

NBER WORKING PAPER SERIES

PARETO EFFICIENCY AND IDENTITY

Christopher Phelan  
Aldo Rustichini

Working Paper 20883  
<http://www.nber.org/papers/w20883>

NATIONAL BUREAU OF ECONOMIC RESEARCH  
1050 Massachusetts Avenue  
Cambridge, MA 02138  
January 2015

We thank Larry Jones, Manuel Amador, Fernando Alvarez, and workshop participants at Stanford University for useful discussions. Aldo Rustichini thanks the National Science Foundation (grant SES-1061817). The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Minneapolis, the Federal Reserve System, or the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2015 by Christopher Phelan and Aldo Rustichini. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Pareto Efficiency and Identity  
Christopher Phelan and Aldo Rustichini  
NBER Working Paper No. 20883  
January 2015  
JEL No. D6,E24,H21

### **ABSTRACT**

Inherent in the definition of Pareto efficiency is the idea that, in dynamic environments, an individual is indexed by the history of events up to his birth (rather than, as usual, the date of birth). Here, we explore the implications of this natural formulation. The set of Pareto efficient allocations that is consistent with this view is potentially larger than those considered so far in the literature. We show that the set of allocations is strictly larger because we do not require individuals to have insurance motives of the Harsanyi-Rawls type regarding risks on their own type realization. We do, however, maintain the insurance motives of parents toward their children. Even in our more general framework, efficiency criteria impose substantial restrictions on the set of allocations. Interestingly, the restrictions are of a new nature. Our different, more natural view has some important policy implications. The first is that some policy criteria (for example, the progressive nature of taxes) cannot be defended on efficiency grounds, once the Harsanyi-Rawlsian insurance criterion is rejected as being normatively unsound. Second, we show that the condition of imposing no taxes of any kind, coupled with each agent owning his own production, results in a Pareto efficient allocation.

Christopher Phelan  
Department of Economics  
University of Minnesota  
1925 4th Avenue South, 4-101 Hanson Hall  
Minneapolis, MN 55455-0462  
and NBER  
cphelan@umn.edu

Aldo Rustichini  
Department of Economics  
University of Minnesota  
Minneapolis, MN 55455  
aldo.rustichini@gmail.com

## 1. Introduction

Inherent in the definition of Pareto efficiency is the idea that individuals have preference orderings over allocations. If allocation  $A$  is Pareto dominated by allocation  $B$ , it is because all individuals weakly prefer allocation  $B$  to allocation  $A$  and at least one individual strictly prefers allocation  $B$  to allocation  $A$ . Further, although it may be possible for two individuals to have the same preference ordering, it is not possible for one individual to have two preference orderings. The individual's preferences over allocations are what they are.

These obvious statements have an important implication. In this paper, we argue that the requirement that a person have a unique preference ordering implies that in dynamic economies, individuals born after the beginning of time must be indexed by the history that preceded their births. That is, an individual born after one history is a *different person* (for the purposes of Pareto comparisons across allocations) than an individual born at the same date but after a different history. We then consider the implications of such a Pareto criterion on a repeated Mirrlees (1971, 1976) economy similar to that of Farhi-Werning (2007), where each period represents a lifetime for an agent, but such an agent cares about the utility of his descendants. Here we show that the set of Pareto efficient allocations that is consistent with this view is potentially larger than those considered so far in the literature. In fact, we show that they are strictly larger, because we do not restrict individuals to having insurance motives of the Harsanyi-Rawls type against risks on their own type realization and the realizations of their ancestors. We do, however, maintain insurance motives of parents toward their children.

Even in our more general framework, efficiency criteria impose substantial restrictions on the set of allocations. Interestingly, the restrictions are of a new nature. Our different, more natural view has some important policy implications. The first is that some policy criteria (for example, the progressive nature of taxes) cannot be defended on efficiency grounds, once the Harsanyi-Rawlsian

insurance criterion is rejected as normatively unsound. Second, we show that imposing no taxes of any kind, coupled with each agent owning his own production, results in a Pareto efficient allocation.

## 2. Individuals as Histories

To illustrate the main idea, that individuals must be indexed by the history that preceded their births, we begin with a simple economy in which one father in the first period has one son in the second. The father knows his own economically relevant type (for example, his own skill), but in the first period, he does not know the type of his son, who may have high or low skill. The father cares about the happiness of the son. Upon his birth, the son knows his own type, immediately and completely, and he is able to make all relevant decisions accordingly. In this model, the father has an insurance motive: when he makes the relevant decisions, he has true uncertainty about the type of the son, and since the son's happiness affects his, he may want to insure against this risk. For the son, the situation is different: at the moment in which he makes choices, he knows his type. Ex ante, in this economy there are three persons, and three persons are relevant for efficiency analysis; one person (the father) has true uncertainty and two persons (each of his potential sons) have not.

We may ask the realization of the son to go back in time, before his birth, and tell us his order over allocations from this behind the veil point of view. If we impose some utility function on him (for example, expected utility, as in Harsanyi (1955), or mini-max, as in Rawls (1971)), he will give us answers, and in this hypothetical choice he would have reasons for insurance. That is, we might impute to him an insurance motive of the Harsanyi-Rawls type. By doing so, note that we are introducing and making relevant preferences that, in our simple model, are not held by any individual at any point in time. Using the real preferences of the father and the hypothetical preferences we have described, we would then be able to select allocations that are efficient in this two-person economy. These allocations are typically efficient in our three potential persons economy; they may be obtained by restricting the Pareto weights to depend only on the date of birth of the

son. They are not, however, all of the Pareto efficient allocations. By introducing the Harsanyi-Rawls thought experiment, we have arbitrarily restricted the set of efficient allocations. If we find properties that all such allocations share, we cannot legitimately conclude that these properties are shared by all efficient allocations.

This model is a simplification, even in some crucial aspects, which we have made explicit. We will hold similar simplified assumptions in the first part of the paper, which considers a repeated Mirrlees type environment, and then we will consider extensions. One very important restriction is that the son knows his type immediately, completely, and at the moment in which he is able to make competent decisions. This assumption allows for an unambiguous discussion about the son and his preference ordering. Reality, however, is more complex: types are revealed progressively during the lifetime; for example, the son may know his gender at birth, know his IQ at 12 years of age, and be able to make competent decisions at 18. We consider these complexities later in the paper (see section 6), and discuss how much, if at all, they change our conclusions.

In our simple illustration, the set of all Pareto efficient allocations can be found if we insist that, as mentioned earlier, individuals born after the beginning of time must be indexed by the history that preceded their births. To further illustrate the implications of this criterion, including the idea that entire histories must be included and not just an individual's own type at birth, consider the following single consumption good endowment economy. At the beginning of date  $t = 1$ , two individuals with names  $i \in \{1, 2\}$  are born — one with a high endowment and one with a low endowment. Let the aggregate state  $\theta_1 \in \{\theta^1, \theta^2\}$  be determined *after* these date  $t = 1$  agents are born and denote which agent has a high endowment. Each  $t = 1$  individual consumes  $c_{i,1}(\theta_1)$  of the consumption good and then dies. At the beginning of date  $t = 2$ , two more individuals with names  $j \in \{1, 2\}$  are born, and then state  $\theta_2 \in \Theta_2$  is realized and each individual consumes  $c_{j,2}(\theta_1, \theta_2)$  of the consumption good and then dies. Thus, an allocation is a vector  $(c_{i,1}(\theta_1), c_{j,2}(\theta_1, \theta_2))$  with

$2(2 + 2N(\Theta_2))$  elements, where  $N(\Theta_2)$  is the number of elements of  $\Theta_2$ .

Next, consider two allocations  $A$  and  $B$  such that  $c_{1,2}^A(\theta^1, \theta_2) > c_{1,2}^B(\theta^1, \theta_2)$  and  $c_{1,2}^A(\theta^2, \theta_2) < c_{1,2}^B(\theta^2, \theta_2)$  for all  $\theta_2$ . That is, person  $j = 1$  born at date  $t = 2$  receives more consumption for each  $\theta_2$  realization under allocation  $A$  than  $B$  if  $\theta_1 = \theta^1$  and more consumption for each  $\theta_2$  under allocation  $B$  than  $A$  if  $\theta_1 = \theta^2$ . (The idea here is that allocations  $A$  and  $B$  are such that person  $j = 1$  at date  $t = 2$  consumes more regardless of the state at date  $t = 2$  if his “father,” the person with the same name at date  $t = 1$ , has a high endowment than if his father has a low endowment.) If at date  $t = 2$  person  $j = 1$  is selfish (certainly an allowable preference ordering), then how he ranks allocations  $A$  and  $B$  depends on the realization of  $\theta_1$ . But if our concept of an individual requires a single preference ordering over allocations, then person  $j = 1$  born at date  $t = 2$  after realization  $\theta_1 = \theta^1$  is a *different person* than person  $j = 1$  born at date  $t = 2$  after realization  $\theta_1 = \theta^2$ , for he ranks allocations differently. One cannot make Pareto comparisons between allocations  $A$  and  $B$  without indexing date  $t = 2$  agents by  $\theta_1$ . In this example, where  $\theta_1$  can take on two values, there are six individuals (persons 1 and 2 at date  $t = 1$  and persons 1 and 2 at date  $t = 2$  for each value of  $\theta_1$ ).

But what if we change this example so that each person  $i$  at date  $t = 1$  lives for two periods, and persons  $j$  at  $t = 2$  are never born? In this second environment, an allocation is exactly the same object as in the previous environment — a vector  $(c_{i,1}(\theta_1), c_{j,2}(\theta_1, \theta_2))$ . However, most standard analysis would define Pareto efficiency relative to only two preference orderings: those of persons 1 and 2 at the beginning of date  $t = 1$ . But the other four preference orderings (those of each person for each realization of  $\theta_1$  at the beginning of date  $t = 2$ ) still exist. To compare allocations under the Pareto criterion taking into account only the first two preference orderings (those of persons 1 and 2 at the beginning of date  $t = 1$ ) is to ignore these other four people. For most cases, we believe this approach is correct. For instance, when an economist argues that outlawing insurance is inefficient,

he or she is ignoring the preference ordering of those individuals whose houses did not burn down and would prefer not to pay to rebuild the houses that did. But when these “ex post” individuals are actually separate human beings and not simply the future self of a pre-existing human, it is no longer clear to us that their preference orderings should be ignored.

A more specific example may be useful. Let the preferences of person  $i$  born at date  $t = 1$  be represented by

$$\sum_{\theta_1, \theta_2} \pi(\theta_1)\pi(\theta_2) \left( u(c_{i,1}(\theta_1)) + u(c_{i,2}(\theta_1, \theta_2)) \right) \quad (1)$$

and the preferences of person  $i$  born at date  $t = 2$  after history  $\theta_1$  be represented by

$$\sum_{\theta_2} \pi(\theta_2) u(c_{i,2}(\theta_1, \theta_2)). \quad (2)$$

Such an example represents a world in which a parent cares about the consumption of his child (but the children care only for themselves). Or, if individuals live for two periods, equation (1) represents person  $i$ 's preferences at the beginning of date  $t = 1$  and equation (2) represents person  $i$ 's preferences at the beginning of period  $t = 2$  after history  $\theta_1$ .

At date  $t = 1$ , let  $\theta_1 \in \Theta_1 = \{\theta^1, \theta^2\}$  be revealed. At date  $t = 2$ , let  $\Theta_2 = \{\theta_2\}$  (a singleton) denoting that all uncertainty is revealed at date  $t = 1$ . If  $\theta_1 = \theta^1$ , person 1 has an endowment of three in date  $t = 1$  and an endowment of one in date  $t = 2$ , and person 2 has an endowment of one in date  $t = 1$  and an endowment of three in date  $t = 2$ . If  $\theta_1 = \theta^2$ , person 1 has an endowment of one in date  $t = 1$  and an endowment of three in date  $t = 2$ , and person 2 has an endowment of three in date  $t = 1$  and an endowment of one in date  $t = 2$ .

Consider two allocations. In allocation A (autarky),  $c_{1,1}(\theta^1) = 3$ ,  $c_{1,2}(\theta^1) = 1$ ,  $c_{1,1}(\theta^2) = 1$ ,  $c_{1,2}(\theta^2) = 3$ ,  $c_{2,1}(\theta^1) = 1$ ,  $c_{2,2}(\theta^1) = 3$ ,  $c_{2,1}(\theta^2) = 3$ , and  $c_{2,2}(\theta^2) = 1$ . Let allocation B (borrowing and lending) have  $c_{i,t}(\theta_1) = 2$  for  $i \in \{1, 2\}$ ,  $t \in \{1, 2\}$ , and  $\theta_1 \in \{\theta^1, \theta^2\}$ . Assuming standard assumptions on  $u$  and no societal ability to transfer the consumption good across periods, this is the equilibrium outcome when individuals can borrow and lend.

In this example, allocation B is strictly preferred to allocation A by both individuals born at date  $t = 1$ . But allocation A is strictly preferred to allocation B from the perspective of person 2 at date  $t = 2$  after history  $\theta_1 = \theta^1$  as well as by person 1 at date  $t = 2$  after history  $\theta_1 = \theta^2$ . If these date  $t = 2$  people are interpreted as simply the future selves of those born at date  $t = 1$ , then this reversal simply expresses the desire to not repay a loan. But if these date  $t = 2$  people are interpreted as new individuals, then this reversal expresses the desire to not repay a loan taken out by someone else.

In both Phelan (2006) and Farhi and Werning (2007, 2010), Pareto efficiency is examined from another perspective as well. Applied to this example economy, both papers consider the rankings of individuals born in period  $t = 1$  as well as mythical individuals who, in essence, exist from the beginning of time but know they will be born at date  $t = 2$ . Thus, they do not know the state  $\theta_1$  and therefore rank allocations according to

$$\sum_{\theta_1, \theta_2} \pi(\theta_1)\pi(\theta_2)u(c_{i,2}(\theta_1, \theta_2)). \quad (3)$$

Like those born at date  $t = 1$ , allocation B is strictly preferred over allocation A by an individual with a preference ordering represented by (3).

One implication of considering the ranking of such a mythical individual is that any allocation that is Pareto efficient in relation to individuals with rankings represented by equations (1) and (3) is also Pareto efficient in relation to individuals with rankings represented by equations (1) and (2). (If for a given allocation, an alternative allocation is weakly worse for all  $i$  and  $\theta_1$  under criterion (2), then it is weakly worse given criterion (3) since (3) is an average of (2) over realizations  $\theta_1$ .) However, allocations that are Pareto efficient in relation to individuals with rankings represented by equations (1) and (2) are not necessarily Pareto efficient in relation to individuals represented by rankings (1) and (3). In the preceding example, allocation B is preferred to allocation A by these “pre-existing” individuals with a ranking represented by (3). Thus, if only allocations A and B are



feasible, only allocation B is Pareto efficient if preference orderings (1) and (3) are the only ones considered. But allocation A is Pareto efficient as well if preference ordering (2) is considered for all  $\theta_1$ .

It is now time to proceed with a formal analysis.

### 3. The Model

Consider a repeated Mirrlees economy with  $T + 1$  dates (generations),  $t \in \{0, 1, \dots, T\}$  where at each date  $t$ , there exists a unit continuum of agents with *family names*  $i \in [0, 1]$ . At each date  $t$ , each agent  $i$  has a *type*  $\theta_{i,t} \in \Theta$ , where  $\Theta$  is a finite set of positive real numbers. For a type  $\theta$  agent to produce  $y \geq 0$  units of the single consumption good, he must exert  $\frac{y}{\theta}$  units of labor effort. Agents live for one period, but an agent from family  $i$  at each date  $t < T$  (a father) is associated with a single agent from family  $i$  at date  $t + 1$  (his son). Thus, associated with each agent from family  $i$  at date  $t$  is the history of shocks of his dynasty,  $\theta_i^t = \{\theta_{i,0}, \dots, \theta_{i,t}\}$ . Let  $\pi(\theta|\theta^-)$  denote the probability that a type  $\theta^-$  father has a type  $\theta$  son. The transition matrix  $\pi$  is such that for all  $\theta \in \Theta$ ,  $f(\theta) = \sum_{\theta^-} \pi(\theta|\theta^-)f(\theta^-)$ , so the fraction of agents of each type is constant over time, and such that for all  $(\theta, \theta^-)$ ,  $\pi(\theta|\theta^-) > 0$ , which ensures that all future  $\theta$  paths are possible given past realizations.

An allocation is a collection  $\{\{c_{i,t}(\theta^t), y_{i,t}(\theta^t)\}_{t=0}^{T-1}\}_{i \in [0,1]}$ . A *symmetric* allocation is an allocation such that for all  $(i, j)$ ,  $t$  and  $\theta_i^t = \theta_j^t$ ,  $c_{i,t}(\theta_i^t) = c_{j,t}(\theta_j^t)$  and  $y_{i,t}(\theta_i^t) = y_{j,t}(\theta_j^t)$ , and thus is denoted  $\{c_t(\theta^t), y_t(\theta^t)\}_{t=0}^{T-1}$ . Let  $\theta^{s,t} = \{\theta_s, \dots, \theta_t\}$  denote a family's realized types between dates  $s$  and  $t$ . Symmetric allocations are feasible if

$$\sum_{t=0}^T \sum_{\theta^t} \frac{1}{R^t} f(\theta_0) \pi(\theta_1|\theta_0) \dots \pi(\theta_t|\theta_{t-1}) (c_t(\theta^t) - y_t(\theta^t)) \leq 0, \quad (4)$$

where  $R > 1$ . From here on, we consider only symmetric allocations.

Agents have identical preferences over the consumption-labor pairs of themselves and their

descendants. We assume an agent born at date  $t$  with family history  $\theta^t$  ranks allocations according to

$$\begin{aligned}
U_t(\theta^t) \equiv & u(c_t(\theta^t)) - h\left(\frac{y_t(\theta^t)}{\theta_t}\right) + \sum_{s=t+1}^T \beta^{s-t} \sum_{\theta^{t+1,s}} \pi(\theta_{t+1}|\theta_t) \dots \pi(\theta_s|\theta_{s-1}) \\
& [u(c_s(\theta^t, \theta^{t+1,s})) - h\left(\frac{y_s(\theta^t, \theta^{t+1,s})}{\theta_s}\right)],
\end{aligned} \tag{5}$$

where  $u$  is assumed differentiable, strictly increasing, strictly concave, and with  $\lim_{c \rightarrow 0} u'(c) = \infty$ , and  $h(\ell) = \ell^\psi$  for some  $\psi > 1$ . That  $h$  takes this functional form ensures that the utility possibilities set is convex when  $\theta$  realizations are private information. This ranking also assumes restrictions regarding altruism toward descendants. In particular, we assume that a father cares about the expectation of the discounted dynastic utility of his son, where  $\beta \in (0, 1)$ . Exponential discounting ensures time consistency regarding the preferences of a father toward his grandchild and the preferences of his son toward that grandchild. This assumption is restrictive and important for some efficiency results detailed later in this paper.

The type realization  $\theta_t$  of an agent born at date  $t$  to family  $i$  is alternatively assumed to be publicly observed (the *full information* case) or privately observed by that agent (the *private information* case). In both cases, all other objects, such as an agent's consumption,  $c$ , and production,  $y$ , are assumed publicly observed with one exception: in the case of private information, an agent's labor effort,  $\frac{y}{\theta}$ , is also assumed to be privately observed by that agent. (Otherwise, one could infer an agent's  $\theta$  type from his labor effort and output.)

In the full information case, a symmetric allocation  $\{c_t(\theta^t), y_t(\theta^t)\}_{t=0}^T$  specifies consumption and required output levels as functions of the date and a family's *realized* type outcomes. In the private information case, an allocation is understood to specify consumption and required output levels as functions of the date, and the family's *announcements* of type outcomes, where the announcement of  $\theta_t$  is made by the family member living at date  $t$ .

In the case of private information, an allocation is considered *incentive compatible* if for all  $t$ ,  $\theta^t$ , and  $\hat{\theta} \neq \theta_t$ ,

$$U_t(\theta^t) \geq u(c_t(\theta^{t-1}, \hat{\theta})) - h\left(\frac{y_t(\theta^{t-1}, \hat{\theta})}{\theta_t}\right) + \sum_{s=t+1}^T \beta^{s-t} \sum_{\theta^{t+1,s}} \pi(\theta_{t+1}|\theta_t) \dots \pi(\theta_s|\theta_{s-1}) \quad (6)$$

$$[u(c_s(\theta^{t-1}, \hat{\theta}, \theta^{t+1,s})) - h\left(\frac{y_s(\theta^{t-1}, \hat{\theta}, \theta^{t+1,s})}{\theta_s}\right)].$$

In words, this condition requires that an agent born at date  $t$  to a family with history  $\theta^t$  weakly prefers to announce his true type given that his descendants will truthfully announce their true types.

Let the full information *symmetric utility possibilities set* be the set of collections  $\{\{U_t(\theta^t)\}_{\theta^t \in \Theta^t}\}_{t=0}^T$  that can be generated by an allocation satisfying the resource condition (4). Let the private information utility possibilities set be such collections that also satisfy the incentive condition (6). A symmetric allocation is considered Pareto efficient if no other feasible symmetric allocation generates a profile  $\{\{U_t(\theta^t)\}_{\theta^t \in \Theta^t}\}_{t=0}^T$  that weakly dominates it and strictly dominates it for at least one  $t$  and  $\theta^t$ . Likewise, a symmetric allocation is considered constrained Pareto efficient if no other feasible and incentive compatible symmetric allocation generates a dominating utility profile.

**Lemma 1.** *The full information and private information utility possibilities sets are each convex.*

*Proof.* Define  $\bar{u}_t(\theta^t) \equiv u(c_t(\theta^t))$  and  $\bar{h}_t(\theta^t) \equiv \left(\frac{y_t(\theta^t)}{\theta_t}\right)^\psi$ . Then a symmetric allocation can be considered a specification  $\{\bar{u}_t(\theta^t), \bar{h}_t(\theta^t)\}_{t=0}^T$ . The utility of each type  $U_t(\theta^t)$  is then a linear function of the allocation. The incentive condition becomes for all  $t$ ,  $\theta^t$ , and  $\hat{\theta} \neq \theta_t$ ,

$$U_t(\theta^t) \geq \bar{u}_t(\theta^{t-1}, \hat{\theta}) - \left(\frac{\hat{\theta}}{\theta_t}\right)^\psi \bar{h}_t(\theta^{t-1}, \hat{\theta}) + \sum_{s=t+1}^T \beta^{s-t} \sum_{\theta^{t+1,s}} \pi(\theta_{t+1}|\theta_t) \dots \pi(\theta_s|\theta_{s-1}) \quad (7)$$

$$[\bar{u}_s(\theta^{t-1}, \hat{\theta}, \theta^{t+1,s}) - \bar{h}_s(\theta^{t-1}, \hat{\theta}, \theta^{t+1,s})],$$

which is a linear function of the allocation. Finally, the resource constraint becomes

$$\sum_{t=0}^T \sum_{\theta^t} \frac{1}{R^t} f(\theta_0) \pi(\theta_1|\theta_0) \dots \pi(\theta_t|\theta_{t-1}) (u^{-1}(\bar{u}_t(\theta^t)) - \theta_t \bar{h}_t(\theta^t)^{\frac{1}{\psi}}) \leq 0, \quad (8)$$

a condition that a convex function of the allocation be weakly less than zero. ■

Define the planner's problem as that of maximizing a weighted sum of lifetime utilities, where  $\gamma_t(\theta^t) \geq 0$  is the weight the planner gives to an agent born at date  $t$  into a family with history  $\theta^t$ :

$$PP : \quad \max_{c_t(\theta^t), y_t(\theta^t)} \sum_{t=0}^T \sum_{\theta^t} f(\theta_0) \pi(\theta_1|\theta_0) \dots \pi(\theta_t|\theta_{t-1}) \gamma_t(\theta^t) U_t(\theta^t) \quad (9)$$

subject to  $c_t(\theta^t)$  and  $y_t(\theta^t)$  non-negative for all  $t$  and  $\theta^t$ , resource feasibility (4), and, for the case of private information, incentive compatibility (6). The principal difference between our approach and that in Phelan (2006) or Farhi-Werning (2007, 2010) is that here  $\gamma_t(\theta^t)$  is allowed to depend on  $\theta^t$ . Farhi-Werning (2007) essentially restricts  $\gamma_t(\theta^t)$  to be equal across all  $\theta^t$ , whereas Phelan (2006) essentially restricts  $\gamma_t$  to be equal across dates  $t$  as well.

Note that if we let

$$\Gamma_t(\theta^t) \equiv \gamma_t(\theta^t) + \sum_{s=1}^t \beta^s \gamma_{t-s}(\theta^{t-s}) = \gamma_t(\theta^t) + \beta \Gamma_{t-1}(\theta^{t-1}),$$

(or  $\Gamma_t(\theta^t)$  is the weight the planner puts on the *instantaneous* utility of a type  $(t, \theta^t)$  agent), then the planner's problem becomes

$$PP : \quad \max_{c_t(\theta^t), y_t(\theta^t)} \sum_{t=0}^T \sum_{\theta^t} f(\theta_0) \pi(\theta_1|\theta_0) \dots \pi(\theta_t|\theta_{t-1}) \Gamma_t(\theta^t) (u(c_t(\theta^t)) - h(\frac{y_t(\theta^t)}{\theta_t})) \quad (10)$$

again subject to  $c_t(\theta^t)$  and  $y_t(\theta^t)$  non-negative for all  $t$  and  $\theta^t$ , resource feasibility (4), and, for the case of private information, incentive compatibility (6). Note that given weights  $\Gamma_t(\theta^t)$ , one can also back out the implied weights  $\gamma_t(\theta^t) = \Gamma_t(\theta_t) - \beta \Gamma_{t-1}(\theta^{t-1})$ . In particular, note that although arbitrary weights  $\gamma_t(\theta^t) \geq 0$  imply weights  $\Gamma_t(\theta^t) \geq 0$ , that  $\Gamma_t(\theta^t) \geq 0$  for all  $t$ ,  $\theta^t$  does *not* imply  $\gamma_t(\theta^t) \geq 0$  for all  $t$ ,  $\theta^t$ .

#### 4. Full Information

In this section, we consider the implications of Pareto efficiency when  $\theta$  realizations are public information.

**Lemma 2.** *Every Pareto efficient symmetric allocation solves the planner's problem for some weights  $\Gamma_t(\theta^t) > 0$  for all  $t \geq 0$  and  $\theta^t \in \Theta^t$ . Likewise, if a symmetric allocation solves the planner's problem given weights  $\Gamma_t(\theta^t) > 0$  for all  $t \geq 0$  and  $\theta^t \in \Theta^t$ , then it is Pareto efficient if and only if  $\Gamma_t(\theta^t) - \beta\Gamma_{t-1}(\theta^{t-1}) \geq 0$  for all  $t \geq 0$  and  $\theta^t \in \Theta^t$  (or, alternatively, if  $\gamma_t(\theta^t) \geq 0$  for all  $t$  and  $\theta^t$ ).*

*Proof.* If a symmetric allocation  $\{c_t(\theta^t), y_t(\theta^t)\}_{t=0}^T$  is Pareto efficient, then it must lie on the frontier of the utility possibilities set. The convexity of the utility possibilities set then ensures that it solves PP for some specification  $\gamma_t(\theta^t) \geq 0$  for all  $t$  and  $\theta^t$ . Next, the assumptions on  $u$  and  $h$  ensure any Pareto efficient allocation has  $c_t(\theta^t)$  and  $y_t(\theta^t)$  each strictly positive for all  $t$  and  $\theta^t$ . (Otherwise, one could generate a Pareto improvement by marginally increasing both  $y_t(\theta^t)$  and  $c_t(\theta^t)$ .) The planner's first order condition with respect to  $c_0(\theta^0)$  implies that

$$\gamma_0(\theta_0)u'(c_0(\theta^0)) = \lambda > 0, \quad (11)$$

where  $\lambda$  is the Lagrange multiplier on the resource constraint. Thus,  $\gamma_0(\theta_0) = \Gamma_0(\theta_0) > 0$  for all  $\theta_0$ . From the definition of  $\Gamma_t(\theta^t)$ , this ensures  $\Gamma_t(\theta^t) > 0$  for all  $t$  and  $\theta^t$ .

Next assume that  $\{c_t(\theta^t), y_t(\theta^t)\}_{t=0}^T$  solves PP for weights  $\Gamma_t(\theta^t) > 0$  for all  $t$  and  $\theta^t$  such that  $\Gamma_t(\theta^t) - \beta\Gamma_{t-1}(\theta^{t-1}) \geq 0$ . This then implies that  $\gamma_t(\theta^t) \geq 0$  for all  $t$  and  $\theta^t$ . That the allocation is Pareto efficient is then immediate given that  $u$  is strictly increasing (non-satiation). ■

We now turn to the main characterization result for the full information economy.

**Proposition 1.** *Assume for all dates  $t$ , a dynasty's history of shocks  $\theta^t$  is public information. Then allocation  $\{c_t(\theta^t), y_t(\theta^t)\}_{t=0}^T$  is Pareto efficient if and only if it satisfies the resource condition (4) with equality and for all  $t$ , and  $(\theta^t, \theta_{t+1})$ ,*

$$u'(c_t(\theta^t))\theta_t = h'\left(\frac{y_t(\theta^t)}{\theta_t}\right), \quad (12)$$

and

$$u'(c_t(\theta^t)) \geq \beta R u'(c_{t+1}(\theta^t, \theta_{t+1})). \quad (13)$$

*Proof.* From Lemma 2, if  $\{c_t(\theta^t), y_t(\theta^t)\}_{t=0}^T$  is Pareto efficient, then it solves PP for some  $\Gamma_t(\theta^t) > 0$  for all  $t, \theta^t$ . That the resource condition must hold with equality is then immediate. That (12) must hold follows from comparing the planner's first order conditions with respect to  $c_t(\theta^t)$  and  $y_t(\theta^t)$ . Next, the planner's first order condition with respect to  $c_0(\theta^0)$  implies that

$$\gamma_0(\theta_0) u'(c_0(\theta^0)) = \lambda, \quad (14)$$

where  $\lambda$  is the Lagrange multiplier on the resource constraint. Thus  $\gamma_0(\theta_0) > 0$  for all  $\theta_0$ . Next, the planner's first order condition with respect to  $c_t(\theta^t)$  implies for all  $t \geq 0$  that

$$\gamma_{t+1}(\theta^{t+1}) = \frac{\Gamma_t(\theta^t)(u'(c_t(\theta^t)) - \beta R u'(c_{t+1}(\theta^{t+1})))}{R u'(c_{t+1}(\theta^{t+1}))}. \quad (15)$$

Applying this sequentially from  $t = 0$  on then delivers that  $u'(c_t(\theta^t)) \geq \beta R u'(c_{t+1}(\theta^{t+1}))$  is both necessary and sufficient for  $\gamma_{t+1}(\theta^{t+1})$  to be non-negative.

Next, assume that  $\{c_t(\theta^t), y_t(\theta^t)\}_{t=0}^T$  satisfies (4) at equality, condition (12) for all  $t \leq T$  and  $\theta^t$ , and (13) for all  $t < T$  and  $\theta^{t+1}$ . Since PP is a concave programming problem, these are sufficient for  $\{c_t(\theta^t), y_t(\theta^t)\}_{t=0}^T$  to solve PP with weights  $\gamma_t(\theta^t) \geq 0$  defined by (15). Lemma 2 then implies Pareto efficiency. ■

The intuition behind this result is that an allocation is Pareto efficient if it does not waste resources, that no individual can be made better off by varying his labor effort and letting him consume the resulting variation in output, and finally, that no type  $\theta^t$  individual can be made better off by reducing his consumption by  $\epsilon$  and increasing the consumption of his type  $\theta_{t+1}$  son by  $\epsilon \frac{R}{\pi(\theta_{t+1}|\theta_t)}$ . Note that these conditions are *exactly* the same conditions as when periods are interpreted

as dates in a  $T + 1$  lived individual's life as opposed to generations, with the single exception that condition (13) is a weak inequality instead of an equality. This condition implies that no father has an unsatisfied desire to give a positive gift to any particular type son at an actuarially fair rate of return. We next show that progressive estate taxes are not necessary for Pareto efficiency.

### A. Full Information Tax Implementation

From Proposition (1), zero marginal labor taxes are necessary for any decentralized implementation of a Pareto efficient allocation.

Next, for any given allocation, define the implicit inheritance tax  $\tau_{t+1}(\theta^{t+1})$  on a type  $\theta^{t+1}$  son as solving

$$u'(c_t(\theta^t)) = \beta R u'(c_{t+1}(\theta^{t+1}))(1 - \tau_{t+1}(\theta^{t+1})). \quad (16)$$

This is the inheritance tax necessary to induce a type  $\theta^t$  father to make exactly a zero transfer to his  $\theta^{t+1}$  son if he takes the allocation as given but then is allowed to make positive *or negative* type conditional transfers to his possible sons at actuarially fair rates of return.

The next result characterizes full information implied inheritance taxes in terms of the planner weights.

**Proposition 2.** *The full information inheritance tax is*

$$\tau_{t+1}(\theta^{t+1}) = \frac{-\gamma_{t+1}(\theta^{t+1})}{\beta \Gamma_t(\theta^t)}. \quad (17)$$

*Proof.* The proof is immediate from the first order conditions of the full information planner's problem. ■

Proposition (2) implies that in any Pareto efficient full information allocation, the implicit inheritance tax on a type  $\theta^{t+1}$  son is weakly negative and strictly negative if the son receives positive weight  $\gamma_{t+1}(\theta^{t+1})$  in the social planner's problem.

If one sets  $T = 1$  (so  $t \in \{0, 1\}$ ), the resulting two period economy is very close to being a full-information version of the private information two-period model of Farhi-Werning (2010) (the only difference being that in Farhi-Werning (2010), second period agents have no  $\theta$  type and cannot work) and thus sheds some light on their result regarding the optimality of progressive inheritance taxes.

In particular, Farhi-Werning (2010) assumes a social welfare function that is equivalent to assuming that the Pareto weight on a son,  $\gamma_1$ , cannot depend on the type of the father,  $\theta_0$ . Letting  $T = 1$  and requiring  $\gamma_1(\theta_0, \theta_1) = \gamma_1$  (a constant) delivers

$$\tau_1(\theta_0) = \frac{-\gamma_1}{\beta\gamma_0(\theta_0)}. \quad (18)$$

Thus, the higher the consumption of the date  $t = 0$  father (which follows directly from a higher  $\gamma_0(\theta_0)$ ), the higher (closer to zero) the common inheritance tax on his sons. That is, when  $\gamma_1$  is restricted to be a constant, inheritance taxes are progressive.

But next consider allowing the Pareto weight of the son,  $\gamma_1(\theta_0, \theta_1)$ , to depend on the type of the father,  $\theta_0$ . (We can continue to restrict it to not depend on  $\theta_1$ , the type of the son, for this example.) Then (17) becomes

$$\tau_1(\theta_0) = \frac{-\gamma_1(\theta_0)}{\beta\gamma_0(\theta_0)}, \quad (19)$$

and implicit inheritance taxes in a full information Pareto efficient allocation can be regressive through an appropriate choice of  $\gamma_1(\theta_0)$ . In particular, if we let  $\theta_0 \in \{\underline{\theta}, \bar{\theta}\}$  with  $\gamma_0(\bar{\theta}) > \gamma_0(\underline{\theta}) > 0$  and set  $\gamma_1(\bar{\theta}) > 0$  and  $\gamma_1(\underline{\theta}) = 0$ , then inheritances taxes are negative for the “rich” ( $\bar{\theta}$  types) and zero for the “poor” ( $\underline{\theta}$  types) and thus regressive.

## B. Full Information Property Rights Implementation

In this section, we establish that although inheritance taxes may be necessary to achieve a *particular* Pareto efficient allocation, no taxes of any kind are necessary to achieve Pareto efficiency



in general.

To this end, define the full information *laissez-faire allocation*  $\{c_t^\ell(\theta^t), y_t^\ell(\theta^t)\}_{t=0}^T$  as follows:

First, for all  $\theta_T \in \Theta$ , let

$$V_T^\ell(\theta_T) \equiv \max_y u(y) - h\left(\frac{y}{\theta_T}\right). \quad (20)$$

That is,  $V_T^\ell(\theta_T)$  is the value to a date  $T$  born agent of type  $\theta_T$  of receiving a zero inheritance but having a property right to his own production. Next, for  $t \in \{0, \dots, T-1\}$  (and working backwards from  $t = T-1$  to  $t = 0$ ), define

$$\begin{aligned} V_t^\ell(\theta_t) \equiv & \max_{c_t, y_t, \{c_s(\theta^{t+1,s}), y_s(\theta^{t+1,s})\}_{s=t+1}^T} u(c_t) - h\left(\frac{y_t}{\theta_t}\right) + \sum_{s=t+1}^T \beta^{s-t} \sum_{\theta^{t+1,s}} \pi(\theta_{t+1}|\theta_t) \dots \pi(\theta_T|\theta_{T-1}) \\ & [u(c_s(\theta^{t+1,s})) - h\left(\frac{y_s(\theta^{t+1,s})}{\theta_s}\right)], \end{aligned} \quad (21)$$

subject to a budget condition

$$c_t - y_t + \sum_{s=t+1}^T \frac{1}{R^{s-t}} \sum_{\theta^{t+1,s}} \pi(\theta_{t+1}|\theta_t) \dots \pi(\theta_T|\theta_{T-1}) (c_s(\theta^{t+1,s}) - y_s(\theta^{t+1,s})) \leq 0, \quad (22)$$

and a no-enslaving-your-descendants condition, for all  $s \geq t+1$  and  $\theta^{t+1,s}$ ,

$$\begin{aligned} & u(c_s(\theta^{t+1,s})) - h\left(\frac{y_s(\theta^{t+1,s})}{\theta_s}\right) + \sum_{n=s+1}^T \beta^{n-s} \sum_{\theta^{s+1,n}} \pi(\theta_{s+1}|\theta_s) \dots \pi(\theta_T|\theta_{T-1}) \\ & [u(c_n(\theta^{t+1,n})) - h\left(\frac{y_n(\theta^{t+1,n})}{\theta_n}\right)] \geq V_s^\ell(\theta_s). \end{aligned} \quad (23)$$

Here,  $V_t^\ell(\theta_t)$  is the value to a date  $t < T$  born agent of type  $\theta_t$  of receiving a zero inheritance, again having a property right to his own production, and further having the ability to control the production and consumption of his descendants, as long as the expected discounted consumption of himself and his descendants does not exceed their production (which implies the ability to make actuarially fair conditional transfers to descendants) and these descendants do no worse than what they could achieve on their own with a zero inheritance and these same rights. (Thus, condition (23) implies that conditional inheritances be non-negative.) Once  $V_t^\ell(\theta_t)$  is thus calculated for all

$t \in \{1, \dots, T\}$ , the laissez-faire allocation then corresponds to the solution to the  $t = 0$  problem above for all  $\theta_0$ .

**Proposition 3.** *The full information laissez-faire allocation is Pareto efficient.*

*Proof.* It is straightforward to show that the laissez-faire allocation has (22) hold as an equality for each  $\theta_0$  type. (Otherwise, each  $\theta_0$  type could simply consume the extra resources himself.) That is, under the laissez-faire allocation, the expected discounted dynastic consumption of each  $\theta_0$  type equals its expected discounted dynastic output.

Suppose there exists an allocation that Pareto dominates the laissez-faire allocation. That  $\beta > 0$  and  $\pi(\theta|\theta^-) > 0$  for all  $(\theta, \theta^-)$  implies that at least one first generation  $\theta_0$  type agent is strictly better off in this Pareto-improving allocation. Further, since the allocation is a Pareto improvement, it satisfies (23) for all  $s \geq 1$  and  $(\theta^{1,s})$ . Thus, it must be the case that for this first generation  $\theta_0$  type, (22) is violated, otherwise the first generation  $\theta_0$  type would have chosen it. That is, under the Pareto-improving allocation, the expected discounted dynastic consumption of at least one  $\theta_0$  type exceeds its expected discounted dynastic output. The society-wide resource constraint then implies that for at least one  $\theta_0$  type, say  $\hat{\theta}_0$ , its expected discounted dynastic output exceeds its expected discounted dynastic consumption. Further, since the new plan is a Pareto improvement, conditions (23) are satisfied for the  $\hat{\theta}_0$  type's problem. Thus, the  $\hat{\theta}_0$  type could have chosen his part of the Pareto-improving allocation with (22) holding as a strict inequality, which is a contradiction. ■

The Pareto efficient full information laissez-faire allocation can be implemented with no taxes whatsoever (no labor taxes, no inheritance taxes). It does require property rights and sufficient insurance markets, however. In particular, fathers need to be able to make conditional transfers to their sons, and the sons need to be protected from fathers making negative conditional transfers.

In fact, these property rights for children not to receive a negative inheritance is where

positive Pareto weights  $\gamma_t(\theta^t)$  occur. Recall from (15) that

$$\gamma_{t+1}(\theta^{t+1}) = \frac{\Gamma_t(\theta^t)(u'(c_t(\theta^t)) - \beta Ru'(c_{t+1}(\theta^{t+1})))}{Ru'(c_{t+1}(\theta^{t+1}))}. \quad (24)$$

That is, a type  $\theta^{t+1}$  agent is associated with a strictly positive Pareto weight when  $u'(c_t(\theta^t)) > \beta Ru'(c_{t+1}(\theta^{t+1}))$  or precisely when, under the laissez-faire allocation, his father wishes he could make him a negative conditional transfer. Likewise, sons who receive a positive inheritance, which from the father's intertemporal first order condition when (23) is slack ensures that  $u'(c_t(\theta^t)) = \beta Ru'(c_{t+1}(\theta^{t+1}))$ , have  $\gamma_{t+1}(\theta^{t+1}) = 0$ .

## 5. Private Information

With private information, characterization of the set of Pareto efficient allocations is less straightforward. In this section, we present three main results. First, we completely characterize the set of Pareto efficient allocations for the two-period economy of Farhi-Werning (2010) subject to a restriction that  $\theta \in \{\underline{\theta}, \bar{\theta}\}$  and show that, like the economy with full information, in constrained efficient allocations, implicit inheritance taxes can be either progressive or regressive, depending on the particular constrained efficient allocation. That is, we show the result of Farhi-Werning (2010) that optimal implicit inheritance taxes are progressive depends crucially on their assumption that societal preferences toward unborn generations reflect, in their words, a “preference for equality” regarding the consumption of unborn generations. In our environment, in which a son born to a father of one  $\theta$  type is considered a different person (for the purpose of Pareto efficiency comparisons) than a son born to a father of a different  $\theta$  type, this progressivity result no longer holds.

Second, we show that the “inverse Euler condition” result of Golosov, Kocherlakota, and Tsyvinski (2003) (hereafter, GKT) (itself a generalization of Rogerson (1985)) holds as a necessary condition for any incentive constrained Pareto efficient allocation, but as an inequality as opposed to an equality as in GKT. This result is an interesting parallel to our similar result given full

information. There, the normal (non-inverse) Euler condition  $u'(c_t(\theta^t)) \geq \beta R u'(c_{t+1}(\theta^{t+1}))$  holds necessarily as an equality when only the first generation receives positive Pareto weights and holds as an inequality when future generations are given positive direct, history-dependent Pareto weights. Here, for the case of private information, we again show that the exact same necessary condition that holds as an equality when only the first generation receives positive Pareto weights — now the inverse Euler condition — holds as a weak inequality when future generations are given positive direct, history-dependent Pareto weights.

Finally, we consider the efficiency of the laissez-faire allocation given private information. Here, we first show that if  $\theta_0$  (the skill realization of the initial generation) is public but  $\theta_t$  is private for all  $t \geq 1$ , then the laissez-faire allocation is constrained efficient. We then consider the case in which  $\theta_0$  is private as well.

#### A. Farhi-Werning (2010)

Farhi-Werning (2010) (hereafter, FW) is a two-period version of our economy with one basic exception: in FW, second period agents (children) cannot produce. In this section, we completely characterize the set of constrained Pareto efficient allocations in such an economy (when types are restricted to be either low or high) and show that implicit inheritance taxes are progressive if the Pareto weights on children are restricted to be independent of their father's type, as FW essentially assume in their social welfare function approach, but can be either progressive or regressive if the Pareto weights on children can depend on the type of their father (or children of high types are assumed to be *different people*, for the purposes of Pareto comparisons, than the children of low types). The next proposition characterizes the set of constrained efficient allocations.

**Proposition 4.** *For the two-period economy  $t \in \{0, 1\}$  such that only  $t = 0$  agents can produce, an incentive compatible, resource feasible allocation  $(c_0^*(\theta), y^*(\theta), c_1^*(\theta))$ ,  $\theta \in \{\underline{\theta}, \bar{\theta}\}$  is constrained*

Pareto efficient if and only if

1. the resource constraint holds with equality,
2. the low type has a weak residual motive to work more, or

$$u'(c_0^*(\underline{\theta}))\underline{\theta} \geq h'\left(\frac{y^*(\underline{\theta})}{\underline{\theta}}\right) \quad (25)$$

(and with equality if the incentive constraint,  $u(c_0^*(\bar{\theta})) - h(\frac{y^*(\bar{\theta})}{\bar{\theta}}) + \beta u(c_1^*(\bar{\theta})) > u(c_0^*(\underline{\theta})) - h(\frac{y^*(\underline{\theta})}{\underline{\theta}}) + \beta u(c_1^*(\underline{\theta}))$ , is slack),

3. the high type has a weak residual motive to work less, or

$$u'(c_0^*(\bar{\theta}))\bar{\theta} \leq h'\left(\frac{y^*(\bar{\theta})}{\bar{\theta}}\right) \quad (26)$$

(and with equality if the incentive constraint,  $u(c_0^*(\underline{\theta})) - h(\frac{y^*(\underline{\theta})}{\underline{\theta}}) + \beta u(c_1^*(\underline{\theta})) > u(c_0^*(\bar{\theta})) - h(\frac{y^*(\bar{\theta})}{\bar{\theta}}) + \beta u(c_1^*(\bar{\theta}))$ , is slack),

4. for each  $\theta \in \{\underline{\theta}, \bar{\theta}\}$  the date  $t = 0$  father has a weak residual motive to transfer consumption from his son to himself, or

$$u'(c_0^*(\theta)) \geq \beta R u'(c_1^*(\theta)). \quad (27)$$

5. Finally, let  $MC(\theta) \equiv \frac{f(\theta)}{u'(c_0^*(\theta))}$ , or  $MC(\theta)$  equals the marginal societal cost of providing utility from consumption to type  $\theta$ , and let  $M\bar{U}(\theta) \equiv f(\theta) \left( \frac{1}{u'(c_0^*(\theta))} - \frac{\theta}{h'(y^*(\theta)/\theta)} \right)$ , or  $M\bar{U}(\theta)$  equals the marginal societal cost of increasing both consumption and output for type  $\theta$  such that the utility for type  $\theta$  stays constant. Then (as part of the necessary and sufficient conditions for  $(c_0^*(\theta), y^*(\theta), c_1^*(\theta))$  to be constrained efficient), one needs for  $\theta = \underline{\theta}$ ,  $\theta^c = \bar{\theta}$  as well as  $\theta = \bar{\theta}$ ,  $\theta^c = \underline{\theta}$ ,

$$MC(\theta) \geq \frac{M\bar{U}(\theta^c)}{(\frac{\theta^c}{\theta})^\psi - 1} - \frac{M\bar{U}(\theta)}{(\frac{\theta}{\theta^c})^\psi - 1}. \quad (28)$$

*Proof.* That the resource constraint must hold as an equality is immediate. (Increasing first period consumption to raise the utility of each type of father by an equal amount is incentive compatible and

Pareto improving.) Thus,  $(c_0^*(\theta), h^*(\theta), c_1^*(\theta))$  is Pareto efficient if and only if there exist  $\gamma_0(\theta) \geq 0$  and  $\gamma_1(\theta) \geq 0$  such that  $(\bar{u}_0^*(\theta), \bar{h}^*(\theta), \bar{u}_1^*(\theta)) \equiv (u(c_0^*(\theta)), h(\frac{y^*(\theta)}{\theta}), u(c_1^*(\theta)))$  solve

$$\min_{\bar{u}_0(\theta), \bar{h}(\theta), \bar{u}_1(\theta)} \sum_{\theta} f(\theta) [u^{-1}(\bar{u}_0(\theta)) - \theta h^{-1}(\bar{h}(\theta)) + \frac{1}{R} u^{-1}(\bar{u}_1(\theta))] \quad (29)$$

subject to the incentive conditions that for  $\theta = \underline{\theta}$ ,  $\theta^c = \bar{\theta}$  as well as  $\theta = \bar{\theta}$ ,  $\theta^c = \underline{\theta}$ ,

$$\bar{u}_0(\theta) - \bar{h}(\theta) + \beta \bar{u}_1(\theta) \geq \bar{u}_0(\theta^c) - \bar{h}(\theta^c) \left(\frac{\theta^c}{\theta}\right)^\psi + \beta \bar{u}_1(\theta^c), \quad (30)$$

and

$$\begin{aligned} & \sum_{\theta} f(\theta) [\gamma_0(\theta) (\bar{u}_0(\theta) - \bar{h}(\theta) + \beta \bar{u}_1(\theta)) + \gamma_1(\theta) \bar{u}_1(\theta)] \\ & \geq \sum_{\theta} f(\theta) [\gamma_0(\theta) (u_0^*(\theta) - h^*(\theta) + \beta u_1^*(\theta)) + \gamma_1(\theta) u_1^*(\theta)]. \end{aligned} \quad (31)$$

If each of the constraints is subtracted from the objective function to form a Lagrangian (with multipliers  $f(\theta)\mu(\theta, \theta^c)$  on the incentive constraints and  $\lambda$  on equation (31)), then the derivatives of this Lagrangian with respect to  $\bar{u}_0(\theta)$ ,  $\bar{u}_1(\theta)$ , and  $\bar{h}(\theta)$  give rise to the following first order conditions:

$$f(\theta) u^{-1'}(\bar{u}_0(\theta)) = \frac{f(\theta)}{u'(c_0(\theta))} = f(\theta)\mu(\theta, \theta^c) - f(\theta^c)\mu(\theta^c, \theta) + \lambda f(\theta)\gamma_0(\theta), \quad (32)$$

$$\frac{f(\theta)}{R} u^{-1'}(\bar{u}_1(\theta)) = \frac{f(\theta)}{R u'(c_1(\theta))} = f(\theta)\beta\mu(\theta, \theta^c) - f(\theta^c)\beta\mu(\theta^c, \theta) + \lambda f(\theta)(\beta\gamma_0(\theta) + \gamma_1(\theta)), \quad (33)$$

$$-f(\theta)h^{-1'}(\bar{h}(\theta))\theta = -\frac{f(\theta)\theta}{h'(\frac{y(\theta)}{\theta})} = -f(\theta)\mu(\theta, \theta^c) + f(\theta^c)\left(\frac{\theta}{\theta^c}\right)^\psi \mu(\theta^c, \theta) - \lambda f(\theta). \quad (34)$$

If one adds (32) and (34), one can solve for

$$\mu(\theta, \theta^c) = \frac{f(\theta^c)}{f(\theta)} \frac{\frac{1}{u'(c_0(\theta^c))} - \frac{\theta^c}{h'(\frac{y(\theta^c)}{\theta^c})}}{\left(\frac{\theta^c}{\theta}\right)^\psi - 1}. \quad (35)$$

If  $\mu(\theta, \theta^c) = 0$ , this then implies that  $u'(c_0(\theta^c))\theta^c = h'(\frac{y(\theta^c)}{\theta^c})$ . If  $\mu(\theta, \theta^c) > 0$ , it is necessary that the sign of the numerator,  $\frac{1}{u'(c_0(\theta^c))} - \frac{\theta^c}{h'(\frac{y(\theta^c)}{\theta^c})}$ , agrees with the sign of the denominator,  $\left(\frac{\theta^c}{\theta}\right)^\psi - 1$ . If  $\theta = \underline{\theta}$  and  $\theta^c = \bar{\theta}$ , then the denominator is positive and thus  $\frac{1}{u'(c_0(\bar{\theta}))} > \frac{\bar{\theta}}{h'(\frac{y(\bar{\theta})}{\bar{\theta}})}$ , or  $u'(c_0(\bar{\theta}))\bar{\theta} < h'(\frac{y(\bar{\theta})}{\bar{\theta}})$ . If  $\theta = \bar{\theta}$  and  $\theta^c = \underline{\theta}$ , then the denominator is negative and thus  $u'(c_0(\underline{\theta}))\underline{\theta} > h'(\frac{y(\underline{\theta})}{\underline{\theta}})$ .

Next, if one solves (32) for  $f(\theta)\mu(\theta, \theta^c) - f(\theta^c)\mu(\theta^c, \theta)$  and substitutes into (33), one has

$$\frac{1}{u'(c_0(\theta))} = \frac{1}{\beta Ru'(c_1(\theta))} - \lambda\gamma_1(\theta). \quad (36)$$

Since  $\gamma_1(\theta) \geq 0$ , this implies that  $u'(c_0(\theta)) \geq \beta Ru'(c_1(\theta))$ .

Finally, substituting (35) into (32) delivers

$$MC(\theta) - \frac{M\bar{U}(\theta^c)}{(\frac{\theta^c}{\theta})^\psi - 1} + \frac{M\bar{U}(\theta)}{(\frac{\theta}{\theta^c})^\psi - 1} = f(\theta)\lambda\gamma_0(\theta). \quad (37)$$

Since  $f(\theta)\lambda\gamma_0(\theta) \geq 0$ , this delivers (28) as a necessary condition.

For sufficiency, note that if one finds an allocation  $(c_0(\theta), y(\theta), c_1(\theta))$  satisfying the derived necessary conditions, these solve the Lagrangian above (which in the transformed choice variables has a convex objective function and linear constraints and thus satisfies the Kuhn-Tucker conditions) with  $\lambda\gamma_0(\theta) \geq 0$  defined by (37) and  $\lambda\gamma_1(\theta) \geq 0$  defined by (36). ■

Next we show that implicit inheritance taxes are progressive if the Pareto weight on the child of a high type parent is restricted to be equal to the Pareto weight on the child of a low type parent, or  $\gamma_1(\bar{\theta}) = \gamma_1(\underline{\theta})$ , but can be progressive or regressive if these Pareto weights are not so restricted.

**Proposition 5.** *Let  $(c_0^*(\theta), y^*(\theta), c_1^*(\theta))$ ,  $\theta \in \{\underline{\theta}, \bar{\theta}\}$  be a constrained efficient allocation for the two period economy  $t \in \{0, 1\}$  such that only  $t = 0$  agents can produce, and let its associated implicit inheritance tax  $\tau(\theta)$  solve  $u'(c_0^*(\theta)) = \beta Ru'(c_1^*(\theta))(1 - \tau(\theta))$ . Then, if Pareto weights are restricted such that  $\gamma_1(\underline{\theta}) = \gamma_1(\bar{\theta}) > 0$ ,  $\tau(\underline{\theta}) < \tau(\bar{\theta}) < 0$  (or inheritance taxes are negative and progressive). If instead  $\gamma_1(\underline{\theta}) \geq 0$  and  $\gamma_1(\bar{\theta}) \geq 0$  are unrestricted, then inheritance taxes remain (weakly) negative but can be either progressive or regressive.*

*Proof.* First, note that if one solves for  $\tau(\theta)$  from its definition, one derives

$$\tau(\theta) = \frac{\beta Ru'(c_1(\theta)) - u'(c_0(\theta))}{\beta Ru'(c_1(\theta))}. \quad (38)$$

Next, consider the primal problem

$$\max_{c_0(\theta), y(\theta), c_1(\theta)} \sum_{\theta \in \{\underline{\theta}, \bar{\theta}\}} f(\theta) [\gamma_0(\theta)(u(c_0(\theta)) - h(\frac{y(\theta)}{\theta})) + (\beta\gamma_0(\theta) + \gamma_1(\theta))u(c_1(\theta))], \quad (39)$$

subject to (for  $\theta = \underline{\theta}$  and  $\theta^c = \bar{\theta}$  as well as for  $\theta = \bar{\theta}$  and  $\theta^c = \underline{\theta}$ )

$$u(c_0(\theta)) - h(\frac{y(\theta)}{\theta}) + \beta u(c_1(\theta)) \geq u(c_0(\theta^c)) - h(\frac{y(\theta^c)}{\theta}) + \beta u(c_1(\theta^c)), \quad (40)$$

and

$$\sum_{\theta \in \{\underline{\theta}, \bar{\theta}\}} f(\theta) [c_0(\theta) - y(\theta) + \frac{1}{R}c_1(\theta)] \leq 0, \quad (41)$$

where  $f(\theta)\mu(\theta, \theta^c)$  and  $\lambda$  are the respective Lagrange multipliers.

The necessary first order conditions with respect to  $c_0(\theta)$  and  $c_1(\theta)$  are

$$f(\theta)\gamma_0(\theta)u'(c_0(\theta)) + f(\theta)\mu(\theta, \theta^c)u'(c_0(\theta)) - f(\theta^c)\mu(\theta^c, \theta)u'(c_0(\theta)) - f(\theta)\lambda = 0 \quad (42)$$

and

$$f(\theta)(\beta\gamma_0(\theta) + \gamma_1(\theta))u'(c_1(\theta)) + \beta f(\theta)\mu(\theta, \theta^c)u'(c_1(\theta)) - \beta f(\theta^c)\mu(\theta^c, \theta)u'(c_1(\theta)) - \frac{1}{R}f(\theta)\lambda = 0 \quad (43)$$

Solving (42) and (43) for  $f(\theta)\lambda$  and equating delivers

$$\begin{aligned} & \beta R u'(c_1(\theta)) [f(\theta)\gamma_0(\theta) + f(\theta)\mu(\theta, \theta^c) - f(\theta^c)\mu(\theta^c, \theta)] + R f(\theta)\gamma_1(\theta)u'(c_1(\theta)) \\ &= u'(c_0(\theta)) [f(\theta)\gamma_0(\theta) + f(\theta)\mu(\theta, \theta^c) - f(\theta^c)\mu(\theta^c, \theta)], \end{aligned} \quad (44)$$

or

$$\begin{aligned} & (\beta R u'(c_1(\theta)) - u'(c_0(\theta))) [f(\theta)\gamma_0(\theta) + f(\theta)\mu(\theta, \theta^c) - f(\theta^c)\mu(\theta^c, \theta)] \\ &= -R f(\theta)\gamma_1(\theta)u'(c_1(\theta)). \end{aligned} \quad (45)$$

Dividing each side by the expression in square brackets and  $\beta R u'(c_1(\theta))$  delivers

$$\tau(\theta) \equiv \frac{\beta R u'(c_1(\theta)) - u'(c_0(\theta))}{\beta R u'(c_1(\theta))} = \frac{-f(\theta)\gamma_1(\theta)}{\beta(f(\theta)\gamma_0(\theta) + f(\theta)\mu(\theta, \theta^c) - f(\theta^c)\mu(\theta^c, \theta))}, \quad (46)$$



or

$$\tau(\theta) = \frac{-f(\theta)\gamma_1(\theta)}{\beta(f(\theta)\gamma_0(\theta) + f(\theta)\mu(\theta, \theta^c) - f(\theta^c)\mu(\theta^c, \theta))}. \quad (47)$$

Note from (42) that

$$f(\theta)\gamma_0(\theta) + f(\theta)\mu(\theta, \theta^c) - f(\theta^c)\mu(\theta^c, \theta) = \frac{f(\theta)\lambda}{u'(c_0(\theta))}. \quad (48)$$

Thus,

$$\tau(\theta) = \frac{-1}{\beta\lambda}\gamma_1(\theta)u'(c_0(\theta)). \quad (49)$$

Next, assume that  $\gamma_1(\theta) = \gamma_1$  (a constant) for all  $\theta$ . This, (42), and (45) then imply that

$$\beta R \left( \frac{1}{u'(c_0(\underline{\theta}))} - \frac{1}{u'(c_0(\bar{\theta}))} \right) = \frac{1}{u'(c_1(\underline{\theta}))} - \frac{1}{u'(c_1(\bar{\theta}))}. \quad (50)$$

Thus, if  $c_0(\underline{\theta}) \geq c_0(\bar{\theta})$ , then  $c_1(\underline{\theta}) \geq c_1(\bar{\theta})$ .

Next, from  $u'(c_0(\underline{\theta}))\underline{\theta} \geq h'(\frac{y(\underline{\theta})}{\underline{\theta}})$  and  $u'(c_0(\bar{\theta}))\bar{\theta} \leq h'(\frac{y(\bar{\theta})}{\bar{\theta}})$ , if  $c_0(\underline{\theta}) \geq c_0(\bar{\theta})$  then  $\frac{y(\underline{\theta})}{\underline{\theta}} < \frac{y(\bar{\theta})}{\bar{\theta}}$ .

Thus, if  $c_0(\underline{\theta}) \geq c_0(\bar{\theta})$ , then  $c_1(\underline{\theta}) \geq c_1(\bar{\theta})$  and  $\frac{y(\underline{\theta})}{\underline{\theta}} < \frac{y(\bar{\theta})}{\bar{\theta}}$ , which violates incentive compatibility.

Thus,  $c_0(\underline{\theta}) < c_0(\bar{\theta})$ , which (49) then implies  $\tau(\underline{\theta}) < \tau(\bar{\theta})$ , or inheritance taxes are progressive when

$\gamma_1(\underline{\theta}) = \gamma_1(\bar{\theta}) > 0$ .

Finally, if  $\gamma_1(\underline{\theta}) = 0$  and  $\gamma_1(\bar{\theta}) > 0$ , then (49) implies  $\tau(\underline{\theta}) = 0$  and  $\tau(\bar{\theta}) < 0$ , and thus  $\tau(\underline{\theta}) > \tau(\bar{\theta})$  or inheritance taxes are regressive. ■

## B. The Inverse Euler Condition

In this section we establish that the inverse Euler condition of Golosov, Kocherlakota and Tsvynski (2003) must hold as an *inequality* in any constrained efficient allocation.

**Proposition 6.** *For the  $T+1$ -period economy with  $t \in \{0, \dots, T\}$ , an incentive compatible, resource feasible allocation  $\{c_t^*(\theta^t), y_t^*(\theta^t)\}_{t=0}^T$  is constrained Pareto efficient only if for all  $t < T$  and  $\theta^t$ ,*

$$\frac{1}{u'(c_t(\theta^t))} \leq \frac{1}{\beta R} \sum_{\theta_{t+1}} \pi(\theta_{t+1}|\theta^t) \frac{1}{u'(c_{t+1}(\theta^{t+1}))}. \quad (51)$$

*Proof.* Since the resource constraint must hold with equality in any constrained efficient allocation,  $\{c_t^*(\theta^t), y_t^*(\theta^t)\}_{t=0}^T$  is constraint efficient only if  $\{\bar{u}_t^*(\theta^t), \bar{h}_t^*(\theta^t)\}_{t=0}^T \equiv \{u(c_t^*(\theta^t)), h(\frac{y_t^*(\theta^t)}{\theta_t})\}_{t=0}^\infty$  solves

$$\min_{u_t(\theta^t), h_t(\theta^t)} \sum_{t=0}^T \sum_{\theta^t} \frac{1}{R^t} f(\theta_0) \pi(\theta_1|\theta_0) \dots \pi(\theta_t|\theta_{t-1}) (u^{-1}(\bar{u}_t(\theta^t)) - \theta_t h^{-1}(\bar{h}_t(\theta^t))) \quad (52)$$

subject to for all  $t$ ,  $\theta^t$ , and  $\hat{\theta} \neq \theta_t$

$$\begin{aligned} & \bar{u}_t(\theta^t) - \bar{h}_t(\theta^t) + \sum_{s=t+1}^T \beta^{s-t} \sum_{\theta^{t+1,s}} \pi(\theta_{t+1}|\theta_t) \dots \pi(\theta_s|\theta_{s-1}) [\bar{u}_s(\theta^s) - \bar{h}_s(\theta^s)] \\ & \geq \bar{u}_t(\theta^{t-1}, \hat{\theta}) - \left(\frac{\hat{\theta}}{\theta_t}\right)^\psi \bar{h}_t(\theta^{t-1}, \hat{\theta}) + \sum_{s=t+1}^T \beta^{s-t} \sum_{\theta^{t+1,s}} \pi(\theta_{t+1}|\theta_t) \dots \pi(\theta_s|\theta_{s-1}) \\ & \quad [\bar{u}_s(\theta^{t-1}, \hat{\theta}, \theta^{t+1,s}) - \bar{h}_s(\theta^{t-1}, \hat{\theta}, \theta^{t+1,s})], \end{aligned} \quad (53)$$

and for all  $t$  and  $\theta^t$

$$\begin{aligned} & \bar{u}_t(\theta^t) - \bar{h}_t(\theta^t) + \sum_{s=t+1}^T \beta^{s-t} \sum_{\theta^{t+1,s}} \pi(\theta_{t+1}|\theta_t) \dots \pi(\theta_s|\theta_{s-1}) [\bar{u}_s(\theta^s) - \bar{h}_s(\theta^s)] \\ & \geq \bar{u}_t^*(\theta^t) - \bar{h}_t^*(\theta^t) + \sum_{s=t+1}^T \beta^{s-t} \sum_{\theta^{t+1,s}} \pi(\theta_{t+1}|\theta_t) \dots \pi(\theta_s|\theta_{s-1}) [\bar{u}_s^*(\theta^s) - \bar{h}_s^*(\theta^s)]. \end{aligned} \quad (54)$$

Next note that if, for a particular  $t < T$  and  $\theta^t$ , one perturbs  $\{\bar{u}_t^*(\theta^t), \bar{h}_t^*(\theta^t)\}_{t=0}^T$  by decreasing  $\bar{u}_t(\theta^t)$  by  $\Delta \geq 0$  and increasing  $\bar{u}_{t+1}(\theta^{t+1})$  by  $\Delta/\beta$  for all  $\theta_{t+1}$ , and otherwise leaving the allocation unchanged, the perturbed policy remains in the constraint set. Further, it affects the objective function only in the terms for dates  $t$  and  $t+1$  following history  $\theta^t$ . Thus, if  $\{c_t^*(\theta^t), y_t^*(\theta^t)\}_{t=0}^T$  is constrained Pareto efficient, a choice of  $\Delta = 0$  must solve

$$\min_{\Delta} u^{-1}(\bar{u}_t^*(\theta^t) - \Delta) + \frac{1}{R} \sum_{\theta_{t+1}} \pi(\theta_{t+1}|\theta_t) u^{-1}(\bar{u}_{t+1}^*(\theta^{t+1}) + \Delta/\beta) \quad (55)$$

subject to  $\Delta \geq 0$ . The necessary first order condition of this problem with respect to  $\Delta$  (where  $\mu$  is the multiplier on the constraint) is

$$-\frac{1}{u'(c_t(\theta^t))} + \frac{1}{\beta R} \sum_{\theta_{t+1}} \pi(\theta_{t+1}|\theta_t) \frac{1}{u'(c_{t+1}(\theta^{t+1}))} = \mu. \quad (56)$$

That  $\mu \geq 0$  then proves the result. ■

### C. Private Information Property Rights Implementation

In this section, we mimic the results established under full information. In particular, we establish conditions under which no taxes of any kind are necessary to achieve Pareto efficiency. Our first result is that if  $\theta_0$  is public information, but  $\theta_t$  for all  $t \geq 1$  is private, then the private information laissez-faire allocation is Pareto efficient.

Define the laissez-faire allocation exactly as with full information except that at each step of deriving  $V^\ell(\theta_t)$ , a date  $t$ , type  $\theta_t$  agent faces a further constraint that his descendants must be willing to truthfully reveal their types (assuming all of their descendants truthfully reveal), or for all  $s \geq t + 1$ ,  $\theta^{t+1,s}$ , and  $\hat{\theta}$ ,

$$\begin{aligned}
& u(c_s(\theta^{t+1,s})) - h\left(\frac{y_s(\theta^{t+1,s})}{\theta_s}\right) + \sum_{n=s+1}^T \beta^{n-s} \sum_{\theta^{s+1,n}} \pi(\theta_{s+1}|\theta_s) \dots \pi(\theta_T|\theta_{T-1}) \\
& [u(c_n(\theta^{t+1,n})) - h\left(\frac{y_n(\theta^{t+1,n})}{\theta_n}\right)] \geq \\
& u(c_s(\theta^{t+1,s-1}, \hat{\theta})) - h\left(\frac{y_s(\theta^{t+1,s-1}, \hat{\theta})}{\theta_s}\right) + \sum_{s=t+1}^T \beta^{s-t} \sum_{\theta^{t+1,s}} \pi(\theta_{t+1}|\theta_t) \dots \pi(\theta_s|\theta_{s-1}) \quad (57) \\
& [u(c_n(\theta^{t+1,s-1}, \hat{\theta}, \theta^{t+1+s,n}) - h\left(\frac{y_n(\theta^{t+1,s-1}, \hat{\theta}, \theta^{t+1+s,n})}{\theta_n}\right)].
\end{aligned}$$

With private information,  $V_t^\ell(\theta_t)$  is the value to a date  $t < T$  born agent of type  $\theta_t$  of receiving a zero inheritance subject to the same conditions as with full information but with one additional condition: Again, the date  $t$  type  $\theta_t$  agent has a property right to his own production and the ability to control the production and consumption of his descendants. Again, the expected discounted consumption of himself and his descendants must not exceed their production (which implies the ability to make actuarially fair conditional transfers to descendants) and these descendants do no worse than what they could achieve on their own with a zero inheritance and these same rights. But added is a constraint that the date  $t$  agent must choose a consumption-output plan for his descendants that induces them to truthfully reveal their types. Once  $V_t^\ell(\theta_t)$  is thus calculated for all  $t \geq 1$ , the laissez-faire allocation with private information then corresponds to the solution to the

$t = 0$  problem for all  $\theta_0$ . Note that nowhere in this definition of the private information laissez-faire allocation is a condition that a date  $t = 0$ ,  $\theta_0$  type agent be willing to truthfully reveal his type.

**Proposition 7.** *If  $\theta_0$  is publicly observable but  $\theta_t$ ,  $t \geq 1$  is private, the private information laissez-faire allocation is constrained Pareto efficient.*

*Proof.* The proof is essentially identical to the full information case. It is straightforward to show that the laissez-faire allocation with private information has (22) hold as an equality for each  $\theta_0$  type. (Otherwise, each  $\theta_0$  type could simply consume the extra resources himself without upsetting incentive compatibility.) So as with full information, under the private information laissez-faire allocation, the expected discounted dynastic consumption of each  $\theta_0$  type equals its expected discounted dynastic output.

Next suppose there exists a feasible incentive-compatible allocation that Pareto dominates the private information laissez-faire allocation. Again, that  $\beta > 0$  and  $\pi(\theta|\theta^-) > 0$  for all  $(\theta, \theta^-)$  implies that at least one first generation  $\theta_0$  type agent is strictly better off in this Pareto-improving allocation. Further, since the allocation is a Pareto improvement, it satisfies (23) for all  $s \geq 1$  and  $(\theta^{1,s})$ . Finally, since the allocation is incentive compatible, it satisfies (58) for all  $s \geq 1$  and  $(\theta^{1,s})$ . Thus, it must be the case that for this first generation  $\theta_0$  type, (22) is violated, otherwise the first generation  $\theta_0$  type would have chosen it. That is, under the Pareto-improving allocation, the expected discounted dynastic consumption of at least one  $\theta_0$  type exceeds its expected discounted dynastic output. The society-wide resource constraint then implies that for at least one  $\theta_0$  type, say  $\hat{\theta}_0$ , its expected discounted dynastic output exceeds its expected discounted dynastic consumption. Further, since the new plan is an incentive-compatible Pareto improvement, conditions (23) and (58) are satisfied for the  $\hat{\theta}_0$  type's problem. Thus, the  $\hat{\theta}_0$  type could have chosen his part of the Pareto-improving allocation with (22) holding as a strict inequality, which is a contradiction. ■

Of course, a more natural assumption is that  $\theta_t$  is private for all dates  $t$ , including  $t = 0$ . It is not immediate, however, that a type  $\theta_0$  would prefer his own private information laissez-faire allocation over the private information laissez-faire allocation associated with type  $\hat{\theta}_0 \neq \theta_0$ . However, if a type  $\theta_0$  prefers the allocation chosen by a type  $\hat{\theta}_0$ , this implies that the allocation chosen by the  $\hat{\theta}_0$  type is not in his constraint set. (Otherwise, the  $\theta_0$  type would have chosen it.) But a date  $t = 0$  agent's type shows up in the constraint set of the laissez-faire allocation problem only in (22) — the condition that expected discounted consumption of an agent and his descendants not exceed the expected discounted production of the agent and his descendants. And in this condition, the date  $t = 0$  agent's type shows up only in term  $\pi(\theta_1|\theta_0)$ . Thus, the following proposition is immediate.

**Proposition 8.** *If  $\pi(\theta|\theta^-)$  is i.i.d., the private information laissez-faire allocation is constrained Pareto efficient.*

Next, we consider alternative assumptions that ensure that the private information laissez-faire allocation is constrained Pareto efficient. In particular, in the next proposition we assume only two  $\theta$  types (high and low) and that although high  $\theta$  types can claim to be low  $\theta$  types, the reverse is not true. In essence, for this proposition, we assume the existence of a skills test that high skill types can purposely fail if they wish to pretend to be the low skill type, but low skill types cannot pass.

**Proposition 9.** *Suppose  $\Theta = \{\underline{\theta}, \bar{\theta}\}$  and  $\beta R \leq 1$ . Then the private information laissez-faire allocation of the high  $\bar{\theta}_0$  is preferred by him to the private information laissez-faire allocation of the low  $\underline{\theta}_0$  type. Thus, if a low skill type cannot pretend to be high skill type, the private information laissez-faire allocation is constrained Pareto efficient.*

*Proof.* First, define the autarkic allocation  $(c^a(\theta), y^a(\theta))$  such that  $c^a(\theta) = y^a(\theta)$  and  $u'(y^a(\theta))\theta = h'(y^a/\theta)$ . The autarkic allocation is in the constraint set of  $\theta_0$  types. Next, let  $T = 1$  (so  $t \in \{0, 1\}$ )

and let  $\theta_0 = \underline{\theta}$ . Here, one can show that the laissez-faire allocation has the  $\theta_0 = \underline{\theta}$  type choosing the autarkic allocation. (Under the autarkic allocation,  $u'(c_0(\underline{\theta})) \leq \beta R u'(c_1(\underline{\theta}, \underline{\theta}))$  since  $c_0(\underline{\theta}) = c_1(\underline{\theta}, \underline{\theta}) = c^a(\underline{\theta})$  and  $\beta R \leq 1$ . Likewise,  $u'(c_0(\underline{\theta})) < \beta R u'(c_1(\underline{\theta}, \bar{\theta}))$  since  $c_0(\underline{\theta}) = c^a(\underline{\theta}) < c_1(\underline{\theta}, \bar{\theta}) = c^a(\bar{\theta})$  and  $\beta R \leq 1$ . Thus, the  $\theta_0 = \underline{\theta}$  type would, at the margin, weakly prefer to make negative conditional transfers to both son types, which is ruled out by (23). Thus, autarky (or no conditional transfers) satisfies his optimization problem.) But next note that the autarkic allocation is in the constraint set of the high  $\theta_0 = \bar{\theta}$  type as well, ensuring that the private information laissez-faire allocation chosen by the  $\theta_0 = \bar{\theta}$  type is weakly preferred by him to autarky, proving the result for  $T = 2$ .

Next, let  $T = 3$ . Here again, one can show that the private information laissez-faire allocation has the  $\theta_0 = \underline{\theta}$  type choosing zero transfers to both son types. Here, the  $\theta_0 = \underline{\theta}$  type chooses the continuation allocation for his  $\underline{\theta}$  son to be the autarkic allocation and chooses the continuation allocation for his  $\bar{\theta}$  son to correspond to the  $T = 2$  private information laissez-faire allocation. (This is in the constraint set of the  $\theta_0 = \underline{\theta}$  type and has  $u'(c_0(\underline{\theta})) \leq \beta R u'(c_1(\underline{\theta}, \theta_1))$  for both  $\theta_1$  types; thus again, the  $\theta_0 = \underline{\theta}$  type weakly prefers to make negative conditional transfers.) The argument then iterates backward. By noting that if the  $\theta_0 = \underline{\theta}$  type chooses zero conditional transfers to both son types, then the private information laissez-faire allocation for  $\theta_0 = \underline{\theta}$  is in the constraint set of the  $\theta_0 = \bar{\theta}$ , and the proof is complete. ■

## 6. Extensions

In the previous sections, agents were assumed to live for one period and care about their descendants. In reality, however, individuals do not live one day but many, and in each of these days some relevant information about their type is realized or revealed. Thus, an individual after one additional day of his life has a different history from the one he had until yesterday. Should this individual be considered, from the point of view of the efficiency of the allocations, a different

individual from the self he was yesterday? To formally analyze this question, we introduce a simple extension of the model in which an individual lives for two periods and cares about his descendants. (Later, it will become clear how to extend this to  $N$  periods.) Further, we assume that the physical attributes of the environment are unchanged and that preferences over allocations (equation (5)) are also unchanged. To link this current extended model to the previous one, in this section we consider the effects of assuming that for an agent born at  $t = 0$ , date  $t = 1$  refers not to the lifetime of his son but rather to the second half of his own life. Likewise, date  $t = 2$  refers to the first half of his son's life, date  $t = 3$  refers to the second half of his son's life, and so on.

In this construction, it follows immediately that if the preference ordering of an agent's future self is considered for the purposes of Pareto rankings exactly as his sons were in the previous sections, nothing changes. The set of Pareto efficient allocations is the same as when agents live for one period, and the implications for inheritance taxes and the efficiency of laissez-faire are also unchanged, with or without private information.

However, if for the purposes of Pareto rankings we consider only an agent's preference ordering at birth, then the set of Pareto efficient allocations is a strict subset of the set of Pareto efficient allocations when agents live for only one period. Nevertheless, our main results, modified versions of Propositions (1) through (9), go through.

Specifically, recall that under full information and agents living for one period, in Proposition (1) we show that a symmetric allocation is Pareto efficient if and only if the resource constraint (4) holds with equality, that labor is not distorted (condition (12) for all  $t$  and  $\theta^t$ ), and for all  $t$  and  $\theta^{t+1}$ ,

$$u'(c_t(\theta^t)) \geq \beta R u'(c_{t+1}(\theta^t, \theta_{t+1})). \quad (58)$$

When agents live for two periods and only their beginning of life preferences are considered for Pareto rankings, this result is unchanged except that (58) holds only for *odd* dates  $t$ . In words, this

requires that no father, in his second period of life, has an unsatisfied desire to give a positive gift to any particular type son at an actuarially fair rate of return. For *even* dates  $t$ , the inequality in (58) is replaced with an equality. This condition requires that no father, in his first period of life, has an unsatisfied desire to make either positive or negative conditional transfers to his second period of life, again at an actuarially fair rate of return.

Thus modified, Proposition (1) now implies, under full information that an agent's consumption in the second period of life does not depend on his second period  $\theta$  realization; he insures against this. Further, the implicit *savings tax* (the implicit inheritance tax in the previous sections but for even dates  $t$ ) must be zero in any Pareto efficient allocation, whereas the implicit inheritance taxes (or "wedges" for odd dates  $t$ ) are again non-positive and can be progressive or regressive. That is, Proposition (2) continues to hold, but only at odd dates  $t$ . Proposition (3), that the full information laissez-faire allocation is efficient, holds as well, but in modified form. The modification here is that in the definition of the  $t$  date laissez-faire allocation, condition (23) (that agents cannot give descendants born in date  $s$  allocations that to them are worse than the date  $s$  laissez-faire allocation) now must hold only for even dates  $s$ . Finally, with similar modifications, Propositions (6) through (9), which characterize the constrained efficiency of the private information laissez-faire allocation, hold as well.

In this extension, we altered the lifetime of an agent from one period to two. The extension to  $N$  is straightforward. But we have not changed the length of a time period (although our results hold for all  $\beta$ ,  $\pi$ , and  $R$ ). A second extension is to set up our model so that the length of a time period can be brought to zero. We then consider the extreme assumption that an individual lives for only an instant, and at every instant a new person is born with rights in determining Pareto rankings. Even with this extreme assumption, we show that the set of Pareto efficient allocations is non-trivial.



To this end, let  $\Delta$  be the length of a time period and assume that agents discount the future by the continuous time parameter  $b$  so that  $\beta(\Delta) = \exp(-b\Delta)$ . Likewise, if  $r$  is the continuous time analogue to  $R$ , let  $R(\Delta) = \exp(r\Delta)$ . Finally, assume that  $\theta \in \{\underline{\theta}, \bar{\theta}\}$ , let  $\underline{\lambda}$  represent the constant Poisson probability of transiting from state  $\underline{\theta}$  to state  $\bar{\theta}$ , and let  $\bar{\lambda}$  represent the constant Poisson probability of transiting from state  $\bar{\theta}$  to state  $\underline{\theta}$ . Then  $\pi(\underline{\theta}|\underline{\theta})(\Delta) = 1 - \underline{\lambda}\Delta$ ,  $\pi(\bar{\theta}|\underline{\theta})(\Delta) = \underline{\lambda}\Delta$ ,  $\pi(\underline{\theta}|\bar{\theta})(\Delta) = 1 - \bar{\lambda}\Delta$ , and  $\pi(\bar{\theta}|\bar{\theta})(\Delta) = 1 - \bar{\lambda}\Delta$ . With  $\beta$ ,  $R$ , and  $\pi$  thus defined as functions of  $\Delta$ , we can consider sequences of economies with  $\Delta$  converging to zero.

At each point in such a sequence, each of our propositions hold. For instance, the full information efficiency characterization Proposition (1) has that for any Pareto efficient allocation, resources must be exhausted (condition (4) holds as an equality) and the intratemporal labor-consumption efficiency condition (12) must hold; thus these must hold in the limiting economy as well. Regarding intertemporal consumption paths, we have (analogous to (58)) for all  $t$  and  $\theta^{t+\Delta}$

$$u'(c_t(\theta^t)) \geq \exp((r - b)\Delta)u'(c_{t+\Delta}(\theta^{t+\Delta})). \quad (59)$$

Heuristically, as  $\Delta \rightarrow 0$  (assuming  $\dot{c}_t$  is well defined), then

$$\dot{c}_t \geq (r - b)/\left(-\frac{u''(c_t)}{u'(c_t)}\right). \quad (60)$$

In words, when every instant a new person is realized for the purposes of Pareto ranking allocations, a necessary condition for the stochastic process  $c_t$  to be Pareto efficient in the full information economy is that the growth rate of consumption, realization by realization, is bounded below by the difference between technological and preference rates of discount, divided by the absolute level of risk aversion at that consumption level. If  $r = b$  (or  $\beta R = 1$  for all  $\Delta$ ), this condition implies that all consumption paths must be non-decreasing over time.

## 7. Conclusion

In this paper, we have introduced a natural method of considering Pareto efficiency in dynamic economies with altruism toward children and uncertainty regarding their skill. We argued that efficiency requires taking into account the desire of individuals to insure against uncertainty, but only against the real uncertainty they face — for example, that of fathers regarding the skill endowment of their children. The same concept does not require individuals to insure against events that have already happened, and this characteristic in fact defines who they are. For example, there is no uncertainty on the child's part regarding the realization of his skill. If we look at the characterization of efficiency in terms of the weights on individuals used by the social planner, our definition is equivalent to the one in which the weights over an individual may depend on the history of type realizations of his dynasty until his birth date.

We show that by adopting this natural definition, we substantially extend the set of Pareto efficient allocations. That the true set of efficient allocations is strictly larger is easy to see in the case of complete information economies. The set of efficient allocations for these economies can be completely characterized. An allocation is efficient in this case if and only if every father would not want to make, at that allocation, direct transfers to any of his child's type realizations. Although this is a substantial extension of the efficient set, substantial restrictions are still imposed by the efficiency requirement, and they are naturally induced by the altruism of the father.

In general we show that, properly understood, Pareto efficiency does not require special forms of government intervention, such as progressive inheritance taxes, to correct market failures. In other words, progressive taxation cannot be justified on efficiency grounds alone. Even a total absence of government intervention may result in efficient allocations.

## 8. Appendix: Continuous Time Model

The economy lasts for an interval  $[0, T]$ . Individuals are born at every integer time in the interval, live for one time unit, and have a single offspring who lives in the next period. At every time  $t$ , the history of the person is the path of types of his forefathers and his own path up to that moment.

The set of types is a finite set  $\Theta$ . The set of type paths is the set  $\Omega$  of functions from  $[0, T]$  to  $\Theta$  that are right continuous and piecewise constant. Let

$$\lambda = \text{diag}(\lambda_{\theta^1}, \dots, \lambda_{\theta^n}) \quad (61)$$

with  $\lambda_{\theta^i} \geq 0$  be the Poisson rate in the state  $\theta^i$ . Let  $\pi$  be a stochastic matrix, with  $\pi(\cdot|\theta^i) \in \Delta(\Theta)$  the transition probability at  $\theta^i$  if a transition occurs. The process on  $\Theta$  is defined as follows. At state  $\theta^i$ , a change in the state occurs at Poisson rate  $\lambda_{\theta^i}$ ; if a change occurs, the state transits to  $\theta^j$  with probability  $\pi(\theta^j|\theta^i)$ . The probability  $P$  on  $(\Omega, \mathcal{F})$  is induced by this process.

Note that  $\lambda$  and  $\pi$  define an invariant measure on  $\Theta$  as the solution  $\mu$  of

$$\mu(I - \lambda\pi) = 0, \mu(\{\Theta\}) = 1, \quad (62)$$

where  $I$  is the identity matrix. At every time  $t$ , the probability that the state is  $\theta^i$  is  $\mu(\theta^i)$ .

**Definition 1.** *An allocation  $a$  is a pair  $(c, y)$  of functions from  $\Omega \times [0, T]$  to the set  $C \times Y$  of consumption and outputs.*

Note that an allocation defines consumption and output for the  $n^{\text{th}}$  member of the dynasty living in the time interval  $[n, n + 1)$ . We call this member a person. An individual in this economy is identified by the history until the point in time at which the individual appears, namely, the path  $\omega$  restricted to  $[0, t]$ . Thus, each person has a continuum of individual realizations.

Let  $r$  and  $b$  be positive numbers. The technology available allows transfer of the accumulated asset at the rate  $e^{rs}$ . An allocation is feasible if

$$E \int_{t=0}^T e^{-rs} [y(\cdot, s) - c(\cdot, s)] ds \geq 0 \quad (63)$$

Utility is discounted at the rate  $e^{-bs}$ . We impose

$$\lim_{t \rightarrow +\infty} u(e^{rt})e^{-bt} = C < \infty. \quad (64)$$

Consider now a small time interval  $\Delta t$ . For each such interval, consider the set of allocations that are constant in the time interval  $[i\Delta t, (i+1)\Delta t)$  for every  $i$ . Utility is not constant over such interval because the type can change, as can the disutility cost of effort. For such a time interval, we have a discrete economy as defined in the main text. For such an economy, feasibility and efficiency have been defined.

**Definition 2.** *An allocation is efficient if it is the limit of  $\Delta t$ -efficient allocations.*

**Lemma 3.** *If condition (64) holds the set of efficient allocations is non-empty.*

**Theorem 1.** *An allocation  $a \equiv (c, y)$  is efficient if and only if:*

1. **Resource utilization** *For every  $t$ :*

$$E \int_{t=0}^T e^{-rs} [y(\cdot, s) - c(\cdot, s)] ds = 0 \quad (65)$$

2. **No wedge**

$$P - a.e. \ \omega, u'(c(\omega, t))\theta(\omega, t) = h' \left( \frac{y(\omega, t)}{\theta(\omega, t)} \right) \quad (66)$$

3. **No transfer**

$$P - a.e. \ \omega, \forall t, \liminf_{\Delta t \rightarrow 0} \frac{c(\omega, t + \Delta t) - c(\omega, t)}{\Delta t} \geq -(r - b) \frac{u''(c(\omega, t))}{u'(c(\omega, t))} \quad (67)$$

## References

- [1] Fahri, Emmanuel and Iván Werning. “Inequality and Social Discounting.”, *Journal of Political Economy*, 115, 3, 365–402, 2007.
- [2] Fahri, Emmanuel and Iván Werning. “Progressive Estate Taxation.”, *Quarterly Journal of Economics*, 635–673, 2010.
- [3] Golosov, Mikhail, Narayana Kocherlakota, and Aleh Tsyvinski, Optimal Indirect and Capital Taxation, *Review of Economic Studies*, 70 (2003), 569-587.
- [4] Harsanyi, John, Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility, *Journal of Political Economy* 63 (1955): 309-321.
- [5] Mirrlees, James A., An Exploration in the Theory of Optimum Income Taxation, *Review of Economic Studies*, 38 (1971), 175-208.
- [6] Mirrlees, James A., Optimal Tax Theory: A Synthesis, *Journal of Public Economics*, 6 (1976), 327-358.
- [7] Phelan, Christopher, Opportunity and Social Mobility, *Review of Economic Studies*, 73 (2006), 487-505.
- [8] Rawls, John, *A Theory of Justice*, Cambridge, MA: Harvard University Press, 1971.
- [9] Rogerson, William P., Repeated Moral Hazard, *Econometrica*, 53 (1985), 69-76.