

Part *a*: Suppose that $U(c)$ is bounded above and unbounded below (part *b* deals with the bounded below case); normalize so that $\lim_{c \to \infty} U(c) = 0$ so that $\bar v = 0$. Proposition 2 then implies that $k'(g^w(\bar\theta,\ v)) \to \beta/\hat\beta$ as $v \to -\infty$. Moreover, note that, since u is nonpositive, $g^w(\theta,\ v) \geq v/\beta$ for all θ, implying that $\lim_{v \to \bar v} k'(g^w(\bar\theta,\ v)) = -\infty$. Given these limits and the fact that $k'(g^w(\bar\theta,\ \cdot))$ is continuous on $(-\infty,\ \bar v)$, it follows that $\sup_{v \in (-\infty,\bar v)} k'(g^w(\bar\theta,\ v)) < 1$, so there exists a $v_L > -\infty$ such that $g^w(\bar\theta,\ v) > v_L$ for all v. Hence, the sequence generated by the policy functions satisfies $v_t(\theta^{t-1}) \geq v_L$ for $t \geq 1$, implying the limiting condition in theorem 2.

   Part *b*: If utility is bounded below, then $v_t(\theta^{t-1})$ is bounded below and the result follows immediately since the limiting condition (17) in theorem 2 holds.

   Part *c*: Using the optimality conditions (A12) and (A14) from the proof of proposition 2, one can show that

$$\frac{C'(g^u(\bar\theta,\ v))}{C'(g^u(\underline\theta,\ v))} \leq \frac{\bar\theta}{\underline\theta}.$$

With logarithmic utility, this implies that $g^u(\bar\theta,\ v) - g^u(\underline\theta,\ v) \leq \log(\bar\theta/\underline\theta)$. 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

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

In the proof of theorem 2 we established that $\lim_{t \to \infty} \beta^t \mathbb{E}_{-1}[v_t(\theta^{t-1})] = 0$. Since $\lim_{t \to \infty} \beta^t tA = 0$, it follows that $\limsup_{t \to \infty} \beta^t \mathbb{E}_{-1}[v_t(\sigma^{t-1}(\theta^{t-1}))] \geq 0$.

   Part *d*: If $\underline\gamma > 0$, then the bound in (21) implies that $k'(g^w(\bar\theta,\ v)) \leq 1 - \beta/\hat\beta$ and the result follows immediately. If $\bar\gamma < 1$, then we can define $\kappa = 1 - [(1 - \beta/\hat\beta)/(1 - \bar\gamma)]$ and define v_H by $k'(v_H) = \kappa$. Then for all $v \leq v_H$ we have $g^w(\theta,\ v) \leq v$. It follows that the unique ergodic set is bounded above by v_H. We can

now apply the argument in part a so there exists a $v_L > -\infty$ such that $g^w(\overline{\theta}, v) > v_L$.

Proof of Proposition 6

For part a, 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. For interior solutions,

$$1 - k'(v) = \eta \sum_{\theta \in \Theta} C'(g^u(\theta, v)) p(\theta) = \eta \sum_{\theta \in \Theta} G(C(g^u(\theta, v))) p(\theta),$$

where $G(c) \equiv C'(U(c)) = 1/U'(c)$. Since $G(c)$ is convex for $c \geq \overline{c}$, there must exist scalars A, B with $A > 0$ such that $G(c) \geq Ac + B$ for all c. For example, set $A = G'(\overline{c})$ and $B = -G'(\overline{c})\overline{c}$. Then

$$\frac{1}{\eta} = \int \sum_{\theta \in \Theta} G(C(g^u(\theta, v))) p(\theta) d\psi^*(v) \geq A \int \sum_{\theta \in \Theta} C(g^u(\theta, v)) p(\theta) d\psi^*(v) + B,$$

and the result follows.

If utility is bounded below, by proposition 2, corner solutions $C(g^u(\theta, v)) = 0$ occur for $\underline{v} \leq v \leq v^*$ where $k'(v^*) = 1$. For $v > v^*$, solutions are interior $C(g^u(\theta, v)) > 0$ for all $\theta \in \Theta$, and the argument above applies. This establishes that total consumption is finite.

For part b note that $\overline{\gamma} < 1$ implies that the ergodic set for welfare v is bounded away from $\overline{v}$. Then the continuity of the policy function $g^u(\theta, v)$ implies that consumption $C(g^u(\theta, v))$ is bounded by some finite number on the support of ψ^*. Total consumption is then finite.

Proof of Proposition 7

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

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

with equality at steady-state distributions ψ^*. Since $\int k(v; \eta) d\psi(v) + [\eta e/(1 - \hat{\beta})]$ represents the full Lagrangian function (which now includes the omitted term due to e in [12]), by duality (see Luenberger 1969, chap. 8.6),

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

with equality whenever ψ, η, and e are such that at the allocation that attains $k(v, \eta)$ the intertemporal resource constraint (11) holds, which is true at the constructed steady state. Integrating the right-hand side by parts gives

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

It follows that

$$S(\psi; \ e) - \int \ vd\psi(v) - \left[S(\psi^*; \ e) - \int \ vd\psi^*(v)\right]$$

$$\leq S^R(\psi; \ e) - \int \ vd\psi(v) - \left[S^R(\psi^*; \ e) - \int \ vd\psi^*(v)\right]$$

$$\leq \int \ [\psi^*(v) - \psi(v)][k'(v; \ \eta) - 1]dv.$$

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

References

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

Atkeson, Andrew, and Robert E. Lucas Jr. 1992. "On Efficient Distribution with Private Information." *Rev. Econ. Studies* 59 (July): 427–53.

———. 1995. "Efficiency and Equality in a Simple Model of Efficiency Unemployment Insurance." *J. Econ. Theory* 66 (June): 64–88.

Bernheim, B. Douglas. 1989. "Intergenerational Altruism, Dynastic Equilibria and Social Welfare." *Rev. Econ. Studies* 56 (January): 119–28.

Caplin, Andrew, and John Leahy. 2004. "The Social Discount Rate." *J.P.E.* 112 (December): 1257–68.

Farhi, Emmanuel, and Iván Werning. 2006. "Progressive Estate Taxation." Working Paper no. 12600 (October), NBER, Cambridge, MA.

Green, Edward J. 1987. "Lending and the Smoothing of Uninsurable Income." In *Contractual Arrangements for Intertemporal Trade*, edited by Edward C. Prescott and Neil Wallace. Minneapolis: Univ. Minnesota Press.

Grunwald, Gary K., Rob J. Hyndman, Leanna Tedesco, and Richard L. Tweedie. 2000. "Non-Gaussian Conditional Linear AR(1) Models." *Australian and New Zealand J. Statis.* 42 (December): 479–95.

Luenberger, David G. 1969. *Optimization by Vector Space Methods.* New York: Wiley.

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

———. 2006. "Opportunity and Social Mobility." *Rev. Econ. Studies* 73 (April): 487–504.

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

Spear, Stephen E., and Sanjay Srivastava. 1987. "On Repeated Moral Hazard with Discounting." *Rev. Econ. Studies* 54 (October): 599–617.

Stokey, Nancy L., and Robert E. Lucas Jr. 1989. *Recursive Methods in Economic Dynamics.* With Edward C. Prescott. Cambridge, MA: Harvard Univ. Press.

Thomas, Jonathan, and Tim Worrall. 1990. "Income Fluctuation and Asymmetric Information: An Example of a Repeated Principal-Agent Problem." *J. Econ. Theory* 51 (August): 367–90.