Cremer, Helmuth; Gahvari, Firouz; Pestieau, Pierre

Working Paper
Accidental bequests: A curse for the rich and a boon for the poor

CESifo Working Paper, No. 3094

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Cremer, Helmuth; Gahvari, Firouz; Pestieau, Pierre (2010) : Accidental bequests: A curse for the rich and a boon for the poor, CESifo Working Paper, No. 3094, Center for Economic Studies and Ifo Institute (CESifo), Munich

This Version is available at:
http://hdl.handle.net/10419/38890

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.
If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Accidental Bequests: A Curse for the Rich and a Boon for the Poor

Helmuth Cremer
Firouz Gahvari
Pierre Pestieau

CESifo Working Paper No. 3094
Category 1: Public Finance
June 2010

An electronic version of the paper may be downloaded
• from the SSRN website: www.SSRN.com
• from the RePEc website: www.RePEc.org
• from the CESifo website: www.CESifo-group.org/wp
Accidental Bequests: A Curse for the Rich and a Boon for the Poor

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% and may even be negative.


Keywords: accidental bequests, estate tax, tagging, first best, second best.
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. 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. Yet there is little agreement regarding the relative strength of each of these elements. Studies suggest that the size of the estate tax is highly sensitive to the relative importance of different bequest motives. 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).
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 and relies on the understanding that bequests have informational content that should be incorporated in the design of optimal tax structures.

The paper considers a setting where accidental bequests signal otherwise unobservable individual characteristics such as productivity and longevity. In this setup, ignoring the informational content of bequests calls for a 100% tax rate on bequests (within the context of the tax system as a whole and when tax instruments are not artificially restricted).\footnote{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 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.} We show this result is unwarranted when the informational content of accidental bequests is taken into account.

The basic idea is for the tax administration to use the information on individuals’ inheritances 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 consisting of
people who do not receive an inheritance and the other 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. Two interesting and inter-related results emerge. One 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% but that the low-ability types face a tax rate that can be smaller or larger than 100% and may even be negative.

The fundamental message underlying these results is that publicly observable bequests have informational content that should be incorporated in the design of optimal tax structures. This is a general point that applies to all types of bequests regardless of bequest motives. However, the paper considers accidental bequests only and uses a stylized model as the vehicle for demonstrating its results. It does so in order to make this point stark and in a simplified way. That we concentrate on accidental bequests does not mean that other bequest motives do not matter. Nor do we claim that accidental bequests are the most important type of bequests. Their choice 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.

Finally, in emphasizing the informational content of bequests, the paper balances the availability of information with practical considerations. Thus we do not allow the tax liability of an individual to be based directly on his parents’ income. From a purely theoretical (mechanism) design perspective, 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 his full earning history. It is
these practical considerations that motivate our conditioning an individual’s tax liability on his 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 as the information that can be obtained from bequests could also be obtained from these other variables. Put differently, we would have a redundancy in the informational structure.

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 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 $\pi_i$, and taste for future versus present consumption. To model the taste difference, we assign a weight $\beta_i$ 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 $\ell$. We further assume that these characteristics are positively correlated so that there are only two types of people: $h$ and $\ell$ with $w_h > w_\ell, \pi_h > \pi_\ell$, and $\beta_h > \beta_\ell$. Types are

\[ \text{4 It is sufficient to assume } \pi_h \geq \pi_\ell \text{ and } \beta_h \geq \beta_\ell \text{ with one of the two inequalities being strict.} \]
dynastically immutable: if a person is of type $i$, his offsprings will also be of type $i$.\footnote{That bequests have informational content whose use improves the design of bequest taxes remains valid even if this assumption is dropped. However, the underlying tagging problem would become significantly more complicated. See the Conclusion for further discussion of this issue.}

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, $L_i$.$^7$ An individual’s expected utility is given by

$$U_i = \pi_i [c_i + \beta_i \phi (d_i) - \varphi (L_i)] + (1 - \pi_i) [c_i - \varphi (L_i)]$$

$$= c_i + \pi_i \beta_i \phi (d_i) - \varphi (L_i), \quad i = h, \ell,$$  \hspace{1cm} (1)

where $\phi$ is strictly concave while $\varphi$ is strictly convex.$^8$

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 an accidental bequest which, unless taxed away, is inherited by their children. Consequently, besides $w_i, \pi_i$ and $\beta_i$, individuals of the second and forthcoming generations will differ also on the basis of their inherited wealth, $\omega_i$. 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 $\omega_h$ or no inherited wealth and $\ell$-types with either $\omega_\ell$ or no inherited wealth.$^9$

\footnote{This formulation assumes that an individual derives no utility from leaving an accidental bequest for his family if he dies early.}

\footnote{Dependence of the bequest one leaves on the inheritance one receives leads to multiplicity of groups on the basis of $\omega_i$, making existence of a steady state problematic. Observe also that even with quasi-}

\footnote{Preston (1975), and Pritchett and Summers (1996), provide empirical support for the existence of a positive correlation between ability and longevity; and Bommer (2006) for a positive correlation between longevity and preferences for future over present consumption.}

\footnote{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 result, derived below, carry over with only minor modifications to a setting with general preferences. See the last paragraph in the Conclusion.}
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

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

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 $\pi_h > \pi_\ell$, $\beta_h > \beta_\ell$, strict concavity of $\phi$, and strict convexity of $\varphi$, 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 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

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

linear preferences $\omega_i$ 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.
where the symbol \(^\hat{\cdot}\) over a variable indicates that it refers to a person who has received an inheritance, and we have substituted \(w_i \hat{L}_i + \omega_i - \hat{s}_i\) for \(\hat{c}_i\) and \(\hat{s}_i\) for \(\hat{d}_i\). The quasi-linearity of preferences implies that future consumption and labor supply do not depend on inherited wealth so that \(\hat{b}_s = s_i\) and \(\hat{b}_L = L_i\). However, present consumption increases with wealth and \(\hat{b}_c > c_i\). Clearly, with \(\hat{s}_i = s_i\), \(\hat{\omega}_i = \omega_i\) 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 \(\hat{s}_h > \hat{s}_\ell\) and \(\hat{L}_h > \hat{L}_\ell\).

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 \(n_\ell\) 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 \(\omega_i\)), are publicly observable. The first-best policy is attained when the government chooses \((c_i, d_i, L_i)\) and \((\hat{c}_i, \hat{d}_i, \hat{L}_i)\) to maximize social welfare defined by

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

where \(v\) is a strictly concave transformation of the quasi-linear utility function (1); hence \(v' > 0\) and \(v'' < 0\). 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 \(\ell\)-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 \( v(U) = U^{1-\varepsilon} / (1 - \varepsilon) \), suggested by Atkinson (1973) and used in the numerical examples below. In this formulation, \( \varepsilon > 0 \ (\varepsilon \neq 1) \) denotes the inequality aversion index and the higher is \( \varepsilon \) the greater will be the desired redistribution.

The resource constraint for the economy is given by

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

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.\(^{10}\) First-best optimum is then characterized by choosing \((c_i, d_i, L_i)\) and \((\hat{c}_i, \hat{d}_i, \hat{L}_i)\) to maximize (4) subject to (5).\(^{11}\) Let \( \mu \) denote the Lagrangian multiplier associated with the resource constraint (5). The Lagrangian expression associated with this problem is

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

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

\[
v'(U_i) = v'\left(\hat{U}_i\right) = \mu, \tag{7}
\]

\[
\beta_i \phi'(d_i) = \beta_i \phi'(\hat{d}_i) = 1, \tag{8}
\]

\[
\varphi'(L_i) = \varphi'(\hat{L}_i) = \mu w_i. \tag{9}
\]

\(^{10}\)While inheritances one generation receives may in principle be different from the bequests it leaves, such an outcome cannot happen in the steady state.

\(^{11}\)There is also the constraint that \( s_i = w_i L_i - c_i \geq 0 \), which we assume to be non-binding.
Denoting the first-best values by superscript \( FB \), it follows from (8)–(9) that

\[
\hat{U}_i = U_i \equiv U_i^{FB}, \tag{10}
\]
\[
\hat{d}_h = d_h \equiv d_h^{FB} > \hat{d}_\ell = d_\ell \equiv d_\ell^{FB}, \tag{11}
\]
\[
\hat{L}_h = L_h \equiv L_h^{FB} > \hat{L}_\ell = L_\ell \equiv L_\ell^{FB}, \tag{12}
\]

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

\[
\hat{c}_i = c_i \equiv c_i^{FB},
\]

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 \( \ell \)-types differ but either one gets the same allocation regardless of receiving an inheritance or not.

### 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: \( (\hat{t}_h, \hat{t}_\ell) \) for those with an inheritance and \( (t_h, t_\ell) \) for those without. Naturally, these fiscal instruments must also satisfy the government’s budget constraint

\[
\sum_{i=h,\ell} n_i \left[ \pi_i (t_i - \tau_i s_i) + (1 - \pi_i) (\hat{t}_i - \tau_i \hat{s}_i) \right] = 0.
\]
To see how the optimum is decentralized, observe that the presence of $\tau_i, \hat{t}_i,$ and $t_i$ changes the budget constraints of the $i$-type with and without an inheritance to

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

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

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

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

These are identical to their first-best counterparts resulting in $\hat{L}_i = L_i = L_i^{FB}$ and $\hat{s}_i = s_i = s_i^{FB}$. Additionally, to ensure $\hat{c}_i = c_i = c_i^{FB}$, 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 $\hat{t}_i = w_i L_i^{FB} + \omega_i - c_i^{FB} - \pi_i s_i^{FB}$.

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

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

Now, given identical tax rates on savings, it is natural to consider the difference between $\hat{t}_i$ 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 ($\hat{s}_i = s_i$), both definitions lead to the same answer:12 The implementation of first-best allocations requires a 100% taxation of (acci-

---

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

$$T_i = t_i - \tau_i s_i,$$

$$\hat{T}_i = \hat{t}_i - \tau_i \hat{s}_i.$$

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

$$T_i = \hat{T}_i - \omega_i.$$
bequests. The \( i \)-type with an inheritance \( \omega_i, i = h, \ell \), 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 \( \pi_i \), and the utility weight they assign to their consumption in retirement \( \beta_i \). 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: \( \hat{U}_h = U_h = \hat{U}_\ell = U_\ell \), 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: \((\hat{t}_h, \hat{t}_\ell)\) for those with an inheritance and \((t_h, t_\ell)\) for those without.

(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 \((w_i, \pi_i, \beta_i)\), and labor supplies, \( L_i \), are not publicly observable. The observables are gross earnings \((\hat{I}_i = w_i \hat{L}_i, I_i = w_i L_i)\), consumption during working years and retirement \((\hat{c}_i, \hat{d}_i; c_i, d_i)\), and bequests \( \omega_i \). 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.\(^{13}\) The key point

\(^{13}\) Recall that 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-
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.\textsuperscript{14}

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. 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.\textsuperscript{15} Put differently, one assigns a single resource constraint to the two groups. This continues to be represented by (5). Let $\mu$ denote the Lagrange multiplier associated with the economy’s resource constraint, and $\lambda$ the multiplier associated with the incentive constraint in the group of individuals who receive no inheritance. The Lagrangian expression associated with this

\textsuperscript{14}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.

\textsuperscript{15}See, e.g., Cremer et al. (2010).
optimization problem is
\[ L = \sum_{i=h,\ell} n_i \left\{ \pi_i \left[ v \left( c_i + \pi_i \beta_i \phi(d_i) - \varphi\left( \frac{I_i}{w_i} \right) \right) + \mu (I_i - c_i - \pi_i d_i) \right] 
+ (1 - \pi_i) \left[ v \left( \hat{c}_i + \pi_i \beta_i \phi(\hat{d}_i) - \varphi(\hat{L}_i) \right) + \mu \left( w_i \hat{L}_i - \hat{c}_i - \pi_i \hat{d}_i \right) \right] \right\} 
+ \lambda \left[ c_h + \pi_h \beta_h \phi(d_h) - \varphi\left( \frac{I_h}{w_h} \right) - c_\ell - \pi_h \beta_h \phi(d_\ell) + \varphi\left( \frac{I_\ell}{w_\ell} \right) \right]. \] (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 \( \hat{L}_i \)—an observable as in Section 3.\(^{16}\)

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 (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:
\[ \frac{\partial L}{\partial c_i} = n_i (1 - \pi_i) \left[ v'\left( \bar{U}_i \right) - \mu \right] = 0, \] (15)
\[ \frac{\partial L}{\partial d_i} = n_i (1 - \pi_i) \left[ v'\left( \bar{U}_i \right) \pi_i \beta_i \phi' \left( \bar{d}_i \right) - \mu \pi_i \right] = 0, \] (16)
\[ \frac{\partial L}{\partial L_i} = n_i (1 - \pi_i) \left[ -v'\left( \bar{U}_i \right) \phi' \left( \bar{L}_i \right) + \mu w_i \right] = 0. \] (17)

\(^{16}\)Although one can equally express the problem in terms of \( \hat{I}_i \).
and for those without an inheritance:

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

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

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

\[
\frac{\partial L}{\partial c_c} = n_c \pi_c \left[ v'(U_c) - \mu \right] + \lambda = 0,
\]

\[
\frac{\partial L}{\partial d_c} = n_c \pi_c \left[ v'(U_c) \pi_c \beta_c \phi'(d_c) - \mu \pi_c \right] - \lambda \pi_c \beta_c \phi'(d_c) = 0,
\]

\[
\frac{\partial L}{\partial I_c} = n_c \pi_c \left[ -v'(U_c) \frac{1}{w_c} \phi' \left( \frac{I_c}{w_c} \right) + \mu \right] + \frac{\lambda}{w_c} \phi' \left( \frac{I_c}{w_c} \right) = 0.
\]

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

\[
v'(\hat{U}_i) = \mu,
\]

\[
\beta_i \phi'(\hat{d}_i) = 1,
\]

\[
\phi'(\hat{L}_i) = w_i.
\]

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, \( \hat{U}_h = \hat{U}_\ell, \hat{d}_h > \hat{d}_\ell, \) and \( \hat{L}_h > \hat{L}_\ell. \) 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)–
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: $\beta_\ell \phi' (d_\ell) > 1$ and $w_\ell > \varphi' (L_\ell)$. In words, consumption of $d_\ell$ and supply of $L_\ell$ 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 strict concavity of $\phi$ and strict convexity of $\varphi$, 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

$$\hat{d}_h = d_h, \quad \hat{L}_h = L_h.$$
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

$$\hat{U}_h < U_h. \quad (35)$$

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

$$\hat{c}_h < c_h. \quad (36)$$

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 $c$-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

$$\hat{d}_c > d_c, \quad \hat{L}_c > L_c. \quad (37)$$

An $c$-type who receives an inheritance consumes more in the second period and works more in the first as compared to an $c$-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

$$\hat{U}_c > U_c. \quad (39)$$

The $c$-type enjoys a higher level of utility if he receives an inheritance. However, these inequalities do not allow us to compare $\hat{c}_c$ and $c_c$. 

16
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 \( \hat{d}_h = d_h > \hat{d}_\ell > d_\ell \). Given these properties, leaving behind \( \hat{d}_h = d_h \) indicates that the deceased must have been of type \( h \) while leaving either \( \hat{d}_\ell \) or \( d_\ell \) indicates type \( \ell \). The assumption of a dynastically immutable family type then establishes the recipient’s type.

### 4.2 Tax policy

Tax policy is set in order to implement the second-best allocations characterized by (15)--(17) for the \( h \)- and \( \ell \)-types in the positive-inheritance group and (18)--(23) for the \( h \)- and \( \ell \)-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.  

In what follows, we state the properties of this function with respect to income and savings briefly, and then 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 are well-known. The properties of \( T(I, s, \omega) \) with respect to saving is 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 \( \ell \)-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 \( \ell \)-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.

\[ \text{17 One could also include } e \text{ which can always be inferred from observability of the other variables. However, this argument would be redundant.} \]
rate. This subsidy rate is, however, independent of the inheritance level, $\omega$ or $b$.

Turning to bequest taxation, consider first the treatment of high-ability individuals. They inherit either nothing or $\omega_h = d_h = \hat{d}_h$. 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 $\hat{T}_h = T(\hat{I}_h, \hat{s}_h, \omega_h) = \hat{I}_h + \omega_h - \hat{c}_h - \hat{d}_h = \hat{I}_h - \hat{c}_h$ if they do. Now, from equations (33)–(36), $\hat{d}_h = d_h, \hat{I}_h = I_h$, and $\hat{c}_h < c_h$. Hence $\hat{T}_h - T_h = d_h + (c_h - \hat{c}_h) = \omega_h + (c_h - \hat{c}_h)$, and

$$\hat{T}_h - T_h > \omega_h.$$ Consequently, the difference between an $h$-type’s total tax liabilities if he receives an inheritance and if he does not, exceeds the inheritance he may receive. 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 a high-ability person’s type is revealed when he inherits $\omega_h$, but that his type remains unidentified otherwise. With his ability known, an $h$-type who receives an inheritance enjoys no “informational rent” over an $\ell$-type who too 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 $\ell$-types. On the other hand, since the ability of an $h$-type who inherits nothing remains private information, he enjoys some informational rent.

Comparing how the $\ell$-types in positive- and zero-inheritance groups fare is rather more complicated. In any given generation, some of the $\ell$-types have received no inheritance, some $\omega_\ell = d_\ell$, and some $\omega_\ell = \hat{d}_\ell$. But an $\ell$-type’s allocation depends only on whether he has received an inheritance or not; those who have inherited $\omega_\ell$ get an identical allocation to those who have inherited $\omega_\ell$. Specifically, an $\ell$-type who inherits nothing plans for a future consumption level of $d_\ell$, while an $\ell$-type who has inherited

---

18Our 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 alongCremer and GaIhari (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.
either $\omega_\ell$ or $\hat{\omega}_\ell$ plans for a consumption level of $\hat{d}_\ell$.\footnote{Thus an $\ell$-type who has inherited $\omega_\ell$ and an $\ell$-type who has inherited $\hat{\omega}_\ell$ would leave $\hat{\omega}_\ell$ in bequests if they die early.} This requires three tax levels:

$$T_\ell = T(I_\ell, s_\ell, 0) = I_\ell - c_\ell - d_\ell$$ if the $\ell$-type inherits nothing, $\hat{T}_\ell = T(\hat{I}_\ell, \hat{s}_\ell, \omega_\ell) = \hat{I}_\ell + \omega_\ell - \hat{c}_\ell - \hat{d}_\ell$ if he inherits $\omega_\ell$, and $\tilde{T}_\ell = T(\tilde{I}_\ell, \tilde{s}_\ell, \tilde{\omega}_\ell) = \tilde{I}_\ell + \tilde{\omega}_\ell - \tilde{c}_\ell - \tilde{d}_\ell = \tilde{I}_\ell - \tilde{c}_\ell$, if he inherits $\tilde{\omega}_\ell$. Consequently, the inheritance tax paid by those who inherit $\omega_\ell$ is

$$\tilde{T}_\ell - T_\ell = \omega_\ell + (\tilde{I}_\ell - \tilde{c}_\ell - \tilde{d}_\ell) - (I_\ell - c_\ell - d_\ell)$$ and the inheritance tax paid by those who inherit $\tilde{\omega}_\ell$ is

$$\tilde{T}_\ell - T_\ell = \tilde{\omega}_\ell + (\tilde{I}_\ell - \tilde{c}_\ell - \tilde{d}_\ell) - (I_\ell - c_\ell - d_\ell).$$\footnote{The definition of the bequest tax is somewhat 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 $\omega$, 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 $T(I, s_i, \omega_i) - T(I, s_i, 0)$ as well as $T(\hat{I}, \hat{s}, \omega_i) - T(\hat{I}, \hat{s}, 0)$ which may not only yield different answers but more problematically are not even unambiguously defined by the optimal allocation. The problem is that $T(I, s_i, \omega_i)$ in the first expression and $T(\hat{I}, \hat{s}, \omega_i)$ 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: $T(\hat{I}, \hat{s}, \omega_i) - T(I, s_i, 0)$.

} Subtracting $\tilde{T}_\ell - T_\ell$ from $(\tilde{T}_\ell - T_\ell)$ to get

$$\tilde{T}_\ell - T_\ell = \hat{I}_\ell - \hat{c}_\ell - (\hat{I}_\ell + \omega_\ell - \hat{c}_\ell - \hat{d}_\ell) = \omega_\ell - \omega_\ell.$$ (40)

This tells us that the $\ell$-types who inherit $\omega_\ell$ pay, effectively, a tax on their extra inheritance, $\omega_\ell - \omega_\ell$, at a rate of 100%. One can then consider the tax paid on $\omega_\ell$ as consisting of two parts: One part is paid on inheritances up to $\omega_\ell$; this tax is identical to the inheritance tax paid by those who receive only $\omega_\ell$. This is followed by a tax on the remaining $\omega_\ell - \omega_\ell$ inheritance; this is levied at a confiscatory rate. Beyond this, one cannot, at this level of generality, determine whether the tax on $\omega_\ell$ is smaller or larger than 100%. Specifically, the finding $\hat{U}_\ell > U_\ell$ does not allow one to conclude that an $\ell$-type with an inheritance has a larger net disposable income than an $\ell$-type with no inheritance. Put differently, one cannot conclude that $\tilde{T}_\ell - \omega_\ell$ is smaller than $T_\ell$. The reason is that the $\ell$-types with no inheritance face a distortion on their consumption...
Table 1. Second-best allocations and bequest taxes

<table>
<thead>
<tr>
<th></th>
<th>(\varepsilon = 0.50)</th>
<th>(\varepsilon = 13)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>positive inheritance</td>
<td>zero inheritance</td>
</tr>
<tr>
<td>(c_h, c_h)</td>
<td>31.918</td>
<td>33.767</td>
</tr>
<tr>
<td>(d_h, d_h)</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>(I_h, I_h)</td>
<td>40.000</td>
<td>40.000</td>
</tr>
<tr>
<td>(\tilde{c}<em>{\ell}, c</em>{\ell})</td>
<td>17.026</td>
<td>15.052</td>
</tr>
<tr>
<td>(\tilde{d}<em>{\ell}, d</em>{\ell})</td>
<td>0.750</td>
<td>0.733</td>
</tr>
<tr>
<td>(\tilde{I}<em>{\ell}, I</em>{\ell})</td>
<td>10.000</td>
<td>9.504</td>
</tr>
<tr>
<td>(T_h, T_h)</td>
<td>8.082</td>
<td>5.233</td>
</tr>
<tr>
<td>(\tilde{T}<em>{\ell}, T</em>{\ell})</td>
<td>-7.043</td>
<td>-6.281</td>
</tr>
<tr>
<td>(\tilde{T}_{\ell})</td>
<td>-7.026</td>
<td>-6.809</td>
</tr>
</tbody>
</table>

bundle and their lower utility level may very well be due to this distortion. The example below illustrates this point and establishes that bequest taxes can be larger as well as smaller than 100%. More interestingly perhaps, our theoretical results do not guarantee that \(\tilde{T}_{\ell} - T_{\ell}\) is positive. Put differently, the \(\ell\)-types may even face a bequest subsidy. The example below illustrates this possibility as well.

4.3 An example

Consider a quadratic disutility for labor \(\varphi(L) = 0.05L^2\) and a logarithmic utility for future consumption \(\varphi(d) = \ln d\) so that our quasi-linear preferences (1) takes the form \(U_i = c_i + \pi_i \beta_i \ln d_i - 0.05L_i^2\). 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 at: \(w_h = 2, w_{\ell} = 1, n_h = n_{\ell} = 1, \pi_h = \pi_{\ell} = 0.5, \beta_h = 1, \beta_{\ell} = 0.75\). Table 1 reports the numerical solutions for the second-best allocations when \(\varepsilon = 0.5\) and \(\varepsilon = 13\). It also reports the values of total tax liabilities in each case.

The interesting property from our perspective is the behavior of \(\tilde{T}_{\ell} - T_{\ell}\) and its comparison to \(\omega_{\ell}\). With \(\varepsilon = 0.5\), we have \(\tilde{T}_{\ell} = -7.043\) and \(T_{\ell} = -6.281\) so that
\( \tilde{T}_\ell - T_\ell = -0.762 \). That is, bequests of the size \( \omega_\ell = 0.733 \) are subsidized. Interestingly, even the \( \ell \)-types who receive \( \tilde{\omega}_\ell \) are subsidized in this case. They pay a bequest tax of \( \tilde{\omega}_\ell - \omega_\ell = 0.750 - 0.733 = 0.017 \) on their extra inheritance so that they receive a net subsidy of \( 0.762 - 0.017 = 0.745 \). On the other hand, at \( \varepsilon = 13 \), one has \( \tilde{T}_\ell = -6.982 \) and \( T_\ell = -7.601 \) so that \( \tilde{T}_\ell - T_\ell = 0.619 \). This is greater than \( \omega_\ell = d_\ell = 0.577 \) and we have a tax that exceeds 100%. Comparing the numbers for \( \varepsilon = 0.5 \) and \( \varepsilon = 13 \) illustrates that the second-best tax on accidental bequests may fall short of as well as exceed a confiscatory rate.\(^{21}\) A second interesting feature of these numbers is that a higher desired degree of redistribution calls for a higher tax on bequests. Whereas a “small” degree of inequality aversion—a small \( \varepsilon \)—calls for a bequest subsidy, a “large” degree of inequality aversion—a large \( \varepsilon \)—results in 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 \( \hat{U}_h = \hat{U}_\ell \). (c) Individuals in the zero-inheritance group face no distortion at the top and a downward distortion on labor supply and savings for the \( \ell \)-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 \( \hat{c}_h < c_h, \hat{d}_h = d_h, L_h = L_h, \) and we have \( \hat{U}_h < U_h \).

\(^{21}\)Remember that a subsidy also means a less than 100% tax.
(iii) In every generation, some of the \( \ell \)-types have inherited \( \tilde{\omega}_\ell = \tilde{d}_\ell \), some \( \omega_\ell = d_\ell \), and some nothing. Second-best allocation for the \( \ell \)-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 \( \tilde{U}_\ell \). Allocations of the positive- and the zero-inheritance groups differ according to \( \tilde{d}_\ell > d_\ell \), \( \tilde{L}_\ell > L_\ell \), and we have \( \tilde{U}_\ell > U_\ell \).

(iv) Decentralization of second-best allocations requires: (a) Marginal saving subsidies; (b) High-ability individuals face a bequest tax of more then 100\% \( (\tilde{T}_h - T_h > \omega_h) \); (c) Low-ability individuals face a bequest tax that can be smaller as well as larger than 100\% and may even be negative \( (\tilde{T}_\ell - T_\ell < \omega_\ell, \tilde{T}_\ell - T_\ell > \omega_\ell, \text{ or } \tilde{T}_\ell - T_\ell < 0) \).

5 Opting out

A striking feature of our results is that the bequest tax may exceed 100\%. One may consider this possibility as “unrealistic” in that legal systems often allow individuals to refuse an inheritance. From a strictly informational perspective, this should make no difference for our setup. 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 may prevail upon their parents to write a will that leaves their estate, in case of 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, this section reconsiders the question of accidental bequests under the assumption that one can not be made worse off as a result of receiving an inheritance.

To study this problem, all one has to do is to add two new incentive constraints, \( \tilde{U}_h \geq U_h \) and \( \tilde{U}_\ell \geq U_\ell \), 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 \( \tilde{U}_h \geq U_h \) constraint on the problem and verify ex post that the
solution does not violate the other constraint, \( \hat{U}_h \geq U_\ell \). The Lagrangian expression of the government’s problem is then written as

\[
L_V = \sum_{i=h,\ell} n_i \left\{ \pi_i \left[ v \left( c_i + \pi_i \beta_i \phi(d_i) - \varphi\left( \frac{I_i}{w_i} \right) \right) + \mu \left( I_i - c_i - \pi_i d_i \right) \right]
+ (1 - \pi_i) \left[ v \left( \tilde{c}_i + \pi_i \beta_i \phi(d_i) - \varphi\left( \frac{\hat{I}_i}{w_i} \right) \right) + \mu \left( \hat{I}_i - \tilde{c}_i - \pi_i \hat{d}_i \right) \right] \right\}
+ \lambda \left[ c_h + \pi_h \beta_h \phi(\hat{d}_h) - \varphi\left( \frac{\hat{I}_h}{w_h} \right) - c_\ell - \pi_h \beta_h \phi(\hat{d}_\ell) + \varphi\left( \frac{I_\ell}{w_\ell} \right) \right]
+ \gamma \left[ \tilde{c}_h + \pi_h \beta_h \phi(\hat{d}_h) - \varphi\left( \frac{\hat{I}_h}{w_h} \right) - c_h - \pi_h \beta_h \phi(\hat{d}_h) + \varphi\left( \frac{I_h}{w_h} \right) \right].
\]

The first-order conditions are given in Appendix B. On the basis of these equations, one 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 \( \hat{U}_h = U_h \) implies that we no longer offer the individuals with an inheritance a non-distorted solution. In particular, the solution is no longer characterized by equal utilities for \( h \)- and \( \ell \)-types (i.e., we no longer have \( \hat{U}_h = \hat{U}_\ell \)). To see this, substitute for \( \mu \) from (47) into (44) to get

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

This relationship implies \( v' \left( \hat{U}_h \right) - v' \left( \hat{U}_\ell \right) < 0 \) so that \( \hat{U}_h > \hat{U}_\ell \). This is quite interesting. It tells us that constraining \( \hat{U}_h \) not to be smaller than \( U_h \) accords the \( h \)-types of the positive-inheritance group to enjoy a rent that they did not have previously (compared to the \( \ell \)-types in the same group).

Second, consider how the \( h \)-types fare in the two tagged groups. Use (44) and (48) together with \( \hat{U}_h = U_h \) to get

\[
\frac{\gamma}{(1 - \pi_h)} = \frac{\lambda - \gamma}{\pi_h}.
\]

Rearranging this relationship yields

\[
\lambda - \gamma = \pi_h \lambda.
\]
Now replace $\gamma$ in (45) by $(1 - \pi_h)\lambda$, $(\lambda - \gamma)$ in (49) by $\pi_h \lambda$, and compare the resulting expressions, using $\hat{U}_h = U_h$, to show that $\hat{d}_h = d_h$. A similar argument establishes that $\hat{I}_h = I_h$. 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 $\hat{U}_h \geq U_h$ constraint, $\hat{U}_h$ will be smaller than $U_h$ and $h$-types face a more than 100% bequest tax. With the added constraint, and given that pooling is optimal, one cannot go beyond 100%.

Third, turning to the $\ell$-type individuals, we observe that the first-order conditions remain exactly the same as previously (in Section 4). Consequently, all our results pertaining to this group remain valid. This also serves as the ex-post verification that the $\hat{U}_\ell \geq U_\ell$ constraint we had ignored is not violated. As far as the tax treatment of bequests are concerned, we have, as previously, that the tax for $\ell$-types can be anything from a subsidy to a tax that exceeds the 100% mark.

These results are summarized as:

**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, $\hat{U}_h > \hat{U}_\ell$ replaces $\hat{U}_h = \hat{U}_\ell$ result of Proposition 2.

(ii) All high-ability types receive the same allocation regardless of their inheritance status. Thus $\hat{U}_h = U_h$ replaces $\hat{U}_h < U_h$, and bequest tax of 100% replaces a bequest tax of more than 100%, 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% and may even be negative.
6 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 but not with the inheritance
level. High-ability individuals face a bequest tax of more than 100%, while low-ability individuals face a bequest tax that can be smaller as well as larger than 100% and may even be negative.

Finally, we studied the implications for our results if people are able to refuse the inheritances that are due to them. We showed 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 also showed 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% rather than one which exceeds 100%. 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% and may even be negative.

We conclude by revisiting two of the paper’s 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, his 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—for children to 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 the laissez-faire, an individual’s second-period consumption, and thus his 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 
\textit{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). It thus remains optimal to treat all
individuals of a given type identically regardless of the size of their inheritance. We will
then have four bequest levels, rather than three under quasilinear preferences. As long
as one can order the four bequest levels all our results will go through.
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

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

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

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

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

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

Substituting the second-best value of $\pi_i \beta_i \phi'(s)$ from (25) for everyone who receives an inheritance, and from (28) and (31) for $h$- and $c$-types who do not receive an inheritance, yields the following marginal saving subsidies:

$$-T_s(I_h, \tilde{s}_h, \omega) = 1 - \pi_h \beta_h \phi'(\tilde{s}_h) = 1 - \pi_h, \quad i = h, \ell, \quad (41)$$

$$-T_s(I_h, s_h, 0) = 1 - \pi_h \beta_h \phi'(s_h) = 1 - \pi_h, \quad (42)$$

$$-T_s(I_\ell, s_\ell, 0) = 1 - \pi_\ell \beta_\ell \phi'(s_\ell) = 1 - \pi_\ell \left[ \frac{\lambda \phi'(s_\ell)}{\mu n_\ell \pi_\ell} \left( \beta_h \frac{\pi_h}{\pi_\ell} - \beta_\ell \right) \right]. \quad (43)$$
Differentiating $\mathcal{L}_V$ yields the following first-order conditions

\[
\frac{\partial \mathcal{L}_V}{\partial c_h} = n_h (1 - \pi_h) \left[ v' \left( \hat{U}_h \right) - \mu \right] + \gamma = 0, \tag{44}
\]

\[
\frac{\partial \mathcal{L}_V}{\partial d_h} = n_h (1 - \pi_h) \left[ v' \left( \hat{U}_h \right) \pi_h \beta_h \phi' \left( \hat{d}_h \right) - \mu \pi_h \right] + \gamma \pi_h \beta_h \phi' \left( \hat{d}_h \right) = 0, \tag{45}
\]

\[
\frac{\partial \mathcal{L}_V}{\partial b} = n_h (1 - \pi_h) \left[ -v' \left( \hat{U}_h \right) \frac{1}{w_h} \varphi' \left( \hat{I}_h \right) + \mu \right] - \frac{1}{w_h} \varphi' \left( \hat{I}_h \right) = 0, \tag{46}
\]

\[
\frac{\partial \mathcal{L}_V}{\partial b} = n_h (1 - \pi_h) \left[ -v' \left( \hat{U}_h \right) \frac{1}{w_h} \varphi' \left( \hat{I}_h \right) + \mu \right] - \frac{1}{w_h} \varphi' \left( \hat{I}_h \right) = 0, \tag{47}
\]

\[
\frac{\partial \mathcal{L}_V}{\partial d} = n_h (1 - \pi_h) \left[ -v' \left( \hat{U}_h \right) \pi_h \beta_h \phi' \left( \hat{d}_h \right) - \mu \pi_h \right] + \gamma \pi_h \beta_h \phi' \left( \hat{d}_h \right) = 0, \tag{48}
\]

\[
\frac{\partial \mathcal{L}_V}{\partial b} = n_h (1 - \pi_h) \left[ -v' \left( U_h \right) \pi_h \beta_h \phi' \left( \hat{d}_h \right) - \mu \pi_h \right] + \lambda - \gamma = 0, \tag{49}
\]

\[
\frac{\partial \mathcal{L}_V}{\partial b} = n_h (1 - \pi_h) \left[ -v' \left( U_h \right) \pi_h \beta_h \phi' \left( \hat{d}_h \right) - \mu \pi_h \right] + \lambda - \gamma = 0, \tag{50}
\]

\[
\frac{\partial \mathcal{L}_V}{\partial b} = n_h \pi_h \left[ v' \left( U_h \right) - \mu \right] + \lambda - \gamma = 0, \tag{51}
\]

\[
\frac{\partial \mathcal{L}_V}{\partial d} = n_h \pi_h \frac{1}{w_h} \varphi' \left( \hat{I}_h \right) + \mu \right] - \frac{1}{w_h} \varphi' \left( \hat{I}_h \right) = 0, \tag{52}
\]

\[
\frac{\partial \mathcal{L}_V}{\partial d} = n_h \pi_h \left[ v' \left( U_h \right) \pi_h \beta_h \phi' \left( \hat{d}_h \right) - \mu \pi_h \right] + \lambda \pi_h \beta_h \phi' \left( \hat{d}_h \right) = 0, \tag{53}
\]
References


3033 Steven Brakman, Robert Inklaar and Charles van Marrewijk, Structural Change in OECD Comparative Advantage, April 2010

3034 Dirk Schindler and Gutorm Schjelderup, Multinationals, Minority Ownership and Tax-Efficient Financing Structures, April 2010

3035 Christian Lessmann and Gunther Markwardt, Decentralization and Foreign Aid Effectiveness: Do Aid Modality and Federal Design Matter in Poverty Alleviation?, April 2010

3036 Eva Deuchert and Conny Wunsch, Evaluating Nationwide Health Interventions when Standard Before-After Doesn’t Work: Malawi’s ITN Distribution Program, April 2010

3037 Eric A. Hanushek and Ludger Woessmann, The Economics of International Differences in Educational Achievement, April 2010

3038 Frederick van der Ploeg, Aggressive Oil Extraction and Precautionary Saving: Coping with Volatility, April 2010

3039 Ainura Uzagalieva, Evžen Kočenda and Antonio Menezes, Technological Imitation and Innovation in New European Union Markets, April 2010

3040 Nicolas Sauter, Jan Walliser and Joachim Winter, Tax Incentives, Bequest Motives, and the Demand for Life Insurance: Evidence from two Natural Experiments in Germany, April 2010

3041 Matthias Wrede, Multinational Capital Structure and Tax Competition, April 2010

3042 Burkhard Heer and Alfred Maussner, A Note on the Computation of the Equity Premium and the Market Value of Firm Equity, April 2010

3043 Kristiina Huttunen, Jukka Pirtilä and Roope Uusitalo, The Employment Effects of Low-Wage Subsidies, May 2010

3044 Matthias Kalkuhl and Ottmar Edenhofer, Prices vs. Quantities and the Intertemporal Dynamics of the Climate Rent, May 2010

3045 Bruno S. Frey and Lasse Steiner, Pay as you Go: A New Proposal for Museum Pricing, May 2010

3046 Henning Bohn and Charles Stuart, Population under a Cap on Greenhouse Gas Emissions, May 2010

3047 Balázs Égert and Rafal Kierzenkowski, Exports and Property Prices in France: Are they Connected?, May 2010
3048 Thomas Eichner and Thorsten Upmann, Tax-Competition with Involuntary Unemployment, May 2010

3049 Taiji Furusawa, Kazumi Hori and Ian Wooton, A Race beyond the Bottom: The Nature of Bidding for a Firm, May 2010

3050 Xavier Vives, Competition and Stability in Banking, May 2010

3051 Thomas Aronsson and Erkki Koskela, Redistributive Income Taxation under Outsourcing and Foreign Direct Investment, May 2010

3052 Michael Melvin and Duncan Shand, Active Currency Investing and Performance Benchmarks, May 2010

3053 Sören Blomquist and Laurent Simula, Marginal Deadweight Loss when the Income Tax is Nonlinear, May 2010

3054 Lukas Menkhoff, Carol L. Osler and Maik Schmeling, Limit-Order Submission Strategies under Asymmetric Information, May 2010

3055 M. Hashem Pesaran and Alexander Chudik, Econometric Analysis of High Dimensional VARs Featuring a Dominant Unit, May 2010

3056 Rabah Arezki and Frederick van der Ploeg, Do Natural Resources Depress Income Per Capita?, May 2010

3057 Joseph Plasman and Ruslan Lukach, The Patterns of Inter-firm and Inter-industry Knowledge Flows in the Netherlands, May 2010

3058 Jenny E. Ligthart and Sebastian E. V. Werner, Has the Euro Affected the Choice of Invoicing Currency?, May 2010


3060 Richard Cornes, Roger Hartley and Yuji Tamura, A New Approach to Solving Production-Appropriation Games with Many Heterogeneous Players, May 2010

3061 Ronald MacDonald and Flávio Vieira, A Panel Data Investigation of Real Exchange Rate Misalignment and Growth, May 2010

3062 Thomas Eichner and Rüdiger Pethig, Efficient Management of Insecure Fossil Fuel Imports through Taxing(!) Domestic Green Energy?, May 2010

3063 Vít Bubák, Evžen Kočenda and Filip Žikeš, Volatility Transmission in Emerging European Foreign Exchange Markets, May 2010

3064 Leonid V. Azarnert, Après nous le Déluge: Fertility and the Intensity of Struggle against Immigration, May 2010
William E. Becker, William H. Greene and John J. Siegfried, Do Undergraduate Majors or Ph.D. Students Affect Faculty Size?, May 2010

Johannes Becker, Strategic Trade Policy through the Tax System, May 2010

Omer Biran and Françoise Forges, Core-stable Rings in Auctions with Independent Private Values, May 2010

Torben M. Andersen, Why do Scandinavians Work?, May 2010


Simon Gächter, Benedikt Herrmann and Christian Thöni, Culture and Cooperation, June 2010

Mehmet Bac and Eren Inci, The Old-Boy Network and the Quality of Entrepreneurs, June 2010

Krisztina Molnár and Sergio Santoro, Optimal Monetary Policy when Agents are Learning, June 2010

Marcel Boyer and Donatella Porrini, Optimal Liability Sharing and Court Errors: An Exploratory Analysis, June 2010

Guglielmo Maria Caporale, Roman Matousek and Chris Stewart, EU Banks Rating Assignments: Is there Heterogeneity between New and Old Member Countries? June 2010

Assaf Razin and Efraim Sadka, Fiscal and Migration Competition, June 2010

Shafik Hebous, Martin Ruf and Alfons Weichenrieder, The Effects of Taxation on the Location Decision of Multinational Firms: M&A vs. Greenfield Investments, June 2010

Alessandro Cigno, How to Deal with Covert Child Labour, and Give Children an Effective Education, in a Poor Developing Country: An Optimal Taxation Problem with Moral Hazard, June 2010

Bruno S. Frey and Lasse Steiner, World Heritage List: Does it Make Sense?, June 2010

Henning Bohn, The Economic Consequences of Rising U.S. Government Debt: Privileges at Risk, June 2010

Rebeca Jiménez-Rodriguez, Amalia Morales-Zumaquero and Balázs Égert, The VARying Effect of Foreign Shocks in Central and Eastern Europe, June 2010

Stephane Dees, M. Hashem Pesaran, L. Vanessa Smith and Ron P. Smith, Supply, Demand and Monetary Policy Shocks in a Multi-Country New Keynesian Model, June 2010
3082 Sara Amoroso, Peter Kort, Bertrand Melenberg, Joseph Plasman and Mark Vancauteren, Firm Level Productivity under Imperfect Competition in Output and Labor Markets, June 2010

3083 Thomas Eichner and Rüdiger Pethig, International Carbon Emissions Trading and Strategic Incentives to Subsidize Green Energy, June 2010

3084 Henri Fraisse, Labour Disputes and the Game of Legal Representation, June 2010

3085 Andrzej Baniak and Peter Grajzl, Interjurisdictional Linkages and the Scope for Interventionist Legal Harmonization, June 2010

3086 Oliver Falck and Ludger Woessmann, School Competition and Students’ Entrepreneurial Intentions: International Evidence Using Historical Catholic Roots of Private Schooling, June 2010

3087 Bernd Hayo and Stefan Voigt, Determinants of Constitutional Change: Why do Countries Change their Form of Government?, June 2010

3088 Momi Dahan and Michel Strawczynski, Fiscal Rules and Composition Bias in OECD Countries, June 2010

3089 Marcel Fratzscher and Julien Reynaud, IMF Surveillance and Financial Markets – A Political Economy Analysis, June 2010

3090 Michel Beine, Elisabetta Lodigiani and Robert Vermeulen, Remittances and Financial Openness, June 2010

3091 Sebastian Kube and Christian Traxler, The Interaction of Legal and Social Norm Enforcement, June 2010

3092 Volker Grossmann, Thomas M. Steger and Timo Trimborn, Quantifying Optimal Growth Policy, June 2010


3094 Helmuth Cremer, Firouz Gahvari and Pierre Pestieau, Accidental Bequests: A Curse for the Rich and a Boon for the Poor, June 2010