Accidental Bequests: A Curse for the Rich and a Boon for the Poor

HELMUTH CREMER, FIROUZ GAHVARI AND PIERRE PESTIEAU

Accidental bequests: a curse for the rich and a boon for the poor[*]

Helmuth Cremer
Toulouse School of Economics
(IDEI, University of Toulouse and Institut universitaire de France)
31000 Toulouse, France

Firouz Gahvari
Department of Economics
University of Illinois at Urbana-Champaign
Urbana, IL 61801, USA

Pierre Pestieau
CREPP, University of Liège and CORE
4000 Liège, Belgium

November 2009

Abstract

When accidental bequests signal otherwise unobservable individual characteristics such as productivity and longevity, the tax administration should partition the population into two groups: One consisting of people who do not receive an inheritance and the other of those who do. The first tagged group gets a second-best tax à la Mirrlees; the second group a first-best tax schedule. The solution implies that receiving an inheritance makes high-ability types worse off and low-ability types better off. High-ability individuals will necessarily face a bequest tax of more than 100%, while low-ability types face a bequest tax that can be smaller as well as larger than 100%. With a Rawlsian social welfare function, the low-ability types too face a more than 100% tax on bequests.

1 Introduction

One striking feature of the recent evolution of tax policy around the world has been the mounting unpopularity of wealth transfer taxation. An increasing number of countries are without inheritance (or estate) taxes and some, including the US, are contemplating ways to phase them out. At the same time, countries that continue to tax wealth transfers raise little revenue from them, despite the fact that the transmitted wealth is increasing.[1] In academic circles, there is little agreement on how these taxes should be structured. Many believe this issue is part of the broader question of taxation of wealth and wealth income, and must be analyzed in that context. We agree with Kopczuk (2009) that the question is best addressed in the most general framework of designing the economy's overall tax policy.

The controversies surrounding estate taxes reflect, to a great extent, disagreement over bequest motives. The literature has highlighted a number of reasons for bequests: pure altruism, warm glow, strategic (deferred payment for filial attention), and accidental.[2] Yet there is little agreement regarding the relative strength of each of theses elements. Studies suggest that the size of the estate tax is highly sensitive to the relative importance of different bequest motives.[3]Not surprisingly, then, one's view on the question of motive shapes also one's view on the appropriate tax treatment of bequests.

While little agreement exists concerning taxation of bequests in general, there is a widely-held view that accidental bequests should be subjected to a confiscatory tax. Kaplow (2008, pp 264–66) gives a lucid exposition of the reasons behind this view. Seen from the perspective of offsprings, Kaplow argues, full insurance in the face of an uncertain inheritance income is optimal. A 100% bequest tax, with its proceeds rebated equally to all children, leads to no distortions and mimics this insurance scheme.

---

[1]The proceeds do not typically exceed one percent of the aggregate tax revenues; see OECD (2008).

[2]When annuity markets are imperfect, most people will have a positive wealth at death, even when they have no bequest motive per se; see Cremer and Pestieau (2005).

[3]See, e.g., Michel and Pestieau (2002) and Pestieau and Sato (2008).

Alternatively, seen from the parents' perspective, the optimal scheme is one of perfect annuities. Kopczuk (2003) points out that this market too, under certain conditions, can be mimicked by a government policy which includes 100% bequest taxes. To effect the first-best solution, one must give the retirees a wealth supplement and then fully tax the wealth of those who die early. However, the availability of wealth supplements rests on the individuals' having strong bequest motives.

Blumkin and Sadka (2003) is, to our knowledge, the only paper that questions the wisdom of the received view in the context of a second-best tax policy design. They argue that a non-confiscatory tax on accidental bequests has the desirable consequence of making the demogrant of an optimal linear income tax system effectively non-uniform. In this sense, it will act as an additional instrument and increases the efficiency of the tax system. Our challenge to the 100% tax idea is more basic. In particular, its rationalization does not rest on having a tax system that disregards useful information. Quite the opposite! The rationalization comes from Mirrlees' (1971) teaching that optimal tax systems must be designed on the basis of all the information that the tax administration possesses. Discarding useful information leads to a suboptimal tax structure.

The fundamental point that we make is that bequests, when publicly observable, have informational content. This content must be incorporated in the design of optimal tax structures. The point is a general one and applies to all types of bequests regardless of bequest motives. However, to make our case as stark as possible, we concentrate exclusively on accidental bequests. Circumventing the identification of bequest motives, simplifies the analysis drastically and highlights our point most efficiently. Our starting benchmark requires that, within the context of the tax system as a whole, and when tax instruments are not artificially restricted, ignoring the informational content of bequests leads to a 100% tax rate on bequests.[4] We show that this result is unwarranted when the

---

[4]The tax neither distorts the behavior of parents nor their utility as they have no bequest motive. The behavior of recipients is not distorted either as accidental bequests are a windfall for them. Moreover, as

informational content of accidental bequests is taken into account. In many instances, the optimal bequest tax exceeds 100%, although there are also circumstances that call for a less than 100% tax.

The paper considers a setting where accidental bequests signal otherwise unobservable individual characteristics such as productivity and longevity. The tax administration then uses an individual's inheritance as a separation mechanism, or a "tag," when designing an optimal tax system. The population is partitioned into two groups consisting of people who do not receive an inheritance and of those who do. The first group faces a tax schedule determined on the basis of Mirrlees' standard optimal non-linear income tax problem. The second group, on the other hand, faces a non-distortionary tax scheme. The most interesting results that emerge are that the high-ability types will be worse off and the low-ability types better off if they receive an inheritance. Moreover, the high-ability types will face a bequest tax that necessarily exceeds 100% but that the low-ability types face a tax rate that can be smaller or larger than 100%. With a Rawlsian social welfare function, the low-ability types too should face a more than 100% tax on bequests.

2 The setting

2.1 Basic model

Consider a two-period overlapping-generation model wherein individuals of each generation live either for one or two periods. Regardless of their longevity, they work in the first period only. Those whose parents die early (*i.e.*, live for one period) receive an inheritance from their parents; those whose parents live for two periods receive nothing. There are no annuity markets. All individuals, at the beginning of period one, allocate

---

long as the inheritance tax they pay is determined as part of their total tax liabilities, the recipients will not become any worse off. This follows because the existence of a non-distortionary source of revenue reduces the amount of the distortionary tax that the government needs to raise. The optimal *allocations* of different households must be independent of who initially owns the non-distortionary revenue sources.

3

their resources—earnings plus any inheritances—between present consumption and saving to be consumed when retired. If individuals stay alive in the second period, they will consume all their savings and leave no bequests; if they die early, their unused saving is transmitted to their children as an accidental bequest. Saving is channeled into future consumption through a storage technology; there is no appreciation or depreciation of savings so that the interest rate is zero.

Individuals differ in their productivity w_i, their survival probability π_i, and taste for future versus present consumption. To model the taste difference, we assign a weight β_i to future consumption in the individuals' utility functions. We assume that each of these characteristics take only two values: "high" indexed by h and "low" indexed by ℓ. We further assume that these characteristics are positively correlated so that there are only two types of people: h and ℓ with $w_h > w_\ell, \pi_h > \pi_\ell$, and $\beta_h > \beta_\ell$.[5] Types are dynastically immutable: if a person is of type i, his offsprings will also be of type i. There is no population growth and each generation consists of n_i individuals of type i, where $n_h + n_\ell = 1$.

Individuals have additive quasi-linear preferences over present consumption, c_i, future consumption, d_i, and labor supply, L_i. An individual's expected utility is given by

$$
\begin{aligned}
U_i &= \pi_i \left[c_i + \beta_i \phi\left(d_i\right) - \varphi\left(L_i\right) \right] + \left(1 - \pi_i\right) \left[c_i - \varphi\left(L_i\right) \right] \\
&= c_i + \pi_i \beta_i \phi\left(d_i\right) - \varphi\left(L_i\right), \quad i = h, \ell,
\end{aligned}
\tag{1}
$$

where ϕ is strictly concave while φ is strictly convex.[6]

Individuals of the first generation start life with no initial wealth. This will not be the case for members of the generations that follow, however. Those who die early leave

---

[5]It is sufficient to assume $\pi_h \geqq \pi_\ell$ and $\beta_h \geqq \beta_\ell$ with one of the two inequalities being strict. Preston (1975), and Pritchett and Summers (1996), provide empirical support for the existence of a positive correlation between ability and longevity; and Bommier (2006) for a positive correlation between longevity and preferences for future over present consumption.

[6]This formulation assumes that an individual derives no utility from leaving an accidental bequest for his family if he dies early.

an accidental bequest which, unless taxed away, is inherited by their children. Consequently, besides w_i, π_i and β_i, individuals of the second and forthcoming generations will differ also on the basis of their inherited wealth, ω_i. The quasi-linearity of the utility function (1) ensures that the size of one's (accidental) bequest to his children is unaffected by the size of the inheritance that he may have received from his own parents (including zero). We will then have, in each period, four groups of people: h-types with either ω_h or no inherited wealth and ℓ-types with either ω_ℓ or no inherited wealth.[7]

2.2 Laissez-faire

Recall that there is no annuity market and that private saving, s_i, is the only source for financing one's consumption during retirement years. The optimization problem of an i-type individual $(i = h, \ell)$ in the first generation is

$$\max_{s_i, L_i} \quad U_i = w_i L_i - s_i + \pi_i \beta_i \phi\left(s_i\right) - \varphi\left(L_i\right), \tag{2}$$

where we have substituted $w_i L_i - s_i$ for c_i, and s_i for d_i, in the individual's expected utility given by equation (1). The optimization yields $\pi_i \beta_i \phi'\left(s_i\right) = 1$, or $s_i = \phi'^{-1}\left(1/\pi_i \beta_i\right)$, and $w_i = \varphi'\left(L_i\right)$. These relationships, along with the assumptions $\pi_h > \pi_\ell$, $\beta_h > \beta_\ell$, strict concavity of $\phi\left(\cdot\right)$, and strict convexity of $\varphi\left(\cdot\right)$, imply that $s_h > s_\ell$ and $L_h > L_\ell$. Finally, observe that since in the absence of annuity markets every i-type person saves s_i to finance his future consumption, those who die early must leave an accidental bequest of s_i behind. That is, $\omega_i = s_i$.

The optimization problem of individuals belonging to second and forthcoming generations depends on whether they inherit an initial wealth or not. Those who receive no inheritance have an identical optimization problem to that of the first generation. This

---

[7]Dependence of the bequest one leaves on the inheritance one receives leads to multiplicity of groups on the basis of ω_i, making existence of a steady state problematic. Observe also that even with quasi-linear preferences ω_i may take more than two values if the prices (tax rates) that individuals face depend on their inheritance status. We discuss this issue below when addressing second-best allocations and their implementation.

continues to be summarized by (2), resulting in the same solution as those obtained for the first generation. In particular, a second-generation i-type with no inheritance will save the same amount as a first-generation i-type. It then also follows that any bequest left by a second-generation i-type with no inheritance will be identical to that of the first-generation i-type: $\omega_i = s_i$.

Second-generation individuals who inherit an initial wealth have an optimization problem summarized by

$$\max_{\widehat{s}_i, \widehat{L}_i} \quad \widehat{U}_i = w_i \widehat{L}_i + \omega_i - \widehat{s}_i + \pi_i \beta_i \phi\left(\widehat{s}_i\right) - \varphi\left(\widehat{L}_i\right), \tag{3}$$

where the symbol ˆ over a variable indicates that it refers to a person who has received an inheritance, and we have substituted $w_i \widehat{L}_i + \omega_i - \widehat{s}_i$ for $\widehat{c}_i$ and $\widehat{s}_i$ for $\widehat{d}_i$. The quasi-linearity of preferences implies that future consumption and labor supply do not depend on inherited wealth so that $\widehat{s}_i = s_i$ and $\widehat{L}_i = L_i$. However, present consumption increases with wealth and $\widehat{c}_i > c_i$. Clearly, with $\widehat{s}_i = s_i$, $\widehat{\omega}_i = \omega_i$ and equal to the bequest of an i-type of the first generation. Finally, with $s_h > s_\ell$ and $L_h > L_\ell$, it thus also follows that $\widehat{s}_h > \widehat{s}_\ell$ and $\widehat{L}_h > \widehat{L}_\ell$.

The quasi-linearity assumption ensures that, starting with generation two, the economy is in a stationary-state equilibrium. After that, the equilibrium values of all the variables remain invariant to time. Specifically, in every period, there will always be n_ℓ unskilled and n_h skilled individuals. Of the unskilled workers, $n_\ell(1 - \pi_\ell)$ have an initial wealth equal to $\omega_\ell = s_\ell$ and the remaining $n_\ell \pi_\ell$ have no wealth; of the skilled workers, $n_h(1 - \pi_h)$ have an initial wealth $\omega_h = s_h$ and $n_h \pi_h$ have no wealth. And, with $s_h > s_\ell$, it is also the case that $\omega_h > \omega_\ell$.

3   First-best

3.1   Allocation

Assume there is full information. In particular, individual types $i = h, \ell$, as well as the size of their inherited wealth (whether zero or ω_i), are publicly observable. The first-best policy is attained when the government chooses (c_i, d_i, L_i) and $\left(\widehat{c}_i, \widehat{d}_i, \widehat{L}_i\right)$ to maximize social welfare defined by

$$W = \sum_{i=h,\ell} n_i \left[\pi_i v(U_i) + (1 - \pi_i) \, v(\widehat{U}_i) \right], \tag{4}$$

where $v\left(\cdot\right)$ is a strictly concave transformation of the quasi-linear utility function (1); hence $v'\left(\cdot\right) > 0$ and $v''\left(\cdot\right) < 0$. Aside from this transformation, the social welfare function defined by (4) is utilitarian in form in that it aggregates the utilities of the four groups—h- and ℓ-types, each with and without an inheritance—and assigns each a weight according to their numbers in the society. The role of $v\left(\cdot\right)$ is to make the social welfare function redistributive. Without such a transformation, there will be no aversion to inequality and thus no redistribution. Observe also that the more concave $v\left(\cdot\right)$ is the more redistributive the social welfare function will be. One common specification for $v\left(\cdot\right)$ is the iso-elastic case $v(U) = U^{1-\varepsilon}/\left(1 - \varepsilon\right)$, suggested by Atkinson (1973). In this formulation, $\varepsilon > 0$ $(\varepsilon \neq 1)$ denotes the inequality aversion index and the higher is ε the greater will be the desired redistribution.

The resource constraint for the economy is given by

$$\sum_{i=h,\ell} n_i \left[\pi_i \left(w_i L_i - c_i - \pi_i d_i \right) + (1 - \pi_i) \left(w_i \widehat{L}_i - \widehat{c}_i - \pi_i \widehat{d}_i \right) \right] \geqq 0. \tag{5}$$

The specification of this constraint implies that the resources available to any generation are spent in full, with the inheritances that any generation receives offsetting the bequests it leaves.[8] First-best optimum is then characterized by choosing (c_i, d_i, L_i)

---

[8] While inheritances one generation receives may in principle be different from the bequests it leaves, such an outcome cannot happen in the steady state.

and $\left(\widehat{c}_i, \widehat{d}_i, \widehat{L}_i\right)$ to maximize (4) subject to (5).[9] Let μ denote the Lagrangian multiplier associated with the resource constraint (5). The Lagrangian expression associated with this problem is

$$
\begin{aligned}
\mathcal{L} \;=\; & \sum_{i=h,\ell} n_i \Bigg\{ \pi_i \left[v(c_i + \pi_i \beta_i \phi(d_i) - \varphi(L_i)) + \mu(w_i L_i - c_i - \pi_i d_i) \right] + \\
& (1 - \pi_i) \left[v\left(\widehat{c}_i + \pi_i \beta_i \phi\left(\widehat{d}_i \right) - \varphi\left(\widehat{L}_i \right) \right) + \mu\left(w_i \widehat{L}_i - \widehat{c}_i - \pi_i \widehat{d}_i \right) \right] \Bigg\}.
\end{aligned}
\tag{6}
$$

It follows from the first-order conditions of the above problem that, for $i = h, \ell$,

$$
v'(U_i) = v'\left(\widehat{U}_i\right) = \mu,
\tag{7}
$$

$$
\beta_i \phi'(d_i) = \beta_i \phi'\left(\widehat{d}_i\right) = 1,
\tag{8}
$$

$$
\varphi'(L_i) = \varphi'\left(\widehat{L}_i\right) = \mu w_i.
\tag{9}
$$

Denoting the first-best values by superscript FB, it follows from (8)–(9) that

$$
\widehat{U}_i = U_i \equiv U^{FB},
\tag{10}
$$

$$
\widehat{d}_h = d_h \equiv d_h^{FB} > \widehat{d}_\ell = d_\ell \equiv d_\ell^{FB},
\tag{11}
$$

$$
\widehat{L}_h = L_h \equiv L_h^{FB} > \widehat{L}_\ell = L_\ell \equiv L_\ell^{FB},
\tag{12}
$$

where the first inequality sign follows from $\beta_h > \beta_\ell$ and the strict concavity of $\phi(\cdot)$, and the second from $w_h > w_\ell$ and the strict convexity of $\varphi(\cdot)$. Observe also that equations (10)–(12) imply

$$
\widehat{c}_i = c_i \equiv c_i^{FB},
$$

so that, at the first-best, the difference in type affects one's allocation but not the difference in inheritance status. Put differently, allocations of h- and ℓ-types differ but either one gets the same allocation regardless of receiving an inheritance or not.

---

[9]There is also the constraint that $s_i = w_i L_i - c_i \geq 0$, which we assume to be non-binding.

3.2 Tax policy

We now show that the government is able to decentralize the first-best allocations through a combination of saving subsidies and lump-sum taxes. Saving subsidies need only be conditioned on types, but not on inherited wealth, and set at a rate equal to

$$\tau_i = 1 - \pi_i, \tag{13}$$

for type $i = h, \ell$. Lump sum taxes, which can be negative as well as positive, on the other hand, must be conditioned on types as well as inherited wealth: $\left(\widehat{t}_h, \widehat{t}_\ell\right)$ for those with an inheritance and (t_h, t_ℓ) for those without. Naturally, these fiscal instruments must also satisfy the government's budget constraint

$$\sum_{i=h,\ell} n_i \left[\pi_i \left(t_i - \tau_i s_i\right) + (1 - \pi_i)\left(\widehat{t}_i - \tau_i \widehat{s}_i\right)\right] = 0.$$

To see how the optimum is decentralized, observe that the presence of $\tau_i, \widehat{t}_i$, and t_i changes the budget constraints of the i-type with and without an inheritance to

$$\begin{aligned}
c_i + (1 - \tau_i) s_i &= w_i L_i + \omega_i - \widehat{t}_i, \\
c_i + (1 - \tau_i) s_i &= w_i L_i - t_i.
\end{aligned}$$

The first-order conditions for the i-type's optimization problem, with or without an inheritance, yield $\varphi'(L_i) = w_i$ and

$$\pi_i \beta_i \phi'(s_i) = 1 - \tau_i = \pi_i.$$

These are identical to their first-best counterparts resulting in $\widehat{L}_i = L_i = L_i^{FB}$ and $\widehat{s}_i = s_i = s_i^{FB}$. Additionally, to ensure $\widehat{c}_i = c_i = c_i^{FB}$, one must set lump-sum taxes at the rates $t_i = w_i L_i^{FB} - c_i^{FB} - \pi_i s_i^{FB}$ and $\widehat{t}_i = w_i L_i^{FB} + \omega_i - c_i^{FB} - \pi_i s_i^{FB}$.

Finally, observe that the expressions for $\widehat{t}_i$ and t_i show that

$$\widehat{t}_i = t_i + \omega_i.$$

9

Now, given identical tax rates on savings, it is natural to consider the difference between $\widehat{t}_i$ and t_i as the tax on bequests. Alternatively, one can consider the bequest tax to be the difference between one's total tax liabilities when he receives an inheritance and when he does not. There is no tension between these definitions; however. Given identical savings under the two scenarios $(\widehat{s}_i = s_i)$, both definitions lead to the same answer:[10] The implementation of first-best allocations requires a 100% taxation of (accidental) bequests. The i-type with an inheritance ω_i, $i = h, \ell$, sees his entire inheritance "confiscated," after which he is treated like his counterpart with no inheritance.

The main results of this section are summarized in the following proposition.

Proposition 1 *Consider a society with two types of individuals $i = h, \ell$ whose preferences are defined by (1). The types are identified by their productivity w_i, their survival probability π_i, and the utility weight they assign to their consumption in retirement β_i. The type characteristics are dynastically immutable and satisfy the property $(w_h, \pi_h, \beta_h) > (w_\ell, \pi_\ell, \beta_\ell)$.*

(i) First-best allocations are characterized by (a) equalization of utilities of all individuals regardless of their type and inheritance status: $\widehat{U}_h = U_h = \widehat{U}_\ell = U_\ell$, and (b) a higher accidental bequest for individuals of type h: $\omega_h > \omega_\ell$.

(ii) Decentralization of first-best allocations requires: (a) Saving subsidies conditioned on types, but not on inherited wealth, set at a rate equal to $\tau_i = 1 - \pi_i$ for type i, and (b) lump sum taxes conditioned on types as well as inherited wealth: $(\widehat{t}_h, \widehat{t}_\ell)$ for those with an inheritance and (t_h, t_ℓ) for those without.

---

[10]Denote the net taxes an i-type pays, $i = h, \ell$, by T_i if he does not receives an inheritance and by $\widehat{T}_i$ if he does. It must then be the case that

$$
\begin{aligned}
T_i &= t_i - \tau_i s_i, \\
\widehat{T}_i &= \widehat{t}_i - \tau_i \widehat{s}_i.
\end{aligned}
$$

With $t_i = \widehat{t}_i - \omega_i$ and $s_i = \widehat{s}_i$, it also follows that

$$
T_i = \widehat{T}_i - \omega_i.
$$

(iii) All accidental bequests are taxed at 100%.

4 Second-best

4.1 Allocation

Define the second-best as a setting wherein individual types, *i.e.*, the characteristics (w_i, π_i, β_i), and labor supplies, L_i, are not publicly observable. The observables are gross earnings $(\widehat{I_i} = w_i \widehat{L_i}; I_i = w_i L_i;)$, consumption during working years and retirement $(\widehat{c_i}, \widehat{d_i}; c_i, d_i)$, and bequests ω_i. This follows the traditional information structure in optimal non-linear income tax models à la Mirrlees. The only difference is that we have added bequests as an observable variable and thus potentially taxable.[11] The key point that we make is that observability of bequests is sufficient to identify its recipient's type. On the other hand, the type of an individual who receives no inheritance remains unknown to the government. Given this information structure, the tax administration uses an individual's inheritance as a separation mechanism, or a "tag," when designing an optimal tax system.[12]

We start by assuming that the government is able to identify the type of a person who receives a bequest and then show that this is in fact the case. Given this assumption, the government proceeds to partition the population into two groups: Those who receive an inheritance and those who do not (tagged as "positive inheritance" and "zero inheritance"). The zero-inheritance group, consisting of people whose ability remains private information, will have to face a tax schedule determined on the basis of Mirrlees' standard optimal non-linear income tax problem. The positive-inheritance group, on the other hand, need not face a second-best tax schedule. This group consists of people whose characteristics can be inferred from the level of the inheritance they receive.

---

[11] Throughout the paper we assume that an individual's tax liability depends only on variables pertaining to the individual himself: income, consumption, and inheritances received. Socio-political considerations prevent the government to condition a person's tax liability on his parents' characteristics.

[12] Akerlof (1978) is the classic paper on tagging. Boadway and Pestieau (2006), and Cremer *et al.* (2010), are among the more recent contributions to this literature.

Hence a full information solution can be achieved within this group.

To describe the optimal tax policy, we first characterize the optimal allocation constrained by the information structure just sketched. As is commonly done in the literature on tagging, one can formulate the problem within each group independently; connecting the two via the economy's resource constraint.[13] Put differently, one assigns a single resource constraint to the two groups. This continues to be represented by (5). Let λ denote the Lagrange multiplier associated with the economy's resource constraint, and μ the multiplier associated with the incentive constraint in the group of individuals who receive no inheritance. The Lagrangian expression associated with this optimization problem is

$$
\mathcal{L} = \sum_{i=h,\ell} n_i \left\{ \pi_i \left[v\left(c_i + \pi_i \beta_i \phi\left(d_i\right) - \varphi(\frac{I_i}{w_i}) \right) + \mu\left(I_i - c_i - \pi_i d_i \right) \right] \right.
$$
$$
\left. + (1 - \pi_i) \left[v\left(\widehat{c}_i + \pi_i \beta_i \phi(\widehat{d}_i) - \varphi(\widehat{L}_i) \right) + \mu\left(w_i \widehat{L}_i - \widehat{c}_i - \pi_i \widehat{d}_i \right) \right] \right\}
$$
$$
+ \lambda \left[c_h + \pi_h \beta_h \phi\left(d_h\right) - \varphi(\frac{I_h}{w_h}) - c_\ell - \pi_h \beta_h \phi\left(d_\ell\right) + \varphi(\frac{I_\ell}{w_h}) \right]. \qquad (14)
$$

In writing (14), we have followed the common practice of writing the problem in terms of the (observable) pre-tax income for individuals in the zero-inheritance group (writing labor supply as I_i/w_i). People in the positive-inheritance group face no incentive constraint; their earning abilities are observable. Here, to stress the first-best nature of the problem within this group, we specify the decision variable to be $\widehat{L}_i$—an observable as in Section 3.[14] Note also that people in the zero-inheritance group pay a tax equal to $T_i = I_i - c_i - d_i$ (with d_i being left as an accidental bequest if they die early). Those in the positive-inheritance group, on the other hand, pay a tax equal to $\widehat{T}_i = \widehat{I}_i + \omega_i - \widehat{c}_i - \widehat{d}_i$. In the steady state $\omega_i = \widehat{d}_i$, and the expression for the tax payment of people with a positive inheritance simplifies to $\widehat{T}_i = \widehat{I}_i - \widehat{c}_i$.

Observe that bequests do not appear directly in the economy's resource constraint.

---

[13]See, *e.g.*, Cremer *et al.* (2010).

[14]Although one can equally express the problem in terms of $\widehat{I}_i$.

This follows because bequests are simply a transfer between generations, with the inheritances one generation receives offsetting the bequests it leaves. However, bequests appear indirectly in (14), being the defining characteristic of zero- and positive-inheritance groups. It is because of the information that bequests convey that (14) does not contain an incentive constraint for people in the positive-inheritance group.

The first-order conditions for $i = h, \ell$ are, for people with an inheritance:

$$\frac{\partial \mathcal{L}}{\partial \widehat{c}_i} = n_i (1 - \pi_i) \left[v' \left(\widehat{U}_i \right) - \mu \right] = 0, \tag{15}$$

$$\frac{\partial \mathcal{L}}{\partial \widehat{d}_i} = n_i (1 - \pi_i) \left[v' \left(\widehat{U}_i \right) \pi_i \beta_i \phi' \left(\widehat{d}_i \right) - \mu \pi_i \right] = 0, \tag{16}$$

$$\frac{\partial \mathcal{L}}{\partial \widehat{L}_i} = n_i (1 - \pi_i) \left[-v' \left(\widehat{U}_i \right) \varphi' \left(\widehat{L}_i \right) + \mu w_i \right] = 0, \tag{17}$$

and for those without an inheritance:

$$\frac{\partial \mathcal{L}}{\partial c_h} = n_h \pi_h \left[v' \left(U_h \right) - \mu \right] + \lambda = 0, \tag{18}$$

$$\frac{\partial \mathcal{L}}{\partial d_h} = n_h \pi_h \left[v' \left(U_h \right) \pi_h \beta_h \phi' \left(d_h \right) - \mu \pi_h \right] + \lambda \pi_h \beta_h \phi' \left(d_h \right) = 0, \tag{19}$$

$$\frac{\partial \mathcal{L}}{\partial I_h} = n_h \pi_h \left[-v' \left(U_h \right) \frac{1}{w_h} \varphi' \left(\frac{I_h}{w_h} \right) + \mu \right] - \frac{\lambda}{w_h} \varphi' \left(\frac{I_h}{w_h} \right) = 0, \tag{20}$$

$$\frac{\partial \mathcal{L}}{\partial c_\ell} = n_\ell \pi_\ell \left[v' \left(U_\ell \right) - \mu \right] - \lambda = 0, \tag{21}$$

$$\frac{\partial \mathcal{L}}{\partial d_\ell} = n_\ell \pi_\ell \left[v' \left(U_h \right) \pi_\ell \beta_\ell \phi' \left(d_\ell \right) - \mu \pi_\ell \right] - \lambda \pi_h \beta_h \phi' \left(d_\ell \right) = 0, \tag{22}$$

$$\frac{\partial \mathcal{L}}{\partial I_\ell} = n_\ell \pi_\ell \left[-v' \left(U_h \right) \frac{1}{w_\ell} \varphi' \left(\frac{I_\ell}{w_\ell} \right) + \mu \right] + \frac{\lambda}{w_h} \varphi' \left(\frac{I_\ell}{w_h} \right) = 0. \tag{23}$$

Consider first the group who receive an inheritance. Simplifying and rearranging equations (15)–(17) yields, for $i = h, \ell$,

$$v' \left(\widehat{U}_i \right) = \mu, \tag{24}$$

$$\beta_i \phi' \left(\widehat{d}_i \right) = 1, \tag{25}$$

$$\varphi' \left(\widehat{L}_i \right) = w_i. \tag{26}$$

This is the same characterization as in the first best indicating no distortion on either labor supply or future consumption. This is not surprising. We have a first-best solution for the people in positive-inheritance group because their characteristics are observable. Equations (24)–(26) also imply, as with the first-best solution, $\widehat{U}_h = \widehat{U}_\ell, \widehat{d}_h > \widehat{d}_\ell$, and $\widehat{L}_h > \widehat{L}_\ell$. That is, in the positive-inheritance group, the utility levels for high- and low-ability types are equalized, but that a high-ability type consumes more in the second period, and works more in the first, as compared with a low-ability type.

Turning to the group whose members do not receive an inheritance, equations (18)–(23) yield

$$v'\left(U_h\right) = \mu - \frac{\lambda}{n_h \pi_h}, \tag{27}$$

$$\beta_h \phi'\left(d_h\right) = 1, \tag{28}$$

$$\frac{1}{w_h} \varphi'\left(\frac{I_h}{w_h}\right) = 1, \tag{29}$$

$$v'\left(U_\ell\right) = \mu + \frac{\lambda}{n_\ell \pi_\ell}, \tag{30}$$

$$\beta_\ell \phi'\left(d_\ell\right) = 1 + \frac{\lambda \phi'\left(d_\ell\right)}{\mu n_\ell \pi_\ell} \left[\beta_h \frac{\pi_h}{\pi_\ell} - \beta_\ell\right], \tag{31}$$

$$\frac{1}{w_\ell} \varphi'\left(\frac{I_\ell}{w_\ell}\right) = 1 + \frac{\lambda}{\mu n_\ell \pi_\ell} \left[\frac{1}{w_h} \varphi'\left(\frac{I_\ell}{w_h}\right) - \frac{1}{w_\ell} \varphi'\left(\frac{I_\ell}{w_\ell}\right)\right]. \tag{32}$$

Equations (27)–(32) show that taxation of individuals with no initial wealth subscribes to the customary properties of second-best income taxation. Specifically, equations (28)–(29) yield the "no distortion at the top" result (applying to both d_h and L_h). Moreover, with $(w_h, \pi_h, \beta_h) > (w_\ell, \pi_\ell, \beta_\ell)$, the bracketed expressions in the right-hand sides of (31) and (32) are both positive. Consequently, the left-hand sides of (31) and (32) exceed one: $\beta_\ell \phi'\left(d_\ell\right) > 1$ and $w_\ell > \varphi'\left(L_\ell\right)$. In words, consumption of d_ℓ and supply of L_ℓ are distorted downward. Finally, comparing $\beta_h \phi'\left(d_h\right) = 1$ with $\beta_\ell \phi'\left(d_\ell\right) > 1$ and $w_h = \varphi'\left(L_h\right)$ with $w_\ell > \varphi'\left(L_\ell\right)$ in conjunction with strict concavity of $\phi\left(\cdot\right)$ and strict convexity of $\varphi\left(\cdot\right)$, tells us that $d_h > d_\ell$ and $L_h > L_\ell$.

The interesting question from our perspective is to find out in what way the second-

best allocation of an i-type differs on the basis of his tag (belonging to the positive- or zero-inheritance group). To address this issue, compare the first-order conditions (24)–(26) pertaining to the positive-inheritance group with (27)–(32) pertaining to the zero-inheritance group. Consider first, the h-type who faces no distortion regardless of his inheritance status. Comparing (25)–(26) with (28)–(29) informs us that

$$\widehat{d}_h = d_h, \tag{33}$$

$$\widehat{L}_h = L_h. \tag{34}$$

Observe that these equalities arise not only because of the no-distortion at the top property but also the quasi-linearity of preferences. This latter property directs any potential differences in allocations due to income effects towards individuals' first-period consumption levels. Comparison of (24) with (27) reveals the impact of income effects. We have

$$\widehat{U}_h < U_h. \tag{35}$$

Rather counter-intuitively, the high-ability type who is the beneficiary of an accidental bequest ends up with less utility than his counterpart who receives no inheritance. Now, with second-period consumption and leisure being the same for an h-type with and without an inheritance, it also follows that

$$\widehat{c}_h < c_h. \tag{36}$$

The lower level of utility enjoyed by the h-type who receives an inheritance, manifests itself in a lower amount of first-period consumption.

Next consider the ℓ-types with and without an inheritance. The difference in their allocations arise from both income and incentive effects. Comparison of (25)–(26) with (31)–(32) reveals that

$$\widehat{d}_\ell > d_\ell, \tag{37}$$

$$\widehat{L}_\ell > L_\ell. \tag{38}$$

An ℓ-type who receives an inheritance consumes more in the second period and works more in the first as compared to an ℓ-type who receives no inheritance. This is due to his facing no distortions when he receives an inheritance but facing them when he has no inheritance. Turning to utility levels, comparing (24) with (30) informs us that

$$\widehat{U}_\ell > U_\ell. \tag{39}$$

The ℓ-type enjoys a higher level of utility if he receives an inheritance. However, these inequalities do not allow us to compare $\widehat{c}_\ell$ and c_ℓ.

To complete the discussion, we now ascertain the correctness of our initial assumption that observability of bequests identifies recipients' types. One can do this despite the fact that accidental bequests take *three* distinct values but we have only *two* types of recipients. The identification follows from our finding that $\widehat{d}_h = d_h > \widehat{d}_\ell > d_\ell$. Given these properties, leaving behind $\widehat{d}_h = d_h$ indicates that the deceased must have been of type h while leaving either $\widehat{d}_\ell$ or d_ℓ indicates type ℓ. The assumption of a dynastically immutable family type then establishes the recipient's type.

4.2 Tax policy

The tax policy is set to implement the second-best allocations characterized by (15)–(17) for the h- and ℓ-types in the positive-inheritance group and (18)–(23) for the h- and ℓ-types in the zero-inheritance group. To achieve this, the tax policy specifies an implementing tax schedule $T(I, s, \omega)$ as a function of the observable variables.[15] Thus the tax paid by an i-type will be $\widehat{T}_i = T(\widehat{I}_i, \widehat{s}_i, \omega_i)$ if he receives an inheritance and $T_i = T(I_i, s_i, 0)$ if he does not.

In what follows, we concentrate on the properties of $T(I, s, \omega)$ that pertain to the taxation of bequests. The properties of $T(I, s, \omega)$ with respect to income follow those of the Mirrlees optimal income tax problem. The properties of $T(I, s, \omega)$ with respect to

---

[15]One could also include c which can always be inferred from observability of the other variables. However, this argument would be redundant.

saving is derived in the Appendix. Suffice it to say here that implementation requires, as in the first best, saving subsidies. Now recall that in the first best, savings are subsidized at a rate equal to $1 - \pi_h$ for the h-type and $1 - \pi_\ell$ for the ℓ-type whether or not they receive an inheritance. In the second best, only the h-type faces the same subsidy rate regardless of his inheritance status. Moreover, the subsidy rate is the same as in the first best. The treatment of the ℓ-type, on the other hand, depends on whether he receives an inheritance or not. If he does, he will face a marginal subsidy rate of $1 - \pi_\ell$ as in the first best. If he does not, he will face a *smaller* marginal subsidy rate.[16]

In discussing the tax treatment of bequests, begin with the high-ability individuals. We have, from equations (33)–(36), $\widehat{s}_h = s_h, \widehat{L}_h = L_h$, and $\widehat{c}_h < c_h$. These imply

$$\widehat{T}_h > T_h + \omega_h.$$

Consequently, the difference between the taxes an h-type pays if he receives an inheritance and if he does not, exceeds the inheritance he receives. In this sense, the high-ability individuals face a tax on accidental bequests that is higher than 100%. This should not be surprising if one remembers that receiving ω_h identifies the person as a high-ability individual. With their ability known, the high-ability types in the positive-inheritance group enjoy no "informational rent," the extra utility that accrues to them under asymmetric information. The symmetry of the social welfare function then implies that the h-types in this group end up with the same utility level as the ℓ-types. On the other hand, the ability of the h-types in the zero-inheritance group remains private information and they enjoy some informational rent.

Comparing how the ℓ-types in positive- and zero-inheritance groups fare is rather more complicated. What is clear from the results characterized by (37)–(39) is that $\widehat{T}_\ell$

---

[16]Our formulation of second-best assumes public observability of savings at the individual level and thus admits nonlinear taxation. An alternative formulation is to assume that only anonymous savings are observable so that savings must be taxed linearly. This latter formulation can be done along Cremer and Gahvari (1997). The nature of information on savings, and the ensuing properties of marginal savings subsidies, does not change our results with respect to the taxation of bequests which is the focus of our paper. We have thus ignored this issue.

and $T_\ell + \omega_\ell$ are generally different. This implies that the ℓ-types too should not face a 100% tax on their inheritance. However, at this level of generality, it is not possible to determine whether the tax is smaller or larger than 100%. Specifically, the fact that $\widehat{U}_\ell > U_\ell$ is not sufficient to conclude that the ℓ-type in the positive-inheritance group pays a smaller tax than the ℓ-type in the zero-inheritance group. This follows because the ℓ-type in the zero-inheritance group faces a distortion and his lower level of utility may very well be due to this distortion. We will show below, through a numerical example, that the tax rate can be larger as well as smaller than 100%.

4.3   An example

Consider a quadratic disutility for labor $\varphi(L) = L^2/2$ and a logarithmic utility for future consumption $\phi(d) = \ln d$ so that our quasi-linear preferences (1) takes the form $U_i = c_i + \pi_i \beta_i \ln d_i - L_i^2/2$. Assume further that the concave transformation for U is iso-elastic, as previously defined, and given by $v(U) = U^{1-\varepsilon}/(1-\varepsilon)$ with $\varepsilon > 0$. The various parameter of the model are set as follows: $w_h = 2, w_\ell = 1, n_h = n_\ell = 1, \pi_h = \pi_\ell = 1/2, \beta_h = 1, \beta_\ell = 1/2$. Table 1 reports the numerical solutions the second-best allocations for $\varepsilon = 1/2$ and $\varepsilon = 11$. It also reports the values of marginal savings subsidies and net taxes collected in each case.[17]

The interesting property from our perspective is the behavior of $\widehat{T}_\ell - T_\ell$ in relation to ω_ℓ as ε increases. With $\varepsilon = 1/2$, we have $\widehat{T}_\ell = -0.458, T_\ell = -0.872$, and $\omega_\ell = 0.50$. Consequently, $\widehat{T}_\ell < T_\ell + \omega_\ell = -0.372$ and the tax ℓ-types pay on accidental bequests is less than 100%. On the other hand, $\varepsilon = 11$ results in $\widehat{T}_\ell = -0.456, T_\ell = -0.974$, and $\omega_\ell = 0.50$. In this case, $\widehat{T}_\ell > T_\ell + \omega_\ell = -0.474$ and the tax ℓ-types pay on

---

[17]These are given by

$$\begin{aligned} T_i &= I_i - c_i - d_i, \\ \widehat{T}_i &= \widehat{I}_i + \omega_i - \widehat{c}_i - \widehat{d}_i = \widehat{I}_i - \widehat{c}_i. \end{aligned}$$

Table 1. Second-best allocations and bequest taxes

	$\varepsilon = 0.50$		$\varepsilon = 11$	
	positive inheritance	zero inheritance	positive inheritance	zero inheritance
$\widehat{c}_h, c_h$	2.785	2.872	2.783	2.801
$\widehat{d}_h, d_h$	1.000	1.000	1.000	1.000
$\widehat{I}_h, I_h$	4.000	4.000	4.000	4.000
$\widehat{c}_\ell, c_\ell$	1.458	1.361	1.456	1.370
$\widehat{d}_\ell, d_\ell$	0.500	0.474	0.500	0.456
$\widehat{I}_\ell, I_\ell$	1.000	0.963	1.000	0.852
$\widehat{T}_h, T_h$	1.215	0.128	1.217	0.199
$\widehat{T}_\ell, T_\ell$	-0.458	-0.872	-0.456	-0.974

accidental bequests is more than 100%.[18] This finding is interesting in two ways. First, it shows that the second-best tax on accidental bequests may fall short of as well as exceed a confiscatory rate. Second, it is indicative of the impact of the desired degree of redistribution on the size of the tax. Whereas a "small" degree of inequality aversion— a small ε—calls for a bequest tax of less than 100%, a "large" degree of inequality aversion—a large ε—results in a bequest tax that exceeds 100%.

4.4 Rawlsian social welfare

The above example alerts us to the possibility that a higher degree of desired redistribution leads to higher tax rate on accidental bequests (by the ℓ-types). One can

---

[18]The definition of the bequest tax becomes rather ambiguous here. One can think of the bequest tax to be $T(I, s, \omega) - T(I, s, 0)$, i.e., the extra tax paid by an individual with bequest ω, for a given value for every other argument of the tax function. But for a nonlinear tax function, this difference is not (necessarily) constant and generally varies with the values I and s take. There is also an added problem in our two-type setting. Here, one has to calculate $T(I_i, s_i, \omega_i) - T(I_i, s_i, 0)$ as well as $T\left(\widehat{I}_i, \widehat{s}_i, \omega_i\right) - T\left(\widehat{I}_i, \widehat{s}_i, 0\right)$ which may not only yield different answers but more problematically are not even unambiguously defined by the optimal allocation. The problem is that $T(I_i, s_i, \omega_i)$ in the first expression and $T\left(\widehat{I}_i, \widehat{s}_i, 0\right)$ in the second are hypothetical. As usual, in two-type models, there are many degrees of freedom in constructing the implementing tax function and this affects $T(I, s, \omega) - T(I, s, 0)$. We circumvent all these difficulties by defining the bequest tax to be the difference between *net* tax payments for an i-type if he receives an inheritance and if he does not: $T\left(\widehat{I}_i, \widehat{s}_i, \omega_i\right) - T(I_i, s_i, 0)$.

indeed derive a formal result that support this view if the social welfare function is Rawlsian. In our setup, the Rawlsian solution is obtained as the limit for a suitably-defined transformation function $v(\cdot)$. Specifically, it is obtained by letting $\varepsilon \to \infty$ in $v(U) = U^{1-\varepsilon}/(1-\varepsilon)$.

With a Rawlsian social welfare function, the utility level of the low-ability individuals should be equalized across the two tagged groups of positive- and zero-inheritance: $\widehat{U}_\ell = U_\ell$. Recall that the ℓ-type faces a distortion in the zero-inheritance group but not in the positive-inheritance group. Attaining the same utility level despite the distortion in the zero-inheritance group, requires the ℓ-type to have more resources at his disposal. Consequently, $T_\ell < \widehat{T}_\ell - \omega_\ell$ or $\widehat{T}_\ell > T_\ell + \omega_\ell$. In words, the ℓ-type in the positive-inheritance group faces a bequest tax that exceeds 100%.

The main results of this section are summarized in the following proposition.

Proposition 2 *Consider the society described in Proposition 1 but assume that individual types are not publicly observable while income, consumption levels, and bequests are.*

(i) The second best solution has the following properties: (a) Individuals can be partitioned into two groups (tags): Those who receive an accidental bequest and those who do not. The characteristics of people in the first group can be inferred from their bequests and they will be given a first-best tax schedule. The characteristics of people in the second group remain private information and they will face a standard Mirrlees optimal tax problem. (b) The high-ability individuals who receive an inheritance lose all their informational rent and $\widehat{U}_h = \widehat{U}_\ell$. (c) Individuals in the zero-inheritance group face no distortion at the top and a downward distortion on labor supply and savings for the ℓ-types. The h-types in this group enjoy some informational rent so that $U_h > U_\ell$.

(ii) Second-best allocation of the h-types in the positive- and zero-inheritance groups differ according to $\widehat{c}_h < c_h, \widehat{d}_h = d_h, \widehat{L}_h = L_h$, and we have $\widehat{U}_h < U_h$.

(iii) Second-best allocation of the ℓ-types in the positive- and zero-inheritance groups

differ according to $\widehat{d}_\ell > d_\ell, \widehat{L}_\ell > L_\ell$, and we have $\widehat{U}_\ell > U_\ell$.

(iv) Decentralization of second-best allocations requires: (a) Marginal saving subsidies; (b) High-ability individuals face a bequest tax of more then 100%: $\widehat{T}_h > T_h + \omega_h$; (c) Low-ability individuals face a bequest tax that can be smaller as well as larger than 100%; i.e., one can have $\widehat{T}_\ell < T_\ell + \omega_\ell$ or $\widehat{T}_\ell > T_\ell + \omega_\ell$.

(v) With a Rawlsian social welfare function, everyone who receives an inheritance pays an inheritance tax of more than 100%.

5 Summary and conclusion

This paper has questioned the validity of the conventional wisdom that purely accidental bequests should be taxed at a confiscatory rate. It has employed a model wherein individuals of different abilities may live for one or for two periods with different probabilities of survival. It has shown that the proposition is correct in a first-best environment when individuals' productivity and longevity are publicly observable. Under this circumstance, subsidizing each ability-type's saving at a rate equal to his probability of an early death, in conjunction with lump-sum taxes that vary according to individuals' ability types and inheritance status, mimics a perfect annuity market. All accidental bequests are taxed at 100% and all individuals enjoy the same level of utility.

In the second-best, individual abilities and survival probabilities are publicly unobservable. Assuming that types and survival probabilities are positively correlated, individuals can be partitioned into two groups (tags). The first group consists of people who receive an accidental bequest and the second of those who receive nothing. The characteristics of people in the first group can be inferred from their bequests and they will be given a first-best tax schedule. The characteristics of people in the second group remains private information and they will have to face a standard Mirrlees optimal tax problem.

With their ability type being inferred, the high-ability individuals in the group of

people who receive an inheritance enjoy no informational rent and will end up with the same utility level as the low-ability types in this group. On the other hand, high-ability types in the group of people who receive no inheritance enjoy an informational rent. This implies that high-ability types will be better off if they do not receive an inheritance. Similar comparison for low-ability types reveals that they will be better off receiving an inheritance. In this sense, accidental bequests are a curse for the rich and a boon for the poor. Finally, to decentralize these allocations, one must levy marginal saving subsidies that vary with income and inheritance status. High-ability individuals face a bequest tax of more then 100%, while low-ability individuals face a bequest tax that can be smaller as well as larger than 100%. With a Rawlsian social welfare function, everyone who receives an inheritance pays an inheritance tax of more than 100%.

We conclude by observing that this paper has admittedly focused on a highly stylized setting wherein the observability of bequests brings about a drastic change in the structure of information available to the tax authority. In following this approach, we have been led by a desire to convey our point in a crisp fashion with no ambiguity. However, our main point that incorporating the informational content of bequests improves the design of tax structures and the properties of bequest taxes would go through under less drastic assumptions. The design of the optimal tax policy would of course become more complicated.

Appendix A

Characterization of marginal saving subsidies: Faced with the tax function $T(I, s, \omega)$, the i-type in the positive-inheritance group chooses I and s to maximize

$$\widehat{U}_i = I + \omega_i - T(I, s, \omega_i) - d + \pi_i \beta_i \phi(s) - \varphi\left(\frac{I}{w_i}\right),$$

and the i-type in the zero-inheritance group chooses I and s to maximize

$$U_i = I - T(I, s, 0) - d + \pi_i \beta_i \phi(s) - \varphi\left(\frac{I}{w_i}\right),$$

where in both cases we have set $d = s$. Denote the partial derivative of $T(\cdot)$ with respect to s by $T_s(\cdot)$. The first-order condition with respect to s, whether one is in the positive or zero-inheritance group, is then equal to

$$-T_s(I, s, \omega) - 1 + \pi_i \beta_i \phi'(s) = 0.$$

Comparing this condition with its counterparts in the second-best [first-order condition (19) for everyone who receives an inheritance, and (28) and (31) for h- and ℓ-types who do not receive an inheritance] yields the following *marginal* saving subsidies:

$$-T_s\left(\widehat{I}_i, \widehat{s}_i, \omega_i\right) = 1 - \pi_i \beta_i \phi'(\widehat{s}_i) = 1 - \pi_i, \quad i = h, \ell, \tag{A1}$$

$$-T_s(I_h, s_h, 0) = 1 - \pi_h \beta_h \phi'(s_h) = 1 - \pi_h, \tag{A2}$$

$$-T_s(I_\ell, s_\ell, 0) = 1 - \pi_\ell \beta_\ell \phi'(s_\ell) = 1 - \pi_\ell \left[\frac{\lambda \phi'(s_\ell)}{\mu n_\ell \pi_\ell}\left(\beta_h \frac{\pi_h}{\pi_\ell} - \beta_\ell\right)\right]. \tag{A3}$$

References

[1] Akerlof, George A. (1978), The economics of "tagging" as applied to the Optimal income tax, welfare programs, and manpower planning, *American Economic Review,* 68, 8–19.

[2] Atkinson, Anthony B. (1970), On the measurement of inequality, *Journal of Economic Theory,* 2, 244–263.

[3] Atkinson, Anthony B. and Joseph E. Stiglitz (1976), The design of tax structure: direct versus indirect taxation, *Journal of Public Economics*, 6, 55–75.

[4] Blumkin, Tomer and Efraim Sadka (2004), Estate taxation with intended and accidental bequests, *Journal of Public Economics*, 88, 1-21.

[5] Boadway, Robin and Pierre Pestieau (2006), Tagging and redistributive taxation. Annales d'Economie et de Statistique 83/84, 123–150.

[6] Bommier, Antoine (2006), Uncertain lifetime and intertemporal choice: Risk aversion as a rationale for time discounting, *International Economic Review,* 47, 1223–1246.

[7] Cremer, Helmuth, Gahvari, Firouz and Jean-Marie Lozachmeur (2010), Tagging and income taxation: theory and applications, *American Economic Journal: Economic Policy*, forthcoming.

[8] Cremer, Helmuth and Pierre Pestieau (2005), Wealth transfer taxation: a survey of the theoretical literature, in *Handbook on the Economics of Giving, Reciprocity and Altruism*, L.A. Gerard-Varet, S-Ch. Kolm and J. Mercier-Ythier, North-Holland, Amsterdam.

[9] Cremer, Helmuth, Pestieau, Pierre and Jean-Charles Rochet (2003), Capital income taxation when inherited wealth is not observable, *Journal of Public Economics*, 87, 2475–2490.

[10] Kaplow, Louis (2008), *The Theory of Taxation and Public Economics*, Princeton University Press, 2008.

[11] Kopczuk, Wojciech (2003), The trick is to live: Is the Estate tax social security for the rich? *Journal of Political Economy*, 111, 1318–1341.

[12] Kopczuk, Wojciech (2003), Economics of estate taxation, unpublished manuscript, Columbia University.

[13] Michel, Philippe and Pierre Pestieau (2002) Wealth transfer taxation with both accidental and planned bequest, CORE DP 2002/59.

[14] Pestieau, Pierre and Motohiro Sato (2009), Estate taxation with both accidental and planned bequests, *Asia Pacific Journal of Accounting and Economics*, forthcoming.

[15] Preston, Samuel H (1975), The changing relation between mortality and the level of economic development, *Population Studies*, 29, 231–248.

[16] Pritchett, Lant and Lawrence Summers (1996), Wealthier is healthier, *Journal of Human Resources*, 31, 841–868.