Federal Reserve Bank of Minneapolis
Research Department Staff Report 323

November 2003

Opportunity and Social Mobility*

Christopher Phelan

Federal Reserve Bank of Minneapolis

ABSTRACT

This study argues that both unequal opportunity and social mobility are necessary implications of an efficient societal arrangement when incentives must be provided.

*The author thanks Marco Bassetto, V. V. Chari, Narayana Kocherlakota, and Aleh Tsyvinski for helpful comments. The views expressed herein are those of the author and not necessarily those of the Federal Reserve Bank of Minneapolis or the Federal Reserve System.

1. Introduction

The fact that the children of rich parents have better economic prospects than the children of poor parents ("unequal opportunity") is generally thought to be one of the weak points of modern capitalist societies. The ability of the descendants of poor families to eventually become rich and the descendants of rich families to eventually become poor ("social mobility") is commonly considered to be one of the strong points of these societies. Here I argue that both of these characteristics are, in fact, necessary implications of an efficient societal arrangement when incentives to work hard must be provided.

I argue this point using a generational version of the model of Phelan and Townsend (1991), an infinitely repeated, general equilibrium economy with incentive constraints. In my model, each household's stochastic output is a function of its level of effort. Since effort is assumed to be costly and privately observed (households can shirk), higher than minimal effort levels must be induced by making a household's present or future consumption dependent on the household's observed output history.

A large literature considers models similar to this (including Green (1987), Atkeson and Lucas (1992), Phelan (1994, 1995), Wang (1995), Hopenhayn and Nicolini (1997), and Khan and Ravikumar (2002)). The main difference between the economies in that literature and here is that here, instead of a household consisting of a single infinitely lived individual, a household consists of a sequence of altruistically linked individuals (a familial dynasty), each of whom lives for one period.

This difference affects the appropriate societal ranking of allocations. When households consist of a single infinitely lived individual, allocations can be ranked according to their implied distributions of *ex ante* lifetime utilities. (If one allocation delivers a distribution of *ex ante* utilities dominated by another allocation, then the first allocation is inefficient.) However, if a household consists of a sequence of altruistically linked individuals, then the appropriate ranking of allocations is no longer obvious. How those in the first generation rank allocations will, in general, differ from how their descendants rank allocations.

This study addresses this conflict between generations by ranking allocations according to a Rawlsian veil-of-ignorance criterion (Rawls (1971)). That is, here, society seeks to maximize the expected dynastic utility (utility including altruism toward descendants) of an individual who does not know into which generation he will be born and does not know the identity or output levels of his ancestors.[1] I argue that this implies allocations are ranked not by their distributions of *ex ante* utilities, but instead by their implied *limiting* distributions of dynastic utilities. This transforms the social planning problem into one of directly choosing the stationary distribution of dynastic utilities, as well as functions for determining effort levels, consumption, and a child's position in this distribution as a function of his parent's output. This transformed problem is a *static* social choice problem and thus a major simplification.

My first main theorem is that a society using this ranking will always choose the distribution of dynastic utilities to be *nondegenerate;* in other words, it will choose *unequal opportunity.* Some individuals will be born relatively poor (fated to receive relatively low consumption for each output realization), and some will be born relatively rich, even though equal opportunity is feasible. This occurs because it helps with the provision of incentives to make a child poorer if his parent realizes a low output level and richer if his parent realizes

[1] Recent work by Freeman and Sadler (2002) uses a similar objective function. They consider whether an optimal policy can be decentralized through inheritances.

a high output level.

My second main theorem, subject to a condition which can be proved for specific functional forms for utility, is that society will choose to have social mobility. It will never choose to have a *caste system* with one group of families having relatively high average consumption and another having relatively low consumption and with no ability for a family to move between groups. Instead, no matter how poor or rich a parent is, eventually, the expected consumption of that person's descendants equals the unconditional expectation.

The basic intuition behind the first result, unequal opportunity, is that there is zero loss, at the margin, from allowing some dependence of a child's consumption on his parent's output realization. However, a positive marginal benefit results from relaxing the incentive constraints on parents by making a child's future consumption depend on his parent's output realization. The basic intuition behind the second result, social mobility, is that a society with multiple castes simply requires more resources to deliver a given mean utility than a single caste society.

After stating and proving these main theoretical results for a general, additively separable utility function, I discuss specific functional forms for utility and computation. Next I show that computation of the optimal stationary distribution of dynastic utilities (along with the functions determining consumption and mobility) consists of solving a single linear program. Finally, I present a computed example and compare it to the static optimum and the optimal allocation when allocations are ranked by *ex ante* utility, as opposed to limiting, dynastic utility.

At the end, I discuss the generality of these results. I argue that my results do not depend on the particular type of incentive problem discussed here, unobserved effort. My results can apply to unobserved endowment models such as that of Green (1987), unobserved preference shock models such as that of Atkeson and Lucas (1992), and unobserved production models such as that of Khan and Ravikumar (2002).

2. The Basic Model

Consider the following economy. In each time period, $t \in \{0, \ldots, \infty\}$, the economy is populated by a unit mass continuum of identical, infinitely lived households. There is a single consumption good. If a household exerts an effort level $a \in A = \{a_0, \ldots, a_N\}$, then its output (in terms of the consumption good) $q \in \{q_0, \ldots, q_M\}$ occurs with probability $P(q|a)$. Assume that $P(q|a) > 0$ for all $(q, a) \in Q \times A$ and that there exists $(\overline{q}, \underline{q}) \in Q^2$ such that if $a_i < a_j$, then $p(\overline{q}|a_i) < p(\overline{q}|a_j)$ and $p(\underline{q}|a_i) > p(\underline{q}|a_j)$. That is, the probability of one outcome (say, the highest) is increasing in a, and the probability of another outcome (say, the lowest) is decreasing in a. Households are assumed to be able to privately exert effort less than that specified by the allocation. That is, they can shirk.

Household utility in period t is determined by the function $U(c_t, a_t) = u(c_t) - v(a_t)$, where c_t is the household's period t consumption. The function u is assumed to be twice differentiable with $u' > 0$ and $u'' < 0$. For $a_i < a_j$, $v(a_i)$ is assumed to be strictly less than $v(a_j)$. Over time and uncertainty, a household cares about the expected value of $(1 - \beta) \sum_{t=0}^{\infty} \beta^t U(c_t, a_t)$, where $\beta \in (0, 1)$. Let V denote the set of feasible lifetime utilities.

Define the *efficient symmetric static allocation* $\big(a^*, c^*(q)\big)$ as the solution to

$$\max_{a, c(q)} \sum_q P(q|a) u\big(c(q)\big) - v(a)$$

subject to a static resource constraint

$$\sum_q P(q|a)\big(c(q) - q\big) \leq 0$$

and the static incentive constraint (for $\hat{a} < a$)

$$\sum_q P(q|a)u\big(c(q)\big) - v(a) \geq \sum_q P(q|\hat{a})u\big(c(q)\big) - v(\hat{a}).$$

Assume that $a^* > a_0$.

A. Feasible Allocations

Let a *dynamic allocation* (or simply an *allocation*) $\big(\Psi_0, \{a_t(w_t), c_t(w_t, q_t), w_{t+1}(w_t, q_t)\}_{t=0}^{\infty}\big)$ be defined recursively as a measure of initial lifetime utilities, Ψ_0, and a sequence of functions $a_t(w_t)$, $c_t(w_t, q_t)$, and $w_{t+1}(w_t, q_t)$. The function $a_t(w_t)$ specifies the recommended effort level for a household which starts period t with a continuation expected utility of w_t. The function $c_t(w_t, q_t)$ specifies the nonnegative consumption of a household which starts period t with a continuation expected utility of w_t and realizes output q_t. The function $w_{t+1}(w_t, q_t)$ specifies the continuation expected discounted utility at the beginning of period $t+1$ of a household which starts period t with a continuation expected utility of w_t and realizes output q_t.

Note that through the initial distribution of forward-looking utilities Ψ_0, and the functions $a_t(w_t)$ and $w_{t+1}(w_t, q_t)$, an allocation determines, for all $t \geq 1$, the period t distribution of forward-looking utilities Ψ_t.

An allocation is said to satisfy *promise-keeping* if, for all t and w_t,

$$(1) \quad w_t = \sum_{q_t} P\big(q_t | a_t(w_t)\big)\Big((1-\beta)\big[u\big(c_t(w_t, q_t)\big) - v(a_t(w_t))\big] + \beta w_{t+1}(w_t, q_t)\Big).$$

In words, promise-keeping requires that expected dynastic utility of an allocation conditional on w_t is actually w_t. Next, an allocation is said to be *incentive-compatible* if, for all t, w_t, and $\hat{a} < a_t(w_t)$,

$$(2) \quad w_t \geq \sum_{q_t} P\big(q_t | \hat{a}\big)\Big((1-\beta)\big[u\big(c_t(w_t, q_t)\big) - v(a_t(w_t))\big] + \beta w_{t+1}(w_t, q_t)\Big).$$

Here, the left side is the dynastic utility associated with taking action $a_t(w_t)$ and the right side is the dynastic utility associated with taking an alternative action $\hat{a} < a_t(w_t)$. Finally, an allocation is said to be *resource-feasible* if, for all t,

$$(3) \quad 0 \geq \int_V \sum_q P\big(q | a_t(w_t)\big)[c(w_t, q) - q] \; d\Psi_t(w_t).$$

Condition (3) requires that aggregate production be weakly greater than aggregate consumption. An allocation is considered *feasible* if it satisfies all three of these conditions (1)–(3).

B. Ranking Feasible Allocations

In most dynamic contracting work, a household consists of a single infinitely lived individual who discounts the future by β.[2] Given this, an allocation is considered *efficient* if it is feasible (satisfies promise-keeping, incentive-compatibility, and resource-feasibility) and

[2]See Green (1987), Phelan and Townsend (1991), and Atkeson and Lucas (1992) among many others. An exception is Freeman and Sadler (2002).

if no other feasible allocation delivers a distribution of initial utilities which dominates Ψ_0.

In models similar to this one, Atkeson and Lucas (1992) and Phelan (1994) derive several implications of this type of efficiency. First, these studies show that an efficient allocation must, household by household, minimize the discounted resource cost of delivering a given *ex ante* utility w_0 and that this resource cost is a convex function of w_0. Thus, a society maximizing mean *ex ante* utility would choose a degenerate measure of initial utilities Ψ_0 with all mass on the same point.[3] Second, these studies show that efficiency, by this definition, implies extreme results regarding the limiting distribution of consumption and utility. In the model of Atkeson and Lucas (1992), almost all consumption paths go to zero, and mean utility goes to the lower bound of the set of possible utilities (either zero or negative infinity, depending on the level of risk aversion). In the model of Phelan (1994), the variance of consumption grows without bound, and thus mean utility becomes infinitely negative.

Here I consider a different ranking of allocations, supported by the following assumption. Instead of a household consisting of a single infinitely lived individual, suppose that a household consists of a sequence of altruistically linked individuals, each of whom lives for one period. Specifically, assume that the dynastic utility of an individual born in period t consists of weight $1-\beta$ on his own direct utility $U(c_t, a_t)$ and weight β on the dynastic utility of his single child. Thus, his dynastic utility is

$$w_t = \sum_{q_t} P\big(q_t|a_t(w_t)\big)\big[(1-\beta)U\big(c_t(w_t,q_t),a_t(w_t)\big) + \beta w_{t+1}(w_t,q_t)\big].$$

[3]Ranking allocations by mean *ex ante* utility is equivalent to maximizing the utility of a household which does not know where in distribution Ψ_0 it will be, but instead sees itself as having the same probability as all other households of being in any subset of the support of Ψ_0.

(This is equivalent to the individual putting weight $1-\beta$ on his own direct utility, weight $\beta(1-\beta)$ on his child's direct utility, $\beta^2(1-\beta)$ on his grandchild's, and so on.) With this composition of households, the set of feasible allocations is identical to that which holds if households consist of a single infinitely lived individual. However, the appropriate ranking of allocations is no longer obvious. Maximizing *ex ante* dynastic utility puts no direct weight on the utility of generations born later than period $t = 0$. These later generations enter the social calculus only indirectly, through the altruism of those born in period $t = 0$.

Maximizing *ex ante* dynastic utility is equivalent to maximizing the expected utility of individuals who know they will be born in period $t = 0$. Alternatively, the expected dynastic utility of individuals who do not know into which generation they will be born can be maximized. Formally, let $\overline{v}_t = \int_V w_t \, d\Psi_t(w_t)$ and $\overline{v} = \lim_{T\to\infty} \frac{1}{T+1} \sum_{t=0}^{T} \overline{v}_t$. In words, $\overline{v}_t$ is the expected dynastic utility, including altruism toward children, of individuals who know only that they will be born in period t, and $\overline{v}$ is the limit of the means of these period t dynastic utilities. Instead of ranking allocations by v_0 (which puts only indirect weight on later generations) allocations can be ranked by $\overline{v}$. Since the number of periods is infinite, such a weighting scheme puts zero weight on the first T periods (regardless of T). Further, for allocations with defined limiting distributions of dynastic utility, this weighting scheme effectively ranks the allocations according to the mean of this limiting distribution. To this end, attention is from here on restricted to allocations where a limiting distribution of dynastic utility exists. Thus, simply choosing this limiting distribution directly results in a much simpler, *static* planner's problem.

C. Limiting Allocations

Rather than the limiting distribution Ψ being an implication of a dynamic allocation $(\Psi_0, \{a_t(w_t), c_t(w_t, q_t), w_{t+1}(w_t, q_t)\}_{t=0}^{\infty})$, one can transform society's problem to one where Ψ is instead a choice variable.

Let a *limiting allocation* be defined as a collection $\xi = (\Psi, a(w), c(w, q), w'(w, q))$ such that

- Ψ is the distribution of dynastic utilities, a probability measure mapping subsets of $V \to [0, 1]$;
- $a(w)$ is the recommended action as a function of w, or $a(w) : V \to A$;
- $c(w, q)$ determines the household's consumption as a function of w and q, or $c(w, q)$: $V \times Q \to {}_{+}$; and
- $w'(w, q)$ is the transition function for dynastic utilities as a function of w and q, or $w'(w, q) : V \times Q \to V$.

Let $\mathcal{B}(V)$ denote the Borel subsets of the real line. A limiting allocation is considered *stationary* if for all subsets $W \in \mathcal{B}(V)$,

$$(4) \qquad \Psi(W) = \int_V \sum_q P\big(q|a(w)\big)\; I\big(w'(w, q) \in W\big)\; d\Psi(w),$$

where $I(\cdot)$ is the indicator function. Here, the left side is the mass of households on set W today, and the right side is the mass of households on set W tomorrow.

A limiting allocation satisfies *promise-keeping* if for almost all w relative to Ψ,

$$(5) \quad w = \sum_q P(q|a(w))\Big((1-\beta)\big[u\big(c(w,q)\big) - v\big(a(w)\big)\big] + \beta w'(w,q)\Big).$$

In words, condition (5) requires that the functions $a(w)$, $c(w,q)$, and $w'(w,q)$ actually deliver dynastic utility w.

A limiting allocation is considered *incentive-compatible* if for almost all w relative to Ψ and all $\hat{a} < a(w)$,

$$(6) \quad w \geq \sum_q P(q|\hat{a})\Big((1-\beta)\big[u(c(w,q)\big) - v(\hat{a})\big] + \beta w'(w,q)\Big).$$

As before, the left side is the dynastic utility of taking action $a(w)$, and the right side is the dynastic utility of taking alternative action $\hat{a}$.

Finally, a limiting allocation satisfies *resource-feasibility* if

$$(7) \quad 0 \geq \int_V \sum_q P\big(q|a(w)\big)[c(w,q) - q] \; d\Psi(w).$$

Recall that attention is restricted to limiting allocations precisely because society ranks allocations by the mean dynastic utility of the limiting allocation. Given this, a limiting allocation is considered *optimal* if it solves

$$(8) \quad \max_{\Psi, c(w,q), w'(w,q)} \int_V w \; d\Psi(w)$$

subject to (4)–(7). This constrained maximization problem will be referred to as the *primal* *problem.*

3. Characterizing Optimal Limiting Allocations

This section presents my main result: unequal opportunity and social mobility are necessarily characteristics of any optimal limiting allocation. Showing this requires two supporting lemmas. One is that resource feasibility holds with equality in any optimal limiting allocation:

Lemma 1. If a limiting allocation $\xi^* = \big(\Psi^*, a^*(w), c^*(w,q), w'^*(w,q)\big)$ is optimal, then $\int_V \sum_q P\big(q|a(w)\big)[c(w,q) - q]\ d\Psi(w) = 0$.

Proof. Consider an optimal limiting allocation ξ^* such that

$$\int_V \sum_q P\big(q|a^*(w)\big)\left[c^*(w,q) - q\right] d\Psi^*(w) < 0.$$

For a given $\epsilon > 0$, construct an alternative allocation ξ as follows. First, for all intervals $(-\infty, w]$, let $\Psi\big((-\infty, w+\epsilon]\big) = \Psi^*\big((-\infty, w]\big)$. This ensures that the objective function increases by ϵ. Next, let $a(w+\epsilon) = a^*(w)$. This ensures that aggregate production is unchanged. Finally, let $w'(w+\epsilon, q) = w'^*(w,q) + \epsilon$, $c(w+\epsilon, q) = c^*(w,q)$ if $q \neq \overline{q}$, and $c(w+\epsilon, \overline{q})$ be such that

$$u\big(c(w+\epsilon, \overline{q})\big) = u\big(c^*(w, \overline{q})\big) + \frac{\epsilon}{P\big(\overline{q}|a^*(w)\big)}.$$

In words, a household promised w utils under allocation ξ^* is delivered $w + \epsilon$ utils under allocation ξ by increasing all continuation utility promises by ϵ and increasing the utility payment if output $\overline{q}$ occurs by $\epsilon/P\big(\overline{q}|a^*(w)\big)$. This construction ensures that stationarity (4) and promise-keeping (5) are maintained.

This leaves the incentive-compatibility constraints (6) and the resource-feasibility condition (7) to be met. For a given utility point $w+\epsilon$ and $\hat{a}<a$, the incentive constraint for allocation ξ is that

$$w+\epsilon \geq \sum_q P(q|\hat{a})\Big((1-\beta)\big[u\big(c(w+\epsilon,q)\big)-v(\hat{a})\big]+\beta w'(w+\epsilon,q)\Big). \tag{9}$$

From the fact that ξ^* is incentive-compatible, we know that

$$w=\sum_q P(q|\hat{a})\Big((1-\beta)\big[u\big(c^*(w,q)\big)-v(\hat{a})\big]+\beta w'^*(w,q)\Big)+\Delta, \tag{10}$$

where $\Delta \geq 0$ is the amount by which the incentive constraint is slack. Subtracting, side by side, expression (10) from expression (9) and using the definition of ξ delivers that

$$\epsilon \geq (1-\beta)\frac{P(\overline{q}|\hat{a})}{P\big(\overline{q}|a^*(w)\big)}\epsilon+\beta\epsilon-\Delta. \tag{11}$$

This holds because $P(\overline{q}|\hat{a})/P\big(\overline{q}|a^*(w)\big)<1$ and $\Delta \geq 0$. Thus, ξ is incentive-compatible.

Finally, the fact that $\int_V \sum_q P\big(q|a^*(w)\big)[c^*(w,q)-q]d\Psi^*(w)<0$ implies that there exists an $\epsilon>0$ for which equation (7) is satisfied, contradicting the optimality of ξ^*. ■

Given that the resource constraint (7) binds, it is straightforward to show that a plan which maximizes mean utility minimizes the cost of providing any given mean utility. Thus, the second supporting lemma is the following.

LEMMA 2. Suppose a limiting allocation ξ^* is optimal, and let $v=\int_V w\ d\Psi^*(w)$. Then ξ^*

solves the *dual problem*

$$(12)\quad C(v) = \min_{\Psi, a(w), c(w,q), w'(w,q)} \int_V \sum_q P\big(q|a(w)\big)[c(w,q) - q]\; d\Psi(w)$$

subject to (4)–(6), and

$$(13)\quad v \leq \int_V w\; d\Psi(w).$$

Proof. The stationary allocation ξ^* satisfies (4)–(6) immediately since it is optimal and thus in the constraint set of the primal problem. It satisfies (13) by the definition of v. Suppose another stationary allocation ξ satisfying (4)–(6) and (13) has a lower value of the dual objective function (12). Such a plan is in the constraint set of the primal problem since it satisfies (4)–(6) immediately and satisfies (7) with slack from the fact that the dual has a lower objective function value than ξ^*. Stationary allocation ξ has a weakly higher primal objective function value (since it satisfies (13)) and the resource constraint (7) does not bind, contradicting Lemma 1. ■

A. Opportunity

Lemma 2 allows for the first main result, that an optimal plan will always exhibit unequal opportunity. (Some individuals are born with lower expected dynastic utility than others.) The general strategy of the proof is to assume that all incentives are static—that all individuals are born with a blank slate—and show that the cost of introducing a small amount of dependency of children's consumption on parents' outcomes is second-order, while the benefit, or gain, from this dependency (which allows for the better provision of incentives to parents) is first-order.

THEOREM 1. *Let a limiting allocation* $\xi^* = \big(\Psi^*, a^*(w), c^*(w,q), w'^*(w,q)\big)$ *be optimal. Then* Ψ^* *is nondegenerate.*

Proof. The strategy of this proof is similar to that in Rogerson (1985): Suppose no links across periods and show there exists an improving perturbation. (The model of this paper is sufficiently different from the model in Rogerson (1985) that this proof needs a different perturbation than that in Rogerson (1985).) To this end, suppose Ψ^* is degenerate with all mass on point w^*. Let $a^* = a^*(w^*)$. Define an alternative allocation $\xi = \big(\Psi, a(w), c(w,q), w'(w,q)\big)$. First, let Ψ put mass $1 - P(\underline{q}|a^*) - P(\overline{q}|a^*)$ on point w^*, mass $P(\underline{q}|a^*)$ on point $\underline{w} = w^* - \epsilon/P(\underline{q}|a^*)$, and mass $P(\overline{q}|a^*)$ on point $\overline{w} = w^* + \epsilon/P(\overline{q}|a^*)$. By construction, then $\int_V w \; d\Psi(w) = \int_V w \; d\Psi^*(w) = w^*$; thus, ξ satisfies condition (13) for $v = w^*$. Next, assume for $w \in \{\underline{w}, w^*, \overline{w}\}$ that $a(w) = a^*$. This ensures that aggregate production is unchanged.

Next, let $w'(w,\underline{q}) = \underline{w}$, $w'(w,\overline{q}) = \overline{w}$, and for $q \notin \{\underline{q}, \overline{q}\}$, $w'(w,q) = w^*$. This ensures (for all ϵ) that stationarity (4) is satisfied. Lastly, define the functions $c(w,q)$ for $(w,q) \in \{\underline{w}, w^*, \overline{w}\} \times Q$. To do this, for all q, let $c(w^*,q)$ be such that $u\big(c(w^*,q)\big) = u\big(c^*(w^*,q)\big) + \Delta(w^*,q)$. Next, let $c(\underline{w},q)$ be such that $u\big(c(\underline{w},q)\big) = u\big(c^*(w^*,q)\big) - \epsilon/[(1-\beta)P(\underline{q}|a^*)] + \Delta(\underline{w},q)$. Finally, let $c(\overline{w},q)$ be such that $u\big(c(\overline{w},q)\big) = u\big(c^*(w^*,q)\big) + \epsilon/[(1-\beta)P(\overline{q}|a^*)] + \Delta(\overline{w},q)$. Since the original limiting allocation ξ^* is optimal, choosing $\epsilon = 0$ and $\Delta(w,q) = 0$ for all $(w,q) \in \{\underline{w}, w^*, \overline{w}\} \times Q$ must minimize equation (12) subject to the promise-keeping constraint (5) and the incentive-compatibility constraint (6).

Note that the incentive constraint associated with $w = w^*$ in this restricted optimization problem binds; thus, the marginal value of loosening it is strictly positive. To see this, consider choosing ϵ and $\Delta(w, q)$ to minimize (12) subject to the promise-keeping constraint (5) but not the incentive constraint (6). Here, I can strictly improve on ξ^* by setting $\epsilon = 0$ and setting $\Delta(w^*, q)$ such that $c(w^*, q) = \sum_q P(q|a^*)c^*(w^*, q)$ (full consumption insurance) less a constant to compensate for the utility gain associated with full insurance. (Recall the assumption that $a^* > a_0$. Thus, full consumption insurance is not attained by ξ^*.)

Now set $\Delta(w, q) = 0$ for all (w, q). This ensures that the promise-keeping constraint (5) holds for all ϵ. Thus, a choice of $\epsilon \neq 0$ affects only the incentive constraint (6) and the dual objective function (12).

For $w \in \{\underline{w}, w^*, \overline{w}\}$, the derivative, with respect to ϵ, of the left side of the incentive constraint minus the derivative of the right side equals $\beta[P(\overline{q}|\hat{a})/P(\overline{q}|a^*) - P(\underline{q}|\hat{a})/P(\underline{q}|a^*)]$. This derivative is a strictly negative constant (not a function of ϵ) for all $\hat{a} < a^*$, and thus, the incentive constraint for each w is loosened as ϵ increases.

Finally, let $u^{-1}(u)$ denote the consumption payment necessary to deliver utility $u(c)$. The dual objective function with the definition of ξ and $\Delta(w, q) = 0$ substituted in is then

$$
\begin{aligned}
&P(\underline{q}|a^*) \sum_q P(q|a^*) \Big[u^{-1}\big(u(c^*(w^*, q)) - \frac{\epsilon}{P(\underline{q}|a^*)}\big) - q\Big] \\
&+P(\overline{q}|a^*) \sum_q P(q|a^*) \Big[u^{-1}\big(u(c^*(w^*, q)) + \frac{\epsilon}{P(\overline{q}|a^*)}\big) - q\Big] \\
&+[1 - P(\underline{q}|a^*) - P(\overline{q}|a^*)] \sum_q P(q|a^*) \Big[u^{-1}\big(u(c^*(w^*, q))\big) - q\Big].
\end{aligned}
$$

The derivative of this expression with respect to ϵ is

$$
\begin{aligned}
&-\sum_q P(q|a^*)u^{-1'}\big(u(c^*(w^*,q)) - \frac{\epsilon}{P(\underline{q}|a^*)}\big) \\
&+\sum_q P(q|a^*)u^{-1'}\big(u(c^*(w^*,q)) + \frac{\epsilon}{P(\overline{q}|a^*)}\big).
\end{aligned}
$$

This derivative equals zero for $\epsilon = 0$. Thus, the marginal value of increasing ϵ when $\epsilon = 0$ and $\Delta(w, q) = 0$ is strictly positive since it loosens a binding constraint (a first-order benefit) with zero first-order effect on the objective function, contradicting the optimality of the original allocation. ■

B. Social Mobility

Note that an allocation $\xi = \big(\Psi, a(w), c(w, q), w'(w, q)\big)$ defines not only the distribution of dynastic utilities, Ψ, but also the rules under which households move up or down this distribution. Thus, the answers to questions regarding social mobility are embedded in ξ. (Can the descendents of poor, or low w, households eventually become rich?) Now I consider to what extent the efficiency of ξ implies social mobility. In particular, I argue that social mobility is a direct implication of strict convexity of the cost function $C(v)$.

To allow a strict definition of social mobility, let a set $W \in \mathcal{B}(V)$ be called a *caste* under limiting allocation ξ if $\Psi(W) > 0$, and

$$
(14) \quad \Psi(W) = \int_W \sum_q P\big(q|a(w)\big) I\big(w'(w,q) \in W\big) \; d\Psi(w),
$$

which implies a zero exit and entry probability from W. A caste W is called *trivial* (relative to ξ) if $\Psi(W) = 1$, or if $\int_W w \; d\Psi(w) = \int_{W^c} w \; d\Psi(w)$, where W^c denotes the complement of

W. Thus, for a caste system to be nontrivial, its complement must have positive mass and the mean utility of those in the caste must differ from the mean utility of those outside the caste.

My main theorem here is that if $C(v)$ is strictly convex (a condition shown in the next section for particular functional forms), then any caste system must be trivial. That is, having a permanently richer group and a permanently poorer group is never optimal.

THEOREM 2. *Suppose $C(v)$ is strictly convex and a limiting allocation $\xi = \big(\Psi, a(w), c(w,q), w'(w,q)\big)$ minimizes* (12) *subject to* (4)–(6), *and* (13) *for $v = \int w \ \Psi w(w)$. Then any caste W relative to ξ is trivial.*

Proof. Let W_1 be a caste relative to ξ. If $\Psi(W_1) = 1$, then the result is proved; thus, assume that $\Psi(W_1) < 1$. Let $W_2 = W_1^c$. Like W_1, the set W_2 is also a caste. Define two separate allocations ξ_i, $i \in \{1, 2\}$, by choosing Ψ_i such that for all $W \subset W_i$, $\Psi_i(W) = \Psi(W)/\Psi(W_i)$ and leaving the functions $a(w)$, $c(w,q)$, and $w'(w,q)$ unaltered. (That is, proportionally put all mass on one set or the other, but otherwise change nothing). These allocations each satisfy promise-keeping and incentive-compatibility since the original allocation satisfies these conditions, and each satisfies stationarity since W_1 and W_2 do not communicate and the original allocation satisfied stationarity. Put less formally, the fact that the sets W_1 and W_2 don't communicate implies that how those in each set are treated defines a feasible allocation for treating all of society. Thus, each allocation must minimize (12) subject to (4)–(6), and (13) for $v = w_i$, where $w_i = [1/\Psi(W_i)] \int_V w d\Psi_i(w)$, $i \in \{1, 2\}$. If another allocation satisfies (4)–(6) and (13) at a lower cost, then the original allocation ξ could not have been optimal, since this lower cost allocation could have been incorporated into the original allocation, lowering its cost. Finally, if $w_1 \neq w_2$, then the strict convexity of C implies that $C\big(\Psi(W_1)w_1 + \Psi(W_2)w_2\big) < \Psi(W_1)C(w_1) + \Psi(W_2)C(w_2)$. Since the right side of this inequality is the resource cost of the original plan, $w_1 = w_2$. ■

4. Functional Forms

Here I introduce two explicit functional forms for $U(c, a)$. These functional forms allow me to solve for $C(v)$ (up to a constant) and thus prove the strict convexity assumed by Theorem 2. Further, they allow for a relatively complete characterization of the optimal allocation when allocations are ranked by *ex ante* utility, as opposed to limiting, mean utility; they thus help highlight the effect of ranking allocations by the mean of the limiting distribution of dynastic utilities. While these examples are not additively separable (as the earlier sections assumed), Lemmas 1 and 2 and Theorem 1 can be proved using arguments similar to those used earlier.

The simplest example has $U(c, a) = -\exp(-\gamma[c - v(a)])$ with consumption unbounded (or $c \in \mathbb{R}$) and $\gamma > 0$—the constant absolute risk aversion (CARA) utility specification in Phelan (1994). Given this utility function and consumption set, Phelan (1994) shows that if allocations are ranked by mean *ex ante* utility, optimality implies that a_t is constant across households and time (thus, so is aggregate production), and household consumption is the sum of an independent and identically distributed random variable and a term which follows a driftless random walk. Since effort is constant and utility is a concave function of consumption, as the cross-sectional variance of consumption increases due to the random walk term, mean dynastic utility decreases over time without bound. In essence, the optimal allocation from an *ex ante* perspective implies a limiting distribution of dynastic utilities which has all mass on negative infinity. However, that's the worst possible allocation when allocations are ranked, as they are here, by the mean of the limiting distribution of dynastic utilities. (Theorem 1 shows that a finite mean limiting utility can, in fact, be achieved, since one can do better than repeating the static optimum, which itself has a finite mean utility.)

Assuming this specific functional form also allows me to show that $C(v)$ is strictly convex, as assumed by Theorem 2. This is shown in the following lemma:

LEMMA 3. If $U(c,a) = -\exp(-\gamma[c - v(a)])$ with $c \in \mathbb{R}$ and $\gamma > 0$, then for $v < 0$,

$$(15) \quad C(v) = \min_{\Psi, a(w), c(w,q), w'(w,q)} \int_V \sum_q P\big(q|a(w)\big)[c(w,q) - q] \; d\Psi(w)$$

subject to (4)–(6), and (13) satisfies $C(v) = -\log(-v)/\gamma + C(-1)$ and is thus strictly convex.

Proof. See the Appendix. ■

With some extension of the model, I can construct tractable example economies which do not depend on CARA utility. In particular, following Atkeson and Lucas (1992), Khan and Ravikumar (2002), and Phelan (2002), let household output equal kq, where k is the quantity of land allocated to the household for use in production that period. Let $v(a)$ denote the per-acre utility loss to effort, and thus, let $kv(a)$ denote the total utility loss to effort a. Finally, let utility be the constant relative risk aversion (CRRA) specification $U(c,a,k) = [c - kv(a)]^\gamma/\gamma$, where $\gamma = 0$ implies that $U(c,a,k) = \log[c - kv(a)]$. With this specification, if allocations are ranked by *ex ante* utility, almost all household consumption paths converge to zero, the result of Atkeson and Lucas (1992) for a preference shock model. This implies that the limiting distribution of dynastic utilities has either all mass on zero (for the case of $\gamma > 0$) or all mass on negative infinity (for the case of $\gamma \leq 0$). Here, as in the previous example, if allocations are ranked by the mean of the limiting distribution, analogs of Lemmas 1 and 2 and Theorem 1 can be proved.

Introducing land to the model introduces an additional resource constraint into the primal problem. Not only must society not allocate more of the consumption good than is available from production, it must also not allocate more land to households than is exogenously given. However, if society is assumed to be able to trade land for the consumption good at a linear price p (which can be set equal to the ratio of the Lagrange multipliers associated with the separate resource constraints), then I can prove that $C(v)$ is convex. This is shown in the following lemma.

LEMMA 4. If $U(c,a,k) = [c - kv(a)]^\gamma/\gamma$ with $c \geq kv(a)$, then

$$(16) \quad C(v) = \min_{\Psi,k(w),a(w),c(w,q),w'(w,q)} \int_V \left(pk(w) + \sum_q P\big(q|a(w)\big)[c(w,q) - q] \right) d\Psi(w)$$

subject to (4)–(6), and (13) satisfies $C(v) = v^{1/\gamma}C(1)$ if $\gamma > 0$ (and thus $v > 0$), $C(v) = \exp(v)C(0)$ if $\gamma = 0$, and $C(v) = (-v)^{1/\gamma}C(-1)$ if $\gamma < 0$ (and, thus, $v < 0$). In each case, $C(v)$ is strictly convex.

Proof. See the Appendix. ■

5. Computation

My approach of directly choosing the limiting allocation simplifies computation as well. In particular, if Ψ is restricted to a finite support, then an optimal limiting allocation can be computed as a *single* linear program along the lines of Prescott and Townsend (1984).

While the incentive constraints capture the dynamic decision of households, the choice of how to organize society given those constraints is static; thus, dynamic programming along the lines of Phelan and Townsend (1991) or Atkeson and Lucas (1992) is not needed.

The linear program is set up as follows. Let $\hat{V} \subset V$ be a finite grid of points in V restricting the support of Ψ. Next, let $\hat{C}$ be a finite grid of points restricting the range of $c(w, q)$. (The function $a(w)$ has already been assumed to have a finite range.) The key to transforming the choice of the limiting allocation into a linear program is to combine the measure Ψ (now restricted to a finite support) with the rules $a(w)$, $c(w, q)$, and $w'(w, q)$. That is, let $\mu(w, a, q, c, w')$ be the fraction of households who start at point $w \in \hat{V}$, receive action recommendation $a \in A$, experience output realization $q \in Q$, get consumption level $c \in \hat{C}$, and transit to point $w' \subset \hat{V}$.

Choosing $\mu(w, a, q, c, w')$ for all $(w, a, q, c, w') \in \hat{V} \times A \times Q \times \hat{C} \times \hat{V}$ pins down $(\Psi, a(w), c(w, q), w'(w, q))$ if $\mu(w, a, q, c, w')$ satisfies several linear conditions. First, the fractions $\mu(w, a, q, c, w')$ must add to one, or

$$(17) \quad \sum_{w,a,q,c,w} \mu(w, a, q, c, w') = 1.$$

Next, the fraction of households which realize output q must coincide with the fraction determined by the technology $P(q|a)$. This can be enforced by requiring that Bayes' rule holds for all $(\bar{w}, \bar{a}, \bar{q}) \in \hat{V} \times A \times Q$:

$$(18) \quad \sum_{c,w'} \mu(\bar{w}, \bar{a}, \bar{q}, c, w') = \pi(\bar{q}|\bar{a}) \sum_{q,c,w'} \mu(\bar{w}, \bar{a}, q, c, w').$$

The objective function, stationarity, the promise-keeping, incentive-compatibility, and resource-feasibility constraints are, like the previous conditions, linear in the choice variables. The objective function becomes

$$(19) \quad \sum_{w,a,q,c,w'} \mu(w,a,q,c,w')w.$$

A collection $\mu(w,a,q,c,w')$ is stationary if for all $w \in \hat{V}$, the fraction of households at w is the same today and tomorrow, or if for all $\bar{w} \in \hat{V}$,

$$(20) \quad \sum_{a,q,c,w'} \mu(\bar{w},a,q,c,w') = \sum_{w,a,q,c} \mu(w,a,q,c,\bar{w}).$$

Promise-keeping requires that, for all $w \in \hat{V}$,

$$(21) \quad \sum_{a,q,c,w'} \mu(w,a,q,c,w')[(1-\beta)u(c,a) + \beta w' - w] = 0.$$

Incentive-compatibility requires that, for all $(w, a, \hat{a} < a)$,

$$(22) \quad \begin{aligned} &\sum_{q,c,w'} \mu(w,a,q,c,w')[(1-\beta)u(c,a) + \beta w'] \geq \\ &\sum_{q,c,w'} \mu(w,a,q,c,w')[(1-\beta)u(c,\hat{a}) + \beta w'] \frac{P(q|\hat{a})}{P(q|a)}. \end{aligned}$$

Finally, the resource-feasibility constraint is satisfied if and only if

$$(23) \quad \sum_{w,a,q,c,w'} \mu(w,a,q,c,w')(c-q) \leq 0.$$

6. An Example

Now I present an example economy computed using the methods just outlined. For this example economy, I also compute the solutions to the static optimum and the optimum when allocations are ranked by *ex ante* dynastic utility and compare those solutions to that when allocations are ranked by the mean of Ψ, the limiting distribution of dynastic utilities.

In the example, these are the parameter values used: $a \in \{0, 1\}$, $q \in \{0, 1\}$, $\beta = 2/3$, and $U(c, a) = -\exp\big(-(c - 0.3a)\big)$. The high output ($q = 1$) occurs with probability 3/4 if $a = 1$ and probability 1/4 if $a = 0$.[4]

Figure 1 displays Ψ, as well as the utility levels associated with the static optimum, the mean of Ψ, and the optimum when allocations are ranked by *ex ante* dynastic utility.

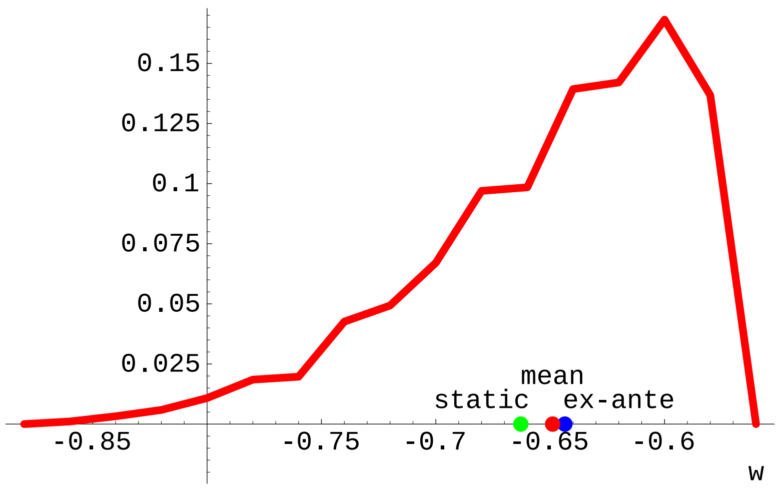

Figure 1: The Distribution of Dynastic Utilities, Ψ.

Since the static optimum is a feasible but suboptimal limiting allocation, its value is strictly lower than the mean of Ψ. Since the optimal limiting allocation is a feasible allocation

[4]The parameters specific to the computation method are as follows: $\hat{V} = \{-0.90, -0.88, \ldots, -0.52, -0.50\}$ and $\hat{C} = \{-0.20, -0.18, \ldots, 1.02, 1.04\}$. The program was written in C using the gnu compiler and the gnu linear programming package. While the resulting linear program has 111,132 variables and 106 constraints, it solves on an Apple 867MHz PowerBook G4 in under four minutes.

when stationarity is not imposed, the mean of Ψ is strictly lower than the utility associated with the optimal *ex ante* allocation.

For all w in the support of Ψ, $a(w) = 1$. (Thus, this function is not graphed.) Figure 2 presents the function $c(w, q)$. Not surprisingly, $c(w, q)$ is increasing in both arguments. Further, for a household receiving the dynastic utility associated with the static optimum, $c(w, q)$ provides less dependence of consumption on current output q than does the static plan. For a household receiving the dynastic utility associated with the *ex ante* optimum, $c(w, q)$ provides more dependence of consumption on current output than does the *ex ante* optimum.

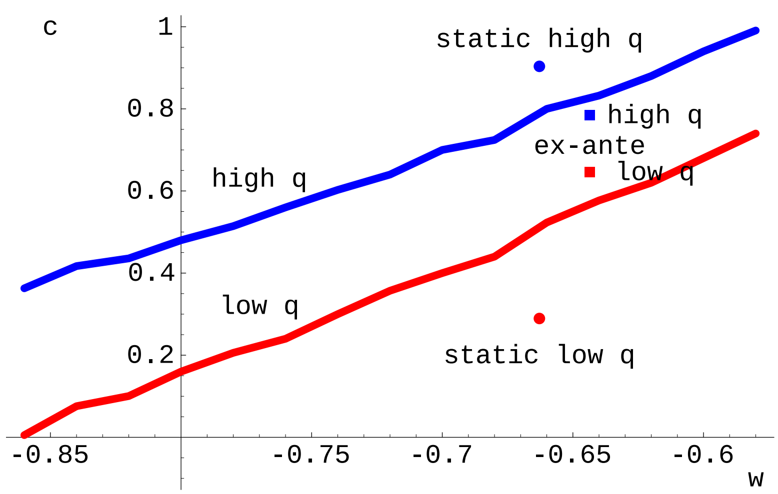

Figure 2: The Consumption Function, $c(w, q)$.

Neither of these characteristics is surprising. That the optimal limiting allocation has less dependence of current consumption on current output comes entirely from the fact that in the static optimum, all incentives must be provided through such dependence, while the optimal limiting allocation allows for incentives to be provided through the function $w'(w, q)$ as well. That the optimal limiting allocation has more dependence of current consumption on current output than does the *ex ante* optimal plan comes from the fact that future generations matter more to society when ranking allocations by mean limiting utility than by *ex ante* utility. Having a household's consumption depend on its ancestors' output costs society because utility is a convex function of consumption. However, such a dependence helps relax the incentive constraints on parents (Theorem 1). The more future generations enter the objective function of society, however, the costlier such intergenerational dependence, and, thus, the less this manner of providing incentives is used.

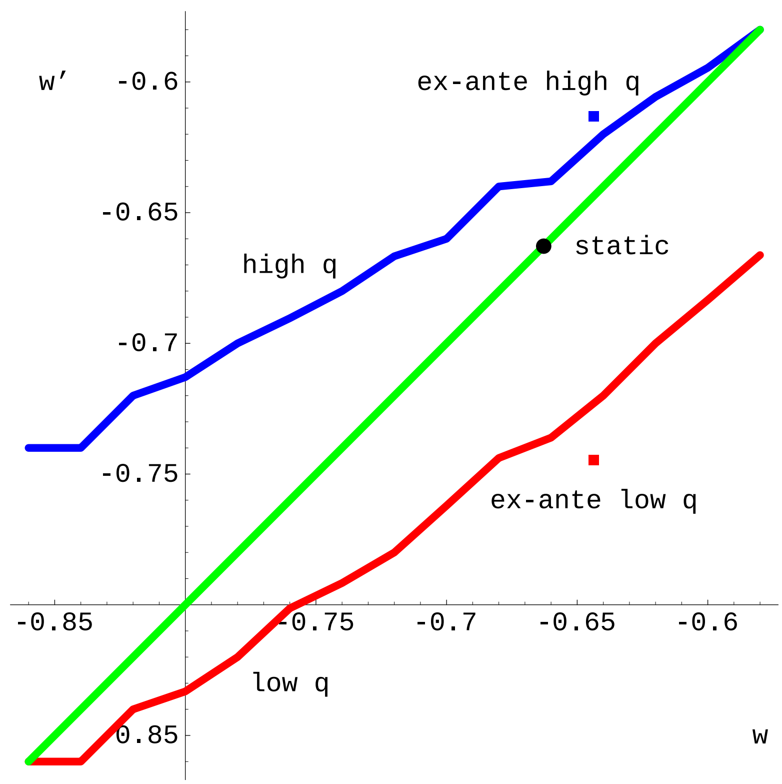

Figure 3: The Utility Transition Function, $w'(w, q)$.

Figure 3 presents the transition function $w'(w, q)$. Like the function $c(w, q)$, the function $w'(w, q)$ is increasing in both arguments. The transition function $w'(w, q)$ provides more dependence of future dynastic utility on current output q than does the static optimum, since, by definition, the static optimum allows for no such dependence. For a household receiving the dynastic utility associated with the *ex ante* optimum, $w'(w, q)$ provides less dependence of future dynastic utility on current output than does the *ex ante* optimum, precisely because $c(w, q)$ is more sensitive to current output than is the *ex ante* optimum.

7. Concluding Remarks

The results here should generalize to environments other than unobserved effort. The idea that at perfect equality the marginal cost of unequal opportunity is second-order but the benefits are first-order appears quite general. The result on social mobility should hold for any incentive model in which the resource cost of providing a mean dynastic utility is strictly convex. For instance, an earlier version of this work proves Theorems 1 and 2 for the taste shock model of Atkeson and Lucas (1992). While the need to provide incentives is fundamental here, the particular source of the incentive problem is not. Technically, all that is needed is a binding incentive constraint. With this, both unequal opportunity and social mobility are necessary implications of an efficient, or optimal, societal arrangement.

References

Atkeson, Andrew, and Robert E. Lucas, Jr. (1992), "On Efficient Distribution with Private Information," *Review of Economic Studies*, vol. 59 (3), pp. 427–453.

Freeman, Scott, and Michael A. Sadler (2002), "On the Nonoptimality of Bequeathing Inequality," manuscript, University of Texas at Austin.

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

Hopenhayn, Hugo A., and Juan Pablo Nicolini (1997) "Optimal Unemployment Insurance," *Journal of Political Economy*, vol. 105 (2), pp. 412–438.

Khan, Aubhik, and B. Ravikumar (2002), "Enduring Relationships in an Economy with Capital," Working Paper 02-5, Federal Reserve Bank of Philadelphia.

Phelan, Christopher (1994), "Incentives and Aggregate Shocks," *Review of Economic Studies*, vol. 61 (4), pp. 681–700.

Phelan, Christopher (1995), "Repeated Moral Hazard and One-Sided Commitment," *Journal of Economic Theory*, vol. 66 (2), pp. 488–506.

Phelan, Christopher (2002), "Inequality and Fairness," *Federal Reserve Bank of Minneapolis Quarterly Review*, vol. 26 (Spring), pp. 2–11.

Phelan, Christopher, and Robert M. Townsend (1991), "Computing Multi-Period, Information-Constrained Optima," *Review of Economic Studies*, vol. 58 (5), pp. 853–881.

Prescott, Edward C., and Robert M. Townsend (1984) "Pareto Optima and Competitive Equilibria with Adverse Selection and Moral Hazard," *Econometrica*, vol. 52 (1), pp.

21–45.

Rawls, John (1971) A Theory of Justice. The Belknap Press of Harvard University Press. Cambridge.

Rogerson, William (1985), “Repeated Moral Hazard,” *Econometrica*, vol. 53, pp. 69–76.

Wang, Cheng (1995), “Dynamic Insurance with Private Information and Balanced Budgets,” *Review of Economic Studies*, vol. 62 (4), pp. 577–596.

8. Appendix: Proofs of Lemmas 3 and 4

Here I provide proofs for Lemmas 3 and 4, asserted and used above. Both assert that under certain conditions, the cost function $C(v)$ is convex.

Lemma 3 Proof.

Proof. (Lemma 3). To begin the proof of Lemma 3, note $V = (-\infty, 0)$. Next, let $\xi_{-1} = \big(\Psi_{-1}, a_{-1}(w), c_{-1}(w,q), w'_{-1}(w,q)\big)$ solve (15) subject to (4)–(6) and (13) for $v = -1$. Next, fix $\Delta > 0$, and define ξ_Δ by scaling ξ_{-1} as follows: First, for all $\{\underline{w}, \overline{w}\} \in \mathbb{R}^2_-$, let $\Psi_\Delta([\underline{w}, \overline{w}]) = \Psi_{-1}([\underline{w}/\Delta,\ \overline{w}/\Delta])$. Next, let $a_\Delta(w) = a_{-1}(w/\Delta)$, $c_\Delta(w,q) = c_{-1}(w/\Delta, q) - \log(\Delta)/\gamma$, and $w'_\Delta(w,q) = w'_{-1}(w/\Delta, q)\Delta$. By construction, ξ_Δ satisfies (4). Next, consider the incentive constraint for a given $w < 0$ and $\hat{a} < a_\Delta(w)$, that

$$\begin{aligned}
&\sum_q P\big(q|a_\Delta(w)\big)\{(1-\beta)\big[-\exp(-\gamma[c_\Delta(w,q) - v(a_\Delta(w))])\big] + \beta w'_\Delta(w,q)\} \\
\geq\ &\sum_q P(q|\hat{a})\{(1-\beta)\big[-\exp\big(-\gamma[c_\Delta(w,q) - v(\hat{a})]\big)\big] + \beta w'_\Delta(w,q)\}.
\end{aligned}$$

With the definition of ξ_Δ substituting into this, it simplifies to

$$\begin{aligned}
&\sum_q P\big(q|a_{-1}(w)\big)\{(1-\beta)\big[-\exp\big(-\gamma\big[c_{-1}(w/\Delta,q) - v\big(a_{-1}(w/\Delta)\big)\big]\big)\big] + \beta w'_{-1}(w/\Delta,q)\} \\
\geq\ &\sum_q P(q|\hat{a})(1-\beta)\big[-\exp\big(-\gamma[c_{-1}(w/\Delta,q) - v(\hat{a})])\big)\big] + \beta w'_{-1}(w/\Delta,q),
\end{aligned}$$

which holds since ξ_{-1} is incentive-compatible, or satisfies (6). Next, confirm the promise-keeping constraint (5) by noting that

$$
\begin{aligned}
& \sum_q P\big(q|a_\Delta(w)\big)\{(1-\beta)\big[-\exp\big(-\gamma[c_\Delta(w,q)-v\big(a_\Delta(w)\big)]\big)\big]+\beta w'_\Delta(w,q)\} \\
= \;& \Delta\sum_q P\big(q|a_{-1}(w)\big)\{(1-\beta)\big[-\exp(-\gamma\big[c_{-1}(w/\Delta,q)-v\big(a_{-1}(w)\big)\big]\big)\big]+\beta w'_{-1}(w/\Delta,q)\} \\
= \;& \Delta w/\Delta = w.
\end{aligned}
$$

Further, ξ_Δ satisfies (13) for $v=-\Delta$ since

$$
\int_V w \; d\Psi_\Delta(w) = \int_V w \; \Delta \; d\Psi_{-1}(w) = \Delta \; \int_V w \; d\Psi_{-1}(w) = -\Delta.
$$

Thus, ξ_Δ is in the constraint set of the dual problem for $v=-\Delta$. The resources consumed by ξ_Δ can be expressed as

$$
\begin{aligned}
& \int_V \sum_q P\big(q|a_\Delta(w)\big)[c_\Delta(w,q)-q]d\Psi_\Delta(w) \\
= \;& \int_V \sum_q P\big(q|a_\Delta(w\Delta)\big)[c_\Delta(w\Delta,q)-q]d\Psi_{-1}(w) \\
= \;& \int_V \sum_q P\big(q|a_{-1}(w)\big)[c_{-1}(w,q)-\log(\Delta)/\gamma-q]d\Psi_{-1}(w) \\
= \;& -\log(-v)/\gamma+\int_V \sum_q P\big(q|a_{-1}(w)\big)[c_{-1}(w,q)-q]d\Psi_{-1}(w) \\
= \;& -\log(-v)/\gamma+C(1).
\end{aligned}
$$

Next, suppose that there exists a plan ξ^* satisfying (4)–(6) and (13) for $v=-\Delta$, which has resource cost $C^*<-\log(-v)/\gamma+C(-1)$. Here, let $\delta=-1/\Delta$, and define ξ_δ by scaling ξ^* by δ as above. The same arguments as above establish that ξ_δ satisfies stationarity, incentive-compatibility (6), and promise-keeping (5). Further,

$$\int_V w\ d\Psi_\delta(w) = \int_V w/\delta\ d\Psi^*(w) = (1/\delta)\int_V w\ d\Psi^*(w) = -1,$$

and thus ξ_δ satisfies (13) for $v = -1$. The resource cost of ξ_δ is, then,

$$\begin{aligned}
& \int_V \sum_q P\big(q|a_\delta(w)\big)[c_\delta(w,q) - q]d\Psi_\delta(w) \\
= & \int_V \sum_q P\big(q|a_\delta(w\delta)\big)[c_\delta(w\delta,q) - q]d\Psi^*(w) \\
= & \int_V \sum_q P\big(q|a^*(w)\big)[c^*(w,q) - \log(\delta)/\gamma - q]d\Psi^*(w) \\
= & -\log(-1/v)/\gamma + \int_V \sum_q P\big(q|a^*(w)\big)[c^*(w,q) - q]d\Psi^*(w) \\
= & \log(-v)/\gamma + C^* < \log(-v)/\gamma - \log(-v)/\gamma + C(-1) = C(-1),
\end{aligned}$$

which contradicts optimality of ξ_{-1}. ■

Lemma 4 Proof.

Proof. (Lemma 4). The proof of Lemma 4 proceeds in the same way as that of Lemma 3. For $\gamma > 0$, $\xi_1 = \{\Psi_1, k_1(w), a_1(w), c_1(w,q), w_1'(w,q)\}$ is defined as the optimal allocation delivering $v = 1$. Then, for $\Delta > 0$, a new allocation ξ_Δ is defined such that for all $[\underline{w}, \overline{w}] \in \mathbb{R}_+^2$, $\Psi_\Delta([\underline{w}, \overline{w}]) = \Psi_1([\underline{w}/\Delta,\ \overline{w}/\Delta])$, $k_\Delta(w) = k_1(w/\Delta)\Delta^{1/\gamma}$, $a_\Delta(w) = a_1(w/\Delta)$, $c_\Delta(w,q) = c_1(w/\Delta, q)\Delta^{1/\gamma}$, and $w_\Delta'(w,q) = w_1'(w/\Delta,q)\Delta$. I can show that ξ_Δ is satisfied for (4)– (6) and (13) for $v = \Delta$. Further, if any other plan had a lower value for the dual objective function (16), it could be used to generate a lower cost plan for delivering $v = 1$. Then $c_\Delta(w,q) = c_1(w/\Delta,q)\Delta^{1/\gamma}$, and $v = \Delta$ delivers $C(v) = v^{1/\gamma}C(1)$.

For $\gamma < 0$, $\xi_{-1} = \{\Psi_{-1}, k_{-1}(w), a_{-1}(w), c_{-1}(w,q), w_{-1}'(w,q)\}$ is defined as the optimal allocation delivering $v = -1$. Then, for $\Delta > 0$, a new allocation ξ_Δ is defined such that for all $[\underline{w}, \overline{w}] \in \mathbb{R}^2_-$, $\Psi_\Delta([\underline{w}, \overline{w}]) = \Psi_{-1}([\underline{w}/\Delta,\ \overline{w}/\Delta])$, $k_\Delta(w) = k_{-1}(w/\Delta)\Delta^{1/\gamma}$, $a_\Delta(w) = a_{-1}(w/\Delta)$, $c_\Delta(w,q) = c_{-1}(w/\Delta, q)\Delta^{1/\gamma}$, and $w'_\Delta(w,q) = w'_{-1}(w/\Delta, q)\Delta$, and the argument proceeds unaltered. Then $c_\Delta(w,q) = c_{-1}(w/\Delta, q)\Delta^{1/\gamma}$, and $v = -\Delta$ delivers $C(v) = (-v)^{1/\gamma}C(-1)$.

Finally, for $\gamma = 0$, or $U(c,a,k) = \log\big(c - kv(a)\big)$, the reference allocation $\xi_0 = \{\Psi_0, k_0(w), a_0(w), c_0(w,q), w'_0(w,q)\}$ is defined as the optimal allocation delivering $v = 0$. Then, for $\Delta \in \mathbb{R}$, a new allocation ξ_Δ is defined such that for all $[\underline{w}, \overline{w}] \in \mathbb{R}^2$, $\Psi_\Delta([\underline{w}, \overline{w}]) = \Psi_0([\underline{w} + \Delta,\ \overline{w} + \Delta])$, $k_\Delta(w) = k_0(w - \Delta)\exp(\Delta)$, $a_\Delta(w) = a_0(w - \Delta)$, $c_\Delta(w,q) = c_0(w - \Delta, q)\exp(\Delta)$, and $w'_\Delta(w,q) = w'_0(w - \Delta, q) + \Delta$, and the argument proceeds unaltered. Then $c_\Delta(w,q) = c_0(w - \Delta, q)\exp(\Delta)$, and $v = \Delta$ delivers $C(v) = \exp(v)C(0)$. ■