

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

Helmuth Cremer

Toulouse School of Economics, FR-31000 Toulouse, France
helmuth.cremer@tse-fr.eu

Firouz Gahvari

University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA
fgahvari@illinois.edu

Pierre Pestieau

University of Liège B-4000 Liège, Belgium
p.pestieau@ulg.ac.be

Abstract

When accidental bequests signal otherwise unobservable individual characteristics, such as productivity and longevity, the population should be partitioned into two groups: those who do not receive an inheritance and those who do. The first tagged group receives a Mirrlees second-best tax schedule; the second group, when its type is fully revealed, faces a first-best tax schedule. Receiving an inheritance makes high-ability types worse off and low-ability types better off. High-ability individuals face a bequest tax of more than 100 percent, while low-ability types face a bequest tax that can be smaller, as well as larger, than 100 percent, and it might even be negative.

Keywords: Accidental bequests; estate tax; first-best; second-best; tagging

JEL classification: H21

I. Introduction

There is little agreement in the literature concerning the taxation of bequests in general.¹ Yet, when it comes to accidental bequests, a widely held view has formed that these should be taxed at 100 percent. The rationalization for this argument is that a confiscatory rate, combined with equal payments to offsprings or parents, is equivalent to providing full insurance for the former or perfect annuities for the latter.² The aim of our paper is to challenge this received view. Our contention is that publicly observable

*We thank Tomer Blumkin and Jean-Marie Lozachmeur for helpful remarks. We are particularly grateful to the referees for their detailed and constructive comments.

¹ See Cremer and Pestieau (2005) for a survey.

² Kaplow (2008, pp. 264–266) has argued that a 100 percent bequest tax, with its proceeds rebated equally to all children, mimics this insurance scheme and is non-distorting. Kopczuk

bequests have informational content that should be incorporated in the design of optimal tax structures. Also, if this is done, the 100 percent tax rate on bequests is no longer optimal.

The basic idea is for the tax administration to use bequests as a separation mechanism, or a “tag”, when designing an optimal tax system. In this way, it can partition the population into two groups: one consists of people who do not receive an inheritance and the other consists of those who do. The information that inheritances convey about the ability and longevity of the individuals allows the government to offer two different Mirrlees tax schedules to the two groups. A consequence of the welfare-enhancing differential tax treatment of the two groups is to undermine the desirability of the 100 percent tax rate on accidental bequests.

To make our point as stark and as simplified as possible, we use a stylized model in which receiving bequests fully reveals one’s type. Under this circumstance, individuals with no inheritances face a standard Mirrlees optimal non-linear tax schedule. However, those who receive an inheritance face a non-distortionary tax scheme. Two interesting and interrelated results emerge. The first is that inheritances make the high-ability types worse off and the low-ability types better off. Second, the high-ability types should face a bequest tax that necessarily exceeds 100 percent, but the low-ability types face a tax rate that can be smaller or larger than 100 percent, and it might even be negative.³

To the best of our knowledge, the paper by Blumkin and Sadka (2004) is the only paper that has questioned the wisdom of the 100 percent tax idea. In their set-up, the tax system is linear. 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 it increases the efficiency of the tax system. Our challenge to the 100 percent tax idea is somewhat more basic, and it relies on the understanding that bequests

(2003a) has pointed out that the market for annuities too, under certain conditions, can be mimicked by a government policy that includes 100 percent bequest taxes. In this scenario, 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.

³ Ignoring the informational content of bequests in this set-up calls for a 100 percent tax rate on bequests (within the context of the tax system as a whole, and when tax instruments are not artificially restricted). The tax neither distorts the behavior of parents nor their utility because we rule out all bequest motives. The behavior of recipients is not distorted either because accidental bequests are a windfall for them. Moreover, as 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 must be independent of who initially owns the non-distortionary revenue sources.

have informational content that should be incorporated into the design of optimal tax structures.

The fact that we ignore the bequest motives of parents does not mean that planned bequests do not matter. Nor do we claim that accidental bequests are the most important type of bequests. Their choice for this study is motivated by the fact that they are the only type of bequests for which there is a consensus as to how they should be taxed. Similarly, our approach is motivated by the fact that circumventing the identification of bequest motives simplifies the analysis drastically and highlights our point most succinctly. Nevertheless, our central point that the informational content of bequests should be incorporated into the design of optimal tax structures is a general one. It applies (i) to all types of bequests, regardless of bequest motives, and (ii) to settings where receiving bequests do not fully reveal one's type. Our result on the undesirability of a 100 percent tax on accidental bequests is also a general result that remains valid under (i) and (ii). The result that might not go through in a more general setting is that of high-ability individuals facing a bequest tax of more than 100 percent.⁴

Finally, in emphasizing the informational content of bequests, we balance the availability of information with practical considerations. Thus, we do not allow the tax liability of an individual to be based directly on the income of their parents. From the perspective of a purely theoretical (mechanism) design, this restriction is arbitrary. However, the assumption is a sensible one from a practical perspective. Tax schedules are often restricted to depend on contemporaneous variables only. A person's income tax liability in a given period, for example, does not even depend on their full earning history. It is these practical considerations that motivate our conditioning an individual's tax liability on their own transactions, income levels, or wealth, alone. Incorporating the available information on the previous generation does not significantly change our formal analysis. However, the interpretation of the results would change, because the information that can be obtained from bequests could also be obtained from these other variables. In other words, we would have a redundancy in the informational structure.

⁴ Even here, the result appears, at first glance, to generalize. Intuitively, one would expect that high-ability individuals enjoy less informational rents, and thus pay more taxes, as members of the group for which better information is available. In our problem, this is the group of individuals who receive an inheritance. This means that our result should hold even in this case. Yet, the literature on tax design with tagging has shown that this type of conjecture has to be interpreted with great care. Additional assumptions are needed to ensure this result.

II. The Setting

Basic Model

Consider a two-period overlapping-generation model wherein individuals of each generation live for either one or two periods. Regardless of their longevity, they work in the first period only. Those whose parents die early (i.e., they 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 their resources – earnings plus any inheritances – between present consumption and savings 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 savings are transmitted to their children as an accidental bequest. Savings are channeled into future consumption through a storage technology; there is no appreciation or depreciation of savings, so that the interest rate is zero.⁵

Individuals differ in their productivity w_i , their survival probability π_i , and taste for future versus present consumption, represented by a weight β_i to future consumption in the individuals' utility functions. We assume, based on empirical evidence, that these characteristics are positively correlated. Preston (1975), Duleep (1986), Pritchett and Summers (1996), Deaton and Paxson (1998), Waldron (2007), and Salm (2007), among others, have all pointed to a positive relationship between income/education and life expectancy.⁶ Saez (2002) has argued that individuals with higher earnings save relatively more, which suggests that high-ability individuals are likely to have a higher taste for savings. Banks and Diamond (2010) have nurtured similar ideas. Finally, Bommier (2006) has provided empirical support for a positive correlation between longevity and preferences for savings.

To simplify, we also assume that each characteristic takes only two values: high, indexed by h , and low, indexed by ℓ . The two-value simplification implies that the assumed positive correlation between the specified

⁵ The important assumption here is the constancy of the interest rate rather than having it to respond to changes in savings and capital accumulation. This is customarily done either by assuming a storage technology or by postulating a small open economy whose savings do not change the interest rate it faces. Observe also that, as long as the interest rate is constant, setting it at a positive rate rather than zero does not change the analysis.

⁶ In their study of US data, Singh and Siahpush (2006) have shown that not only has life expectancy been steadily increasing over the past several decades, but that, accompanying these increases, there has been a growing disparity in life expectancy between individuals with high and low incomes and between those with more and less education. Differences in life expectancy across socioeconomic groups are significantly larger now than in 1980 or 1990. A similar trend has been observed in a number of other countries.

characteristics is perfect. This means that, effectively, we have only two types of agents: h and ℓ with $w_h > w_\ell$, $\pi_h > \pi_\ell$, and $\beta_h > \beta_\ell$.⁷ Moreover, types are assumed to be dynastically immutable: if a person is of type i , their offsprings will also be of type i . These are, of course, oversimplifications, but they have the advantage of greatly simplifying the multidimensional screening problem that arises when individuals differ in more than one dimension.⁸ Nevertheless, their adoption does not affect our central message (although they will affect our specific results below). The fact that informative bequests should serve as a separating mechanism in the design of optimal tax structures, and the fact that doing so dispels the received view of the 100 percent taxation of accidental bequests, are invariant to this assumption.⁹

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 .¹⁰ An individual's expected utility is given by

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

where ϕ is strictly concave while φ is strictly convex.¹¹

Individuals of the first generation start life with no initial wealth. However, this will not be the case for members of the generations that follow. Those who die early leave 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 individuals' (accidental) bequests to their

⁷ It is sufficient to assume $\pi_h \geq \pi_\ell$ and $\beta_h \geq \beta_\ell$ with one of the two inequalities being strict.

⁸ Without perfect correlation, it is often impossible to determine which incentive constraints are binding at the second-best solution. For a discussion of multidimensional heterogeneity in the context of capital and wealth taxation, see Cremer *et al.* (2003) and, more recently, Diamond and Spinnewijn (2011).

⁹ In Cremer *et al.* (2010), for example, the correlation between earnings and gender used as tag is far from perfect. Yet, they have shown that tagging based on gender (i.e., offering males and females different tax schedules) is welfare-improving. However, in the absence of these assumptions, the underlying tagging problem becomes significantly more complicated. See Section V for a further discussion of this issue.

¹⁰ This assumption simplifies the analysis and exposition of our results. It is not crucial for the main results of this paper concerning the properties of first- and second-best allocations (and the associated tax policies). These results, derived below, carry over with only minor modifications to a setting with general preferences. See the last paragraph in Section VI.

¹¹ This formulation assumes that individuals derive no utility from leaving accidental bequests to their families if they die early. Moreover, the fact that the same amount of consumption yields different utility levels in the two periods, c and $\phi(d)$, indicates differing intertemporal preferences. Finally, the concavity of $\phi(d)$ reflects risk aversion.

children is unaffected by the size of the inheritance that they might have received from their own parents (including zero). Then, we will 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.¹²

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

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

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 $\pi_i \beta_i \phi'(s_i) = 1$, or $s_i = \phi'^{-1}(1/\pi_i \beta_i)$, and $w_i = \varphi'(L_i)$. These relationships, along with the assumptions $\pi_h > \pi_\ell$, $\beta_h > \beta_\ell$, the strict concavity of ϕ , and the strict convexity of φ , imply that $s_h > s_\ell$ and $L_h > L_\ell$. Finally, observe that because, in the absence of annuity markets, every i -type person saves s_i to finance their future consumption, those who die early must leave an accidental bequest of s_i behind (i.e., $\omega_i = s_i$).

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 continues to be summarized by equation (2), resulting in the same solution as that 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. Then, it 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 (i.e., $\omega_i = s_i$).

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

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

where the symbol $\widehat{}$ 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

¹² The dependence of the bequest one leaves on the inheritance one receives leads to a multiplicity of groups on the basis of ω_i , making the existence of a steady state problematic. Observe also that, even with quasi-linear preferences, ω_i might 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.

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 is 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 the second generation, 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. Moreover, with $s_h > s_\ell$, it is also the case that $\omega_h > \omega_\ell$.

III. First-Best

Allocation

Assume that 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 $(\widehat{c}_i, \widehat{d}_i, \widehat{L}_i)$ to maximize social welfare, defined by

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

Here, v is a strictly concave transformation of the quasi-linear utility function (1), and hence $v' > 0$ and $v'' < 0$. Aside from this transformation, the social welfare function defined by equation (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 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 is, the more redistributive the social welfare function will be. One common specification for v is the iso-elastic case $v(U) = U^{1-\varepsilon} / (1 - \varepsilon)$, $\varepsilon \neq 1$, suggested by Atkinson (1970). In this formulation, $\varepsilon > 0$ denotes the inequality aversion index, and the higher ε is, the greater the desired redistribution will be.

The resource constraint for the economy is given by

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

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.¹³ The first-best optimum is then characterized by choosing (c_i, d_i, L_i) and $(\widehat{c}_i, \widehat{d}_i, \widehat{L}_i)$ in order to maximize equation (4) subject to equation (5).¹⁴ Let μ denote the Lagrangian multiplier associated with the resource constraint (5). The Lagrangian expression associated with this problem is

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

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

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

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

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

Denoting the first-best values by superscript “FB”, it follows from equations (8) and (9) that

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

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

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

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

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

Thus, at first-best, the difference in type affects one’s allocation but not the difference in inheritance status. In other words, the allocations of h - and ℓ -types differ, but either one receives the same allocation, regardless of receiving an inheritance or not.

¹³ While the inheritances that one generation receives might, in principle, be different from the bequests it leaves, such an outcome cannot occur in the steady state.

¹⁴ There is also the constraint that $s_i = w_i L_i - c_i \geq 0$, which we assume to be non-binding.

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 to be conditioned on types, but not on inherited wealth, and they need to be set at a rate equal to

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

for type $i = h, \ell$.¹⁵ However, lump-sum taxes, which can be negative as well as positive, must be 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. Naturally, these fiscal instruments must also satisfy the government’s budget constraint

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

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

$$c_i + (1 - \tau_i)s_i = w_i L_i + \omega_i - \widehat{t}_i$$

and

$$c_i + (1 - \tau_i)s_i = w_i L_i - t_i.$$

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 $\widehat{L}_i = L_i = L_i^{\text{FB}}$ and $\widehat{s}_i = s_i = s_i^{\text{FB}}$. Additionally, to ensure $\widehat{c}_i = c_i = c_i^{\text{FB}}$, one must set lump-sum taxes at the rates $t_i = w_i L_i^{\text{FB}} - c_i^{\text{FB}} - \pi_i s_i^{\text{FB}}$ and $\widehat{t}_i = w_i L_i^{\text{FB}} + \omega_i - c_i^{\text{FB}} - \pi_i s_i^{\text{FB}}$.

Finally, observe that the expressions for \widehat{t}_i and t_i show that $\widehat{t}_i = t_i + \omega_i$. 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, we can consider the bequest tax to be the difference between individuals’ total tax liabilities when they receive an inheritance and when they do not. There is no tension between these definitions, however. Given identical savings

¹⁵ The subsidy is required because there is no annuity market. The absence of such a market implies that the private cost of future consumption d , as perceived in the first period when individuals do not know if they will be alive in the second period, exceeds its social cost. A Pigouvian subsidy corrects for this divergence.

under the two scenarios ($\widehat{s}_i = s_i$), both definitions lead to the same answer.¹⁶ The implementation of first-best allocations requires a 100 percent taxation of (accidental) bequests. An i -type with an inheritance ω_i , $i = h, \ell$, sees their entire inheritance confiscated, after which they are treated like their 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 equation (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) the equalization of utilities of all individuals regardless of their type and inheritance status (i.e., $\widehat{U}_h = U_h = \widehat{U}_\ell = U_\ell$) and (b) a higher accidental bequest for individuals of type h (i.e., $\omega_h > \omega_\ell$).*
- (ii) *Decentralization of first-best allocations requires (a) saving subsidies conditional on individual survival probabilities, 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 (specifically wages) as well as inherited wealth: $(\widehat{t}_h, \widehat{t}_\ell)$ for those with an inheritance and (t_h, t_ℓ) for those without.*
- (iii) *All accidental bequests are taxed at 100 percent.*

IV. Second-Best

Allocation

We 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 Mirrlees optimal non-linear income tax models. The only difference is that we have added bequests as an observable variable and thus they are potentially taxable.¹⁷

¹⁶ We denote the net taxes an i -type pays, $i = h, \ell$, by T_i if they do not receive an inheritance and by \widehat{T}_i if they do. Then, it must be the case that $T_i = t_i - \tau_i s_i$ and $\widehat{T}_i = \widehat{t}_i - \tau_i \widehat{s}_i$. With $t_i = \widehat{t}_i - \omega_i$ and $s_i = \widehat{s}_i$, it also follows that $T_i = \widehat{T}_i - \omega_i$.

¹⁷ Recall that, throughout the paper, we assume that an individual's tax liability depends only on variables pertaining to the individual: income, consumption, and inheritances received.

The key point that we make is that the observability of a bequest is sufficient to identify its recipient's type. However, 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.¹⁸

We start by assuming that the government is able to identify the type of a person who receives a bequest, and then we 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, respectively). The zero-inheritance group, consisting of people whose ability remains private information, will have to face a tax schedule determined on the basis of the Mirrlees standard optimal non-linear income tax problem. However, the positive-inheritance group 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. 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, we can formulate the problem within each group independently, connecting the two via the economy's resource constraint (e.g., Cremer *et al.*, 2010). In other words, we assign a single resource constraint to the two groups. This continues to be represented by equation (5). Let μ denote the Lagrange multiplier associated with the economy's resource constraint, and let λ be the multiplier associated with the incentive constraint in the group of individuals who receive no inheritance.¹⁹ The Lagrangian expression associated with this optimization problem is

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

Sociopolitical considerations prevent the government from conditioning a person's tax liability on their parents' characteristics.

¹⁸ The paper by 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 body of literature.

¹⁹ In this group, the government does not observe an individual's type. Consequently, the designed tax policy must be incentive-compatible (i.e., the policy must induce self-selection).

In writing equation (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, in order 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 III.²⁰

Observe that bequests do not appear directly in the economy’s resource constraint. 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 equation (14), being the defining characteristic of zero- and positive-inheritance groups. It is because of the information that bequests convey that equation (14) does not contain an incentive constraint for people in the positive-inheritance group.

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

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

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

and

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

For those without an inheritance, they are

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

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

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

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

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

²⁰ However, we can equally express the problem in terms of \widehat{I}_i .

and

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

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

$$v'(\widehat{U}_i) = \mu, \quad \beta_i \phi'(\widehat{d}_i) = 1, \quad \text{and} \quad \phi'(\widehat{L}_i) = w_i. \quad (24)$$

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 the positive-inheritance group because their characteristics are observable. Equations (24) 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 a high-ability type consumes more in the second period, and works more in the first, compared to a low-ability type.

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

$$v'(U_h) = \mu - \frac{\lambda}{n_h \pi_h}, \quad (25)$$

$$\beta_h \phi'(d_h) = 1, \quad (26)$$

$$\frac{1}{w_h} \phi' \left(\frac{I_h}{w_h} \right) = 1, \quad (27)$$

$$v'(U_\ell) = \mu + \frac{\lambda}{n_\ell \pi_\ell}, \quad (28)$$

$$\beta_\ell \phi'(d_\ell) = 1 + \frac{\lambda \phi'(d_\ell)}{\mu n_\ell \pi_\ell} \left(\beta_h \frac{\pi_h}{\pi_\ell} - \beta_\ell \right), \quad (29)$$

$$\frac{1}{w_\ell} \phi' \left(\frac{I_\ell}{w_\ell} \right) = 1 + \frac{\lambda}{\mu n_\ell \pi_\ell} \left[\frac{1}{w_h} \phi' \left(\frac{I_\ell}{w_h} \right) - \frac{1}{w_\ell} \phi' \left(\frac{I_\ell}{w_\ell} \right) \right]. \quad (30)$$

Equations (25)–(30) show that taxation of individuals with no initial wealth subscribes to the customary properties of second-best income taxation. Specifically, equations (26)–(27) yield the “no distortion at the top” result (applying to both d_h and L_h). In other words, for any pair of goods (including leisure), the trade-off that the “top” (i.e., the high-ability individuals) face in the second-best is the same as in the first-best. This implies, in particular, that in the second-best also, the marginal tax on labor income is zero for the high-ability individuals. Turning to equations (29)–(30), with $(w_h, \pi_h, \beta_h) > (w_\ell, \pi_\ell, \beta_\ell)$, the bracketed expressions on

the right-hand sides are both positive. Consequently, the left-hand sides of equations (29) and (30) exceed one: $\beta_\ell \phi'(d_\ell) > 1$ and $w_\ell > \varphi'(L_\ell)$. In other words, the consumption of d_ℓ and supply of L_ℓ are distorted downward. Finally, comparing $\beta_h \phi'(d_h) = 1$ with $\beta_\ell \phi'(d_\ell) > 1$ and $w_h = \varphi'(L_h)$ with $w_\ell > \varphi'(L_\ell)$, in conjunction with the strict concavity of ϕ and the strict convexity of φ , tells us that $d_h > d_\ell$ and $L_h > L_\ell$.

An interesting question from our perspective is to find out in what way the second-best allocation of i -types differs on the basis of their tags (belonging to the positive- or zero-inheritance group). To address this issue, compare the first-order conditions (24) pertaining to the positive-inheritance group with equations (25)–(30) pertaining to the zero-inheritance group. Consider first the h -type who faces no distortion regardless of inheritance status. Comparing equation (24) with equations (26)–(27) informs us that

$$\widehat{d}_h = d_h, \text{ and } \widehat{L}_h = L_h. \quad (31)$$

Observe that these equalities arise not only because of the “no distortion at the top” property but also because of the quasi-linearity of preferences. This latter property directs any potential differences in allocations because of income effects towards the first-period consumption levels of individuals. A comparison of equation (24) with equation (25) reveals the impact of income effects. We have $\widehat{U}_h < U_h$. Rather counter-intuitively, the high-ability type who is the beneficiary of an accidental bequest ends up with less utility than their 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

$$\widehat{c}_h < c_h. \quad (32)$$

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 arises from both income and incentive effects. A comparison of equation (24) with equations (29)–(30) reveals that $\widehat{d}_\ell > d_\ell$ and $\widehat{L}_\ell > L_\ell$. An ℓ -type who receives an inheritance consumes more in the second period and works more in the first, compared to an ℓ -type who receives no inheritance. This is because they face no distortions when they receive an inheritance but they do face distortions when they have no inheritance. Turning to utility levels, comparing equation (24) with equation (28) informs us that

$$\widehat{U}_\ell > U_\ell. \quad (33)$$

The ℓ -type enjoys a higher level of utility if they receive an inheritance. However, these inequalities do not allow us to compare \widehat{c}_ℓ and c_ℓ .

To complete the discussion, we now ascertain the correctness of our initial assumption that observability of bequests identifies the recipients' types. We can do this despite the fact that accidental bequests take three distinct values, whereas 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}_ℓ or d_ℓ indicates type ℓ . The assumption of a dynastically immutable family type then establishes the recipient's type.

Tax Policy

Tax policy is set in order to implement the second-best allocations characterized by equations (15)–(17) for the h - and ℓ -types in the positive-inheritance group and by equations (18)–(23) for the h - and ℓ -types in the zero-inheritance group. To achieve this, the policy specifies an implementing tax schedule, $T(I, s, \omega)$, as a function of the observable variables: income, savings, and inheritance.²¹ In what follows, we state the properties of this function with respect to income and savings briefly, and then we concentrate on the properties 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, and thus they are well-known. The properties of $T(I, s, \omega)$ with respect to saving are derived in the Appendix. Suffice it to say here that implementation requires saving subsidies, as it did in the first-best. 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. However, the treatment of the ℓ -type depends on whether they receive an inheritance or not. If they do, they will face a marginal subsidy rate of $1 - \pi_\ell$, as in the first-best. If they do not, they will face a smaller marginal subsidy rate.²² However, this subsidy rate is independent of the inheritance level, ω_ℓ or $\widehat{\omega}_\ell$.

²¹ As pointed out by a referee, the tax is formally conditioned on inheritances and not on types (even when inheritances fully reveal the types). Note also that the tax can also be conditioned on c (a variable that is inferred from observability of the other variables). However, this would be a redundant argument for the tax function.

²² The reason for the lower marginal subsidy rate comes from the original result of Cremer and Gahvari (1995), that uncertain earnings provide a rationale for taxation of savings (see Banks and Diamond, 2010, p. 564, for a discussion). In the present context, this arises because, with high-skilled individuals valuing second-period consumption more than the low-skilled, taxation of savings relative to the first-best – lowering its first-best subsidy – slackens the otherwise binding self-selection constraint.

Turning to bequest taxation, consider first the treatment of high-ability individuals. They inherit either nothing or $\omega_h = d_h = \widehat{d}_h$. They pay a tax equal to $T_h = T(I_h, s_h, 0) = I_h - c_h - d_h$ if they receive no inheritance and $\widehat{T}_h = T(\widehat{I}_h, \widehat{s}_h, \omega_h) = \widehat{I}_h + \omega_h - \widehat{c}_h - \widehat{d}_h = \widehat{I}_h - \widehat{c}_h$ if they do. Now, from equations (31)–(32), $\widehat{d}_h = d_h$, $\widehat{I}_h = I_h$, and $\widehat{c}_h < c_h$. Hence, $\widehat{T}_h - T_h = d_h + (c_h - \widehat{c}_h) = \omega_h + (c_h - \widehat{c}_h)$, and

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

Consequently, the difference between an h -type's total tax liabilities if they receive an inheritance and if they do not, exceeds the inheritance they might receive. In this sense, the high-ability individuals face a tax on accidental bequests that is higher than 100 percent. This should not be surprising if we remember that a high-ability person's type is revealed when they inherit ω_h , but that their type remains unidentified otherwise. With their ability known, an h -type who receives an inheritance enjoys no "informational rent" over an l -type who also receives an inheritance. 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. However, because the ability of an h -type who inherits nothing remains private information, they enjoy some informational rent.

Comparing how the l -types in positive- and zero-inheritance groups fare is more complicated. In any given generation, some of the l -types have received no inheritance at all, some have received $\omega_\ell = d_\ell$, and some $\widehat{\omega}_\ell = \widehat{d}_\ell$. However, an l -type's allocation depends only on whether they have received an inheritance or not; those who have inherited ω_ℓ have an identical allocation to those who have inherited $\widehat{\omega}_\ell$. Specifically, an l -type who inherits nothing plans for a future consumption level of d_ℓ , while an l -type who has inherited either ω_ℓ or $\widehat{\omega}_\ell$ plans for a consumption level of \widehat{d}_ℓ .²³ This tells us that the l -types who inherit $\widehat{\omega}_\ell$ pay, effectively, a tax on their extra inheritance, $\widehat{\omega}_\ell - \omega_\ell$, at a rate of 100 percent. We can then consider the tax paid on $\widehat{\omega}_\ell$ as consisting of two parts. One part is paid on inheritances up to ω_ℓ ; this tax is identical to the inheritance tax paid by those who receive only ω_ℓ . This is followed by a tax on the remaining $\widehat{\omega}_\ell - \omega_\ell$ inheritance; this is levied at a confiscatory rate.

Beyond this, we cannot, at this level of generality, determine whether the tax on ω_ℓ is smaller or larger than 100 percent. Specifically, the finding $\widehat{U}_\ell > U_\ell$ does not allow us to conclude that the net disposable income of an l -type with an inheritance is larger than that of an l -type with no inheritance.²⁴ The reason for this is that the l -types with no inheritance

²³ Thus, an l -type who has inherited ω_ℓ and an l -type who has inherited $\widehat{\omega}_\ell$ would leave $\widehat{\omega}_\ell$ in bequests if they die early.

²⁴ This amounts to saying that we cannot conclude whether $\omega_\ell - \widehat{T}_\ell$ is larger than $-T_\ell$.

face a distortion on their consumption bundle and their lower utility level might very well be a result of this distortion. More interestingly, perhaps, is that our results do not guarantee that $\widehat{T}_\ell - T_\ell$ is positive. In other words, the ℓ -types might even face a bequest subsidy.

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 $U_h > U_\ell$.*
- (ii) *Second-best allocation of the h -types in the positive- and zero-inheritance groups differs according to $\widehat{c}_h < c_h$, $\widehat{d}_h = d_h$, and $L_h = L_h$, so we have $\widehat{U}_h < U_h$.*
- (iii) *In every generation, some of the ℓ -types have inherited $\widehat{\omega}_\ell = \widehat{d}_\ell$, some $\omega_\ell = d_\ell$, and some nothing. Second-best allocation for the ℓ -types is such that all individuals who have received an inheritance will, regardless of their inheritance level, receive the same consumption bundle and enjoy the same level of utility \widehat{U}_ℓ . Allocations of the positive- and zero-inheritance groups differ according to $\widehat{d}_\ell > d_\ell$ and $\widehat{L}_\ell > L_\ell$, so we have $\widehat{U}_\ell > U_\ell$.*
- (iv) *Decentralization of second-best allocations requires the following: (a) marginal saving subsidies; (b) high-ability individuals who face a bequest tax of more than 100 percent ($\widehat{T}_h - T_h > \omega_h$); (c) low-ability individuals who face a bequest tax that can be smaller, as well as larger, than 100 percent and might even be negative ($\widehat{T}_\ell - T_\ell < \omega_\ell$, $\widehat{T}_\ell - T_\ell > \omega_\ell$, or $\widehat{T}_\ell - T_\ell < 0$).*

V. Opting Out

A striking feature of our results is that the bequest tax might exceed 100 percent. We might consider this possibility as unrealistic, in the sense

that legal systems often allow individuals to refuse an inheritance. From a strictly informational perspective, this should make no difference to our set-up. The individual who refuses a bequest can still be tagged (the information on his ability status has been revealed). On practical grounds, however, this poses a challenge to our results. Another objection is that children might prevail upon their parents to write a will that leaves their estate, in the case of their death, to charitable organizations (and not to their children). To address these objections, and to have implementing tax functions that are compatible with existing legal structures, in this section we reconsider the question of accidental bequests under the assumption that one cannot be made worse off as a result of receiving an inheritance.

To study this problem, all we have to do is to add two new incentive constraints, $\widehat{U}_h \geq U_h$ and $\widehat{U}_\ell \geq U_\ell$, to our previous second-best problem. Recalling that the solution to that problem satisfied the latter of these two constraints, we proceed as follows. We impose only the $\widehat{U}_h \geq U_h$ constraint on the problem and verify *ex post* that the solution does not violate the other constraint, $\widehat{U}_\ell \geq U_\ell$. The Lagrangian expression of the government's problem is then written as

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

On the basis of the first-order conditions, we can derive three conclusions regarding the nature of the solution in this case, and how it compares to the second-best solution we had previously.

First, the presence of the extra constraint $\widehat{U}_h = U_h$ implies that we no longer offer a non-distorted solution to the individuals who have an inheritance. In particular, the solution is no longer characterized by equal utilities for *h*- and *ℓ*-types (i.e., we no longer have $\widehat{U}_h = \widehat{U}_\ell$). To see this, we combine the first-order conditions with respect to \widehat{c}_ℓ and \widehat{c}_h in order to obtain

$$n_h(1 - \pi_h)[v'(\widehat{U}_h) - v'(\widehat{U}_\ell)] + \gamma = 0.$$

This relationship implies $v'(\widehat{U}_h) - v'(\widehat{U}_\ell) < 0$ so that $\widehat{U}_h > \widehat{U}_\ell$. This is quite interesting. It tells us that constraining \widehat{U}_h so that it is not smaller than U_h means that the *h*-types of the positive-inheritance group enjoy a

rent that they did not have previously (compared to the ℓ -types in the same group).

Second, consider how the h -types fare in the two tagged groups. We use the first-order condition with respect to \widehat{c}_h and c_h , together with $\widehat{U}_h = U_h$, and we rearrange to obtain $\lambda - \gamma = \pi_h \lambda$.

Now, we replace γ in the equation $\partial \mathcal{L}_V / \partial \widehat{d}_h = 0$ by $(1 - \pi_h)\lambda$, and $(\lambda - \gamma)$ in equation $\partial \mathcal{L}_V / \partial d_h = 0$ by $\pi_h \lambda$. Then we compare the resulting expressions, using $\widehat{U}_h = U_h$, to show that $\widehat{d}_h = d_h$. Using a similar argument, we establish that $\widehat{I}_h = I_h$. This tells us that the h -types are treated identically whether they receive an inheritance or not; pooling them together is optimal. The implication of this result is that h -types should now face a confiscatory tax on their inheritance. This is not surprising. Without the added $\widehat{U}_h \geq U_h$ constraint, \widehat{U}_h will be smaller than U_h and h -types face a bequest tax of more than 100 percent. With the added constraint, and given that pooling is optimal, we cannot go beyond 100 percent.

Third, turning to the ℓ -type individuals, we observe that the first-order conditions remain exactly the same as previously (in Section IV.). Consequently, all our results pertaining to this group remain valid. This also serves as the *ex post* verification that the $\widehat{U}_\ell \geq U_\ell$ constraint we had ignored is not violated. As far as the tax treatment of bequests is concerned, we find, as previously, that the tax for ℓ -types can be anything from a subsidy to a tax that exceeds the 100 percent mark.

These results are summarized as follows.

Proposition 3. *Consider the society described in Proposition 2, but assume that nobody can be made worse off as a result of receiving an inheritance.*

- (i) *The high-ability types in the positive-inheritance group enjoy a rent as compared to the low-ability types in the group. That is, $\widehat{U}_h > \widehat{U}_\ell$ replaces the $\widehat{U}_h = \widehat{U}_\ell$ result of Proposition 2.*
- (ii) *All high-ability types receive the same allocation, regardless of their inheritance status. Thus, $\widehat{U}_h = U_h$ replaces $\widehat{U}_h < U_h$, and a bequest tax of 100 percent replaces a bequest tax of more than 100 percent, results of Proposition 2.*
- (iii) *All the results of Proposition 2 pertaining to the low-ability types remain valid. In particular, low-ability types face a bequest tax that can be smaller, as well as larger, than 100 percent, and might even be negative.*

VI. Summary and Conclusion

In this paper, we have questioned the validity of the conventional wisdom that purely accidental bequests should be taxed at a confiscatory rate. We have employed a model wherein individuals of different abilities might

live for one or two periods with different probabilities of survival. We have shown that the proposition is correct in a first-best environment, when individuals' productivity and longevity are publicly observable. Under this circumstance, we can mimic a perfect annuity market by subsidizing the saving of each ability type at a rate equal to their probability of an early death, in conjunction with lump-sum taxes that vary according to individuals' ability types and inheritance status. All accidental bequests are taxed at 100 percent 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 group consists 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 remain 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 they will end up with the same utility level as the low-ability types in this group. However, the 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. A 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, we must levy marginal saving subsidies that vary with income and inheritance status, but not with inheritance level. High-ability individuals face a bequest tax of more than 100 percent, while low-ability individuals face a bequest tax that can be smaller, as well as larger, than 100 percent, and might even be negative.

Finally, we have studied the implications for our results if people are able to refuse the inheritances that are due to them. We have shown that, in this case, high-ability types in the positive-inheritance group enjoy a rent as compared to the low-ability types in the group. We have also shown that all high-ability types are pooled together and receive the same allocation, regardless of their inheritance status. Consequently, they will face a bequest tax of 100 percent rather than one that exceeds 100 percent. As far as the low-ability types are concerned, however, all of our previous results remain intact. Specifically, low-ability types continue to face a bequest tax that can be smaller, as well as larger, than 100 percent, and might even be negative.

We conclude by revisiting two of our main assumptions. First, we have considered a highly stylized setting, wherein the observability of bequests

brings about a drastic change in the structure of information available to the tax authority. This drastic change comes about from the assumption that types are dynastically immutable: if a person is of a given type, their offsprings will also be of the same type. In following this approach, we have been led by a desire to convey our point in a crisp fashion with no ambiguity. A more realistic setting posits only that there is a high probability – and not a certainty – that children will be of the same type as their parents. Our main point, that incorporating the informational content of bequests improves the design of tax structures, including bequest taxes, should remain valid in this more realistic setting. The specifics of the optimal tax policy would, of course, become more complicated.

Second, we have assumed that individual's preferences are quasi-linear. Accordingly, in *laissez-faire*, an individual's second-period consumption, and thus their saving, which constitutes the accidental bequest, does not depend on the bequest received (there is no income effect). This assumption simplifies the analysis but it is not crucial for the main results of our paper. To be more precise, its relevance is confined to the *laissez-faire* allocation; limiting the equilibrium levels of positive accidental bequests at the steady state (to two). It plays no such role for first- and second-best allocations. With more general preferences too, the tagging is between those who inherit nothing and those who inherit something (not how much). Thus, it remains optimal to treat all individuals of a given type identically, regardless of the size of their inheritance. Then, we will have four bequest levels, rather than three, under quasi-linear preferences. As long as we can order the four bequest levels, all our results will go through.

Appendix: 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

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

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

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

We denote the partial derivative of $T(\cdot)$ with respect to s by $T_s(\cdot)$. The first-order condition with respect to s , whether an individual is in the positive- or zero-inheritance group, is then equal to

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

Substituting the second-best value of $\beta_i \phi'(s)$ from equation (24) for everyone who receives an inheritance, and from equations (26) and (29) for h - and ℓ -types who do not receive an inheritance, yields the following marginal saving subsidies:

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

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

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

References

- Akerlof, G. A. (1978), The Economics of “Tagging” as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning, *American Economic Review* 68, 8–19.
- Atkinson, A. B. (1970), On the Measurement of Inequality, *Journal of Economic Theory* 2, 244–263.
- Atkinson, A. B. and Stiglitz, J. E. (1976), The Design of Tax Structure: Direct versus Indirect Taxation, *Journal of Public Economics* 6, 55–75.
- Banks, J. and Diamond, P. (2010), The Base for Direct Taxation, in *The Mirrlees Review*, Oxford University Press, Oxford, 548–648.
- Blumkin, T. and Sadka, E. (2004), Estate Taxation with Intended and Accidental Bequests, *Journal of Public Economics* 88, 1–21.
- Boadway, R. and Pestieau, P. (2006), Tagging and Redistributive Taxation, *Annales d’Economie et de Statistique* 83/84, 123–150.
- Bommier, A. (2006), Uncertain Lifetime and Intertemporal Choice: Risk Aversion as a Rationale for Time Discounting, *International Economic Review* 47, 1223–1246.
- Cremer, H. and Gahvari, F. (1995), Uncertainty, optimal taxation and the direct versus indirect tax controversy, *Economic Journal*, 105, 1165–79.
- Cremer, H. and Pestieau, P. (2005), Wealth Transfer Taxation: A Survey of the Theoretical Literature, in L. A. Gerard-Varet, S.-Ch. Kolm, and J. Mercier-Ythier (eds.), *Handbook on the Economics of Giving, Reciprocity and Altruism*, North-Holland, Amsterdam.
- Cremer, H., Pestieau, P., and Rochet, J.-C. (2003), Capital Income Taxation when Inherited Wealth is not Observable, *Journal of Public Economics* 87, 2475–2490.
- Cremer, H., Gahvari, F., and Lozachmeur, J.-M. (2010), Tagging and Income Taxation: Theory and an Application, *American Economic Journal: Economic Policy* 2, 31–50.
- Deaton, A. S. and Paxson, C. H. (1998), Aging and Inequality in Income and Health, *American Economic Review* 88, 248–253.
- Diamond, P. and Spinnewijn, J. (2011), Capital Income Taxes with Heterogeneous Discount Rates, *American Economic Journal: Economic Policy* 3, 52–76.
- Duleep, H. O. (1986), Measuring the Effect of Income on Adult Mortality using Longitudinal Administrative Record Data, *Journal of Human Resources* 21, 238–251.
- Kaplow, L. (2008), *The Theory of Taxation and Public Economics*, Princeton University Press, Princeton, NJ.
- Kopczuk, W. (2003a), The Trick is to Live: Is the Estate Tax Social Security for the Rich?, *Journal of Political Economy* 111, 1318–1341.
- Preston, S. H. (1975), The Changing Relation between Mortality and the Level of Economic Development, *Population Studies* 29, 231–248.

- Pritchett, L. and Summers, L. (1996), Wealthier is Healthier, *Journal of Human Resources* 31, 841–868.
- Saez, E. (2002), The Desirability of Commodity Taxation under Non-Linear Income Taxation and Heterogeneous Tastes, *Journal of Public Economics* 83, 217–230.
- Salm, M. (2007), The Effect of Pensions on Longevity: Evidence from Union Army Veterans, IZA Discussion Paper 2668.
- Singh, G. K. and Siahpush, M. (2006), Widening Socioeconomic Inequalities in US Life Expectancy, 1980–2000, *International Journal of Epidemiology* 35, 969–979.
- Waldron, H. (2007), Trends in Mortality Differentials and Life Expectancy for Male Social Security-Covered Workers, by Socioeconomic Status, *Social Security Bulletin* 67, 1–28 (<http://www.ssa.gov/policy/docs/ssb/v67n3/v67n3p1.pdf>).

First version submitted November 2010;
final version received July 2011.