Non-linear taxation of bequests, equal sharing rules and the tradeoff between intra- and inter-family inequalities

Helmut Cremer\textsuperscript{a,\*}, Pierre Pestieau\textsuperscript{b}
\textsuperscript{a}IDEI and GREMAQ, University of Toulouse, and Institut Universitaire de France, Toulouse, France
\textsuperscript{b}CREPP, University of Liège, CORE and DELTA, Liège, Belgium

Received 30 August 1998; received in revised form 1 February 1999; accepted 28 February 1999

Abstract

This paper studies the design of the tax and regulatory regime applied to bequests. Bequests are observable, while parent’s wealth and children’s earning abilities are not. Parents know their children’s earning abilities. Parents are altruistic; their utility depends on their children’s utilities, but weights may differ between children. The optimal tax schedule strikes a balance between the (often) conflicting ‘incentive’ and ‘corrective’ effects. When parents attach identical weights to their children, an estate taxation is sufficient. Equal sharing rules appear to be appropriate only in extreme cases such as in presence of the so-called Cinderella effect. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Taxation of bequests; Equal sharing rules; Inequality

JEL classification: H21; H23; D64; J13

1. Introduction

Death taxes may be imposed for a variety of reasons. One of them, on which this paper focuses, is redistribution. The institution of inheritance is a major factor responsible for concentration of wealth and, indirectly, for income inequality. According to most recent estimates, inherited wealth accounts for almost half of
the net worth of households.\footnote{For a survey of empirical studies on bequests, see Arrondel et al. (1997)} Death taxes, especially in the form of inheritance taxation, can thus (at least potentially) be used to moderate economic inequalities.

There are, however, more subtle effects of death taxation that might be significant; they are stressed by those who are less enthusiastic about such taxation. In addition to the alleged reduction of savings, there is the concern that death taxation may adversely affect equality. According to Becker (1974, 1991) and Tomes (1981), transfers between generations follow a ‘regression towards the mean’ mechanism with bequests and gifts flowing from well-to-do donors to less well-to-do recipients. Within each family, transfers thus tend to offset inequalities. Where this is true, taxation will mitigate the redistributive effect of wealth transfers.

The question one might raise at this point is why couldn’t the government directly effect the appropriate redistribution across and within generations. In a perfect information setting, it is clear that taxes on wealth transfers could be tuned in such a way that redistribution within families is not discouraged while redistribution across families is fostered. In an asymmetric information setting, however, this is less clear. Well to do families could be induced to leave lower bequests to avoid a too heavy tax burden; at the same time, the government could not be able to implement the ‘right’ redistribution because of imperfect information as to wealth holding.

It should be pointed out that the government can affect bequest behavior not only through the tax schedule, but also through the choice of the tax base or even through restrictions on estate sharing. Polar cases include estate taxes (based on the total amount which is bequeathed) and inheritance taxes (based on individual shares) or even accession taxes (based on individual shares plus other resources). Because tax schedules are typically non-linear (and progressive) the definition of the tax base is of crucial importance. In addition, bequests are subject to a variety of legal rules including more or less stringent equal sharing rules (amongst children).

Our paper studies the optimal design of the tax and regulatory regime applied to bequests. It is based on the observation that the determination of tax rates, tax bases and sharing rules can be treated in an integrated way by considering the design of a non-linear bequest tax schedule. For instance, a schedule which is separable in the individual shares corresponds to an inheritance tax while a tax function which depends on the sum of individual shares yields an estate tax. Sharing rules, on the other hand, can be implemented by imposing sufficiently high tax rates on sharing configurations which are to be ruled out.

Our setting has the following main features. Families are heterogenous and differ in the parent’s wealth. Each parent has two children who in turn differ in their income (earning ability). The tax administration observes bequests (including
individual shares) but neither the parent’s wealth nor the children’s earning abilities. Parents, on the other hand, are perfectly informed about their children’s earning abilities. They are altruistic in the sense that their own utility depends on their children’s utility (possibly in a non-symmetric way).

We start by characterizing the full-information optimum which is an interesting reference point. Under complete information and ‘perfect’ (non-discriminatory) altruism, the taxation of bequest is merely used to reduce inter-family inequalities. The intra-family allocation of wealth can be left to parents. Put differently, the optimal policy resembles an estate tax with no sharing rules. On the other hand, if parents are ‘imperfectly altruistic’ (and may attach a higher weight to one of their children), the optimal policy may include ‘corrective’ measures such as sharing rules.

Next, we study the design of optimal policies under asymmetric information, which constitutes a far more intricate question. Even under perfect altruism, the government may now find it optimal to intervene in the parent’s bequest decisions. Specially, we show that the optimal policy implies a positive (marginal) tax on bequests in low-wealth families. Consequently, there is a trade-off between the inter and intra-family allocation of wealth. It is desirable to distort the allocation of wealth within families in order to relax the self-selection constraint (to reduce the informational rents of high-wealth families). Sharing rules (whether explicitly imposed or implicitly enforced through the tax schedule), on the other hand, are no use within such a setting.

Then, we combine asymmetric information with imperfect altruism by allowing parents to attach different weights to the utilities of their children. The optimal tax schedule now combines incentive and corrective features. Depending on the parameter values, a wide variety of solutions can arise, including the possibility of negative (marginal) taxes on some categories of bequests. The optimal policy now also affects the allocation of bequests between children and in some cases it may be optimal to enforce strict equal sharing.

Finally, we present a variant of our model which shows that our analysis remains valid if other instruments of redistributive policy (such as income taxation) are also considered. While this is essentially a technical exercise (which is therefore relegated to Appendix A), it has important implications for the interpretation of our setting. In particular, it shows that our assumption that parents’ wealth is not observable, though maybe debatable, is effectively a meaningful shortcut for a more realistic information structure under which wealth (or income) is endogenous (and observable) but related to an unobservable characteristic (earning ability). In such a setting, inequalities in parents wealth can be reduced through (say) income taxation, but the unobservability of earning

\[\text{More precisely, it is jointly determined by an adverse selection and a moral hazard variable (ability and labor supply or effort).}\]
abilities precludes a full equalization of wealth. We show that the exogenous (and unobservable) wealth in our setting can effectively be interpreted as the ‘net income’ implied by an optimal (nonlinear) income taxation.

This argument also shows that bequest taxes do have a role to play as part of the optimal redistributive tax mix. They are not made redundant by the availability of an income tax (as long as income taxation is also restricted by informational asymmetries). Intuitively, this is because wealth transfers do provide additional information which proves valuable to achieve a better screening for the underlying unobservable characteristic.

2. The basic model

2.1. Parents’ preferences

Society consists of a large number of one-parent-two-children families, which can be of two types depending on parents’ wealth. There is an equal number of families of either type so that we can formally proceed as if there were only two families. Regardless of parents’ wealth, denoted by \( W_H \) and \( W_L \) for high and low wealth, each family comprises two children with unequal ability or lifetime earnings: \( w_1 \) and \( w_2 \), with \( w_2 > w_1 \). Parents are altruistic and have identical preferences represented by:

\[
u(y_i) + \gamma_1 u(c_{i1}) + \gamma_2 u(c_{i2}) \quad i = H, L,
\]

where \( y_i, c_{i1} \) and \( c_{i2} \) denote consumption of parents of type \( i \) and of children of ability 1 and 2 respectively, \( u \) is a strictly concave utility function, and \( \gamma_1 \) and \( \gamma_2 \) are the weights (altruism parameters) attached to children. We allow for different weights to reflect the possibility of discrimination against one of the children.

Let \( b_{i1} \) and \( b_{i2} \) be the bequests left by parents of wealth \( i \) to children of ability 1 and 2. The parent’s problem in the absence of death taxation is given by:

\[
\max_{b_{i1}, b_{i2}} u(W_i - b_{i1} - b_{i2}) + \gamma_1 u(w_1 + b_{i1}) + \gamma_2 u(w_2 + b_{i2}).
\]

When \( \gamma_1 = \gamma_2 = 1 \), an interior solution implies \( y_i = c_{i1} = c_{i2} \); there is perfect smoothing of consumption levels across generations and across children. On the

\[1\] Recall that in an optimal income tax model after tax incomes are not equalized; see, e.g. Stiglitz (1987).

\[2\] Net of the disutility of labor.

\[3\] See Section 4 and R. Brenner (1985a) and G. Brenner (1985b) for a discussion of the historical relevance of this phenomena. As we show below, the potential for such discrimination may justify specific non-linearities in the tax schedule as well as legal regulations.
other hand, when \( \gamma_1 \) and \( \gamma_2 \) are different and/or smaller than one (asymmetric and/or partial altruism), smoothing is not complete.

2.2. Wealth transfer tax

We now introduce the government. Its objective is a utilitarian social welfare function from which the altruistic components have been laundered.\(^6\) We further assume that the government only observes the amount of bequests but neither the parents’ wealth nor the children’s incomes. The government imposes (positive or negative) taxes on wealth transfers. Taxation is purely redistributive; there is no revenue requirement.

Let us first consider the full information problem to be used as a reference. It amounts to maximizing welfare given by

\[
W = u(W_H - b_{H1} - b_{H2} - T) + u(w_1 + b_{H1}) + u(w_2 + b_{H2}) \\
+ u(W_L - b_{L1} - b_{L2} + T) + u(w_1 + b_{L1}) + u(w_2 + b_{L2}),
\]

(1)

with respect to \( b_{ij} \) and \( T \).

The solution for an interior maximum implies a total equalization of disposable incomes; recall that all parents have identical preferences. Like in the case of the parent’s problem without taxation and with \( \gamma_1 = \gamma_2 = 1 \) there is perfect smoothing of consumption levels across generations and across children. Formally, one has:

\[
u'(y_H) = u'(y_L) = u'(c_{ij}) \quad i = H, L; j = 1, 2, \]

(2)

or

\[
y_H = y_L = c_{ij} = \frac{W_H + W_L + 2(w_1 + w_2)}{6} = \bar{w},
\]

(3)

with

\[b_{ij} = b_{Hij} = \bar{w} - w_j \quad \text{and} \quad T = \frac{W_H - W_L}{2}.
\]

Throughout the paper, we impose non-negative bequests, but assume here and in the next section that these constraints are not binding.\(^7\) In Section 4, where \( \gamma_i = 0 \), the non negativity constraints on \( b_{ij} \)’s (\( i = L, H \)) will necessarily be binding.

\(^6\)See e.g., Harsanyi (1995). By ‘launder’, we mean that in measuring social welfare the planner only takes into account the utility each individual derives from her own consumption, sometimes called felicity. That way, one avoids double counting and furthermore espousing individuals’ discriminatory preferences.

\(^7\)\( W_i \) and \( \gamma_i \) are supposed to be sufficiently large to yield interior solutions for bequests.
2.3. Tax function

We now return to the case of asymmetric information. Now, $T$ must be based on observable variables, namely wealth transfers. The tax function $T(b_1, b_2)$ specifies the (positive or negative) tax imposed on a parent (his estate) when bequest levels are given by $b_1$ and $b_2$. This formulation is quite general and includes a number of special cases. For instance, if $T$ depends on $b_1 + b_2$, the sum of bequests, one has a simple (linear or nonlinear) estate tax. A separable function $T = f(b_1) + f(b_2)$, on the other hand, represents a European style inheritance tax schedule (which, again, can be linear or non linear). When $T$ is not separable, the tax liability may depend on the way the estate is divided between children. Then, the tax function can be used to affect the sharing of bequests; it is a way to implement sharing rules. For instance strict equal sharing ($b_1 = b_2$) can be induced if $T$ is sufficiently large whenever $b_1 \neq b_2$.

With this tax function, the parent’s problem becomes

$$\max_{b_1, b_2} u(W_i - b_{i1} - b_{i2} - T(b_{1i}, b_{2i})) + \gamma_1 u(w_1 + b_{1i}) + \gamma_2 u(w_2 + b_{2i}).$$

Assuming a differentiable tax function, the first order conditions are now given by:

$$u'(y)(1 + T'_{ij}) = \gamma_j u'(c_{ij}) \quad i = L, H; \quad j = 1, 2,$$

where

$$T'_{ij} = \frac{\partial T(b_{1i}, b_{2i})}{\partial b_{ij}}.$$

Note that the bequests $b_{ij}$’s are after-tax (net) amounts. This is a pure accounting convention which has no implications on the results.

The values of $T'_{ij}$ determine the marginal rate of substitution (MRS) between parent’s and children’s consumption, which is given by:

$$\frac{u'(c_{ij})}{u'(y_i)} = \frac{1 + T'_{ij}}{\gamma_j}. \quad (6)$$

---

*The differentiability assumption is made here solely for the ease of exposition. The second-best solution presented below does not rely on this assumption; see footnote 13 for further discussion.*

*The marginal tax on gross bequests, denoted by $\tau$, is then given by

$$\tau = \frac{T'_{ii}}{1 + T'_{ii}}.$$

This is the usual formula to convert (marginal) rates from a tax-exclusive base to a tax-inclusive base.
Similarly, the ratio \( T_{i1}'/T_{i2}' \) determines the MRS between the two children’s consumption levels.

\[
\frac{u'(c_{i1})}{u'(c_{i2})} = \frac{\gamma_{i1} 1 + T_{i1}'}{\gamma_{i2} 1 + T_{i2}'}.
\]  

(7)

The distortions implied by a tax function can be assessed by comparing (6) and (7) with the corresponding MRS in the first-best (all equal to 1) and in the laissez-faire (equal to \( 1/\gamma_{i1} \) and \( \gamma_{i1}/\gamma_{i2} \) respectively).

In particular, comparing (6) to its laissez-faire counterpart one sees that \( T_{i1}' \) (resp. \( < \)) 0 implies less (resp. more) redistribution between parents and children than in a laissez-faire solution. Another relevant comparison is between the decentralized solution with taxation and the first-best solution. Observe that first-best levels of the MRS can be obtained if the marginal tax rates are set such that:

\[
\gamma_{i1} = 1 + T_{i1}'.
\]  

(8)

Consequently, \( 1 + T_{i1}' > \gamma_{i1} \), implies a downward distortion in \( b_{i0} \) compared to the first-best, while the opposite inequality implies an upward distortion.

### 3. Second-best problem

The second-best problem consists in designing the tax function \( T \) which maximizes welfare (taking into account its impact on the parent’s bequest behavior). As usual in optimal tax settings, we first characterize the optimal allocation, given the information available to the government, and then study its implementation via a tax function.

The information structure is formally introduced through the incentive (self-selection) constraints. In our setting, there are two types of incentive constraints. First, the policy must be designed in such a way that wealthy families cannot gain by mimicking poor families. The solution must therefore satisfy the following condition:

\[
\begin{align*}
 u(W_{L2} - b_{i1} - b_{i2} - T) + \gamma_{i1} u(w_1 + b_{i1}) + \gamma_{i2} u(w_2 + b_{i2}) &\geq 0, \\
 u(W_{L2} - b_{i2} + T) + \gamma_{i1} u(w_1 + b_{i1}) + \gamma_{i2} u(w_2 + b_{i2}) &\geq 0.
\end{align*}
\]

(9)

This constraint is necessarily violated at the full information optimum characterized by (3) and (4): all families leave the same bequests, but the rich pay a tax while the poor receive a transfer. Consequently, this constraint will be binding at

\[\text{In words, this requires that the net (after tax) cost of the transfer to child } j, (1 + T_{i0}'), \text{ equals the weight of this child in the parents utility.}\]
the second-best optimum (recall that the government has a utilitarian objective function) and it will restrict the redistributive capacity of the government.\footnote{The concavity of $u$ implies that a utilitarian government will always want to set $b_{i1} > b_{i2}$, so that (10) is necessarily satisfied if $\gamma_i \geq \gamma_2$.}\

In addition to this standard incentive constraint, there is a second type of constraints because the government does not observe a particular child’s income; it does know that there is one child of each type in every family, but (unlike the parents) it does not know which child is poor (unable) and which one is rich (able). Consequently, one has to ensure that parents are not going to cheat by ‘switching’ bequests between their children. This can be achieved by imposing the following constraints:

$$
\gamma_1 u(w_1 + b_{i1}) + \gamma_2 u(w_2 + b_{i2}) \geq \\
\gamma_1 u(w_1 + b_{i2}) + \gamma_2 u(w_2 + b_{i1}) \quad i = L, H.
$$

One can easily show that these constraints are never binding if $\gamma_1 \geq \gamma_2$.\footnote{We ignore the ‘upward’ incentive constraint preventing the poor from mimicking the rich. Because the downward incentive constraint is necessarily binding the upward constraint is relevant only in the case where both constraints bind simultaneously, which is a technical curiosity without much economic interest (and which can be ruled out at the expense of some technical assumptions).} They may become binding only if $\gamma_1$ is sufficiently smaller than $\gamma_2$, that is, when the conflict between the government’s and the parents’ preferences over $b_1$ and $b_2$ is sufficiently strong.

For the time being (and up to Section 4), we assume that differences in weight (if any) are not too significant, so that (10) is not binding. This allows us to focus on the social optimization problem with (9) as single constraint. Formally, we then maximize (1) subject to (9) with respect to $T$ and $b_i, \; i = L, H; \; j = 1, 2$. Differentiating with respect to $T, b_{i1},$ and $b_{i2}$ and rearranging yields the following first-order conditions.

$$
-u'(y_H) + u'(y_L) - \lambda[u'(y_H) + u'(\tilde{y}_H)] = 0, 
$$

$$
-u'(y_H) + u'(c_{H1}) + \lambda[-u'(y_H) + \gamma u'(c_{H1})] = 0, 
$$

$$
-u'(y_L) + u'(c_{L1}) + \lambda[u'(\tilde{y}_H) - \gamma u'(c_{L1})] = 0,
$$

where $\lambda > 0$ is the Lagrangemultiplier associated with constraint (9) and $\tilde{y}_H = W_H - b_{L1} - b_{L2} + T$ is the disposable income of the mimicker.

From the first-order conditions, one obtains the following expressions for the MRS for type $H$ and type $L$ families respectively:

$$
\frac{u'(c_{H1})}{u'(y_H)} = \frac{1 + \lambda}{1 + \lambda \gamma_1} \quad \text{and} \quad \frac{u'(c_{L1})}{u'(c_{L2})} = \frac{1 + \lambda \gamma_2}{1 + \lambda \gamma_1}
$$

$$
(14)
$$
\[
\frac{u'(c_{L1})}{u'(y_{L1})} = \frac{1 - \lambda R}{1 - \lambda \gamma_j} \quad \text{and} \quad \frac{u'(c_{L2})}{u'(y_{L2})} = \frac{1 - \lambda \gamma_j}{1 - \lambda \gamma_i}
\]

where \( R = \frac{u'(\tilde{y}_j)}{u'(y_{Lj})} < 1 \).

The results presented in the remainder of this section are derived from (14) and (15), while making use of (6) and (7) relating the MRS to the (marginal) rates implied by the implementing tax function.\(^{13}\) We consider alternative values of \( \gamma_j \).

To reduce the number of cases, we concentrate on situations where \( 0 \leq \gamma_1 \leq \gamma_2 \leq 1 \).\(^{14}\) In other words, the parent’s objective does not give more weight to his children than the social welfare function; further, when there is discrimination, it is against the lower ability child.\(^{15}\)

3.1. Case 1: \( \gamma_1 = \gamma_2 = 1 \)

In this case, all MRS in the \( H \) family are equal to one, that is their first-best (and \textit{laisser-faire}) levels. Consequently, one has \( T'_{H1} = T'_{H2} = 0 \); marginal tax rates are equal to zero. This is the counterpart in our setting of the standard ‘no distortion at the top’ result. No incentive constraint can be relaxed by distorting the choice of the \( H \) type parents. In addition, parents’ preferences are here perfectly in line with the government’s preferences so that ‘corrective’ measures (aimed at correcting the parents sharing of bequests between children) are not called for either.

For the \( L \) families, one obtains \( T'_{L1} = T'_{L2} > 0 \) and \( c_{Lj} \leq y_{Lj} \). There is a positive marginal tax on bequests so that intergenerational redistribution is lower than in the first-best. This result does not come as a surprise. When both weights are equal to one, there is no conflict between government’s and parents’ preferences for \textit{intra}-family redistribution. The formal structure of the model is then similar to that of the standard two group optimal income taxation model, where the positive marginal tax on the low type is a well known property. The intuition is also the same in both settings. The positive marginal tax and the so induced downward distortion in the individuals (observable) decision variable-bequests in our setting

\(^{13}\)It is well known from the optimal income tax literature that with two types of individuals, the implementing tax function will not in general be differentiable at one point. In our setting this problem occurs at \((b_{Lj}, b_{Lj})\). We nevertheless use the standard terminology and refer to \( \frac{u'(c_{Lj})}{\gamma_j u'(y_{Lj})} - 1 \), \( j = 1, 2 \), as the marginal tax rate for \( b_j \) of the low-wealth family; see Stiglitz (1982), p. 217 and Stiglitz (1987), p. 1003.

\(^{14}\)The first-order conditions do not rely on this assumption and the remaining cases can be dealt with as a straightforward extension of our analysis; see footnote 18.

\(^{15}\)It is the (potential) conflict of interest between government and parents and not the relative weights per se, which drives the result.
and pre-tax income in the standard model is beneficial because it relaxes an otherwise binding incentive constraint. This, in itself, hurts the poor, but it has an even stronger effect on the mimicking individual. Consequently, it opens the door for increased redistribution towards the poor, which will more than compensate them for the initial loss.

Finally, between children the first-best tradeoffs continue to prevail within any given family. Disposable incomes are equalized across siblings ($c_{H1} = c_{H2}$ and $c_{L1} = c_{L2}$) so that bequests provide full compensation for income (ability) differentials.

3.2. Case 2: $\gamma_1 = \gamma_2 < 1$

For the $H$ families, we now obtain the following inequalities:

$$\gamma_i - 1 < T_{m} < 0.$$  

Marginal tax rates are negative so that bequests are subsidized at the margin. However, they are larger than $\gamma_i - 1 < 0$ so that the (marginal) subsidization is not sufficient to achieve a first-best level of the MRS between parent and children (see (8) and the comments provided there). Put differently, there is less redistribution between generations than in the first-best but more than in the laissez-faire solution.

These results are rather intuitive. The government gives more weight to children’s utility than do their own parents. Consequently, it will tend to subsidize bequests to bring them closer to their first-best levels. However, under asymmetric information these ‘corrective’ subsidies do have a cost because they tend to affect the incentive constraint. The optimal (marginal) tax rates then strikes a balance between these two conflicting effects and, not surprisingly, the policy falls short of completely reestablishing a first-best level of inter-generational redistribution.

For the $L$ families, 11 implies that there can be over or under redistribution relative to the first-best (and thus negative or positive marginal tax rates) depending on the sign of $R - \gamma$. This ambiguity arises because the ‘corrective’ subsidy effect conflicts with the incentive effect. Like in the previous case, the relaxation of the incentive constraint calls for positive marginal tax rates. However, the parents now put lower weight on the children than the government, and this can be corrected by a subsidy on bequest (negative marginal tax rate). The

---

It can be easily verified that rich parents tend to prefer larger transfers.

The distortion will reduce utility of $H$ types parents, thus making it harder to satisfy the self-selection constraint.
optimal policy strikes a balance between these conflicting effects and, depending on their relative strength, positive or negative marginal taxes may be in order.\footnote{It can be observed in passing that when $\gamma_1 = \gamma_2 > 1$ – a case we do not consider in detail – marginal tax rates are necessarily positive. In that case the ‘corrective’ effect also calls for a positive tax (parents now put a higher weight on their children than the government) and thus reinforces the incentive effect.}

Redistribution between children of any given family is not affected by the asymmetry of information. Like in the first-best, consumption levels of siblings are equalized.

3.3. Case 3: $\gamma_1 < \gamma_2 < 1$

Let us finally consider the case where the parents put a higher weight on the utility of the able child ($\gamma_1 < \gamma_2 < 1$). First, one can easily derive the following inequality:

$$T_{H1}' < T_{H2}' < 0.$$  

Like in the previous case, bequests are thus subsidized at the margin. Further, one notices that the marginal subsidy is higher for the bequest towards the low ability child. Both results are in line with the corrective subsidy argument presented above. The crucial difference with the previous case is that (unlike the government) parents now put different weights on their children, and this is potentially the source of a second type of corrective subsidy (aimed at altering the sharing of bequests between children).

This is shown by the following expression which compares the MRS between children obtained at the different types of solutions:\footnote{To simplify the interpretation, it also recalls (7) specifying the relationship between MRS and marginal tax rates.}

$$1 < \frac{1 + \lambda \gamma_2}{1 + \lambda \gamma_1} \left( \frac{\gamma_2}{\gamma_1} \frac{1 + T_{H1}'}{1 + T_{H2}'} \right) < \frac{\gamma_2}{\gamma_1}.$$  

(16)

These relationships imply that the (between children) MRS in the second-best is larger than 1, the MRS in the first-best (so that $c_{H1} < c_{H2}$), and smaller than $\gamma_2/\gamma_1$, the MRS in the *laisser-faire*. Once again, the second-best allocation strikes a compromise between the *laisser-faire* and the first-best. The tax policy provides partial correction for the parents’ bias in favor of child 2. However, for incentive reasons it falls short of completely eliminating this bias.

The following (admittedly somewhat loose) argument is helpful for achieving a more precise understanding of the impact of informational asymmetries. Observe
that the incentive constraint affects (16) through its shadow price $\lambda$. If $\lambda$ is close to 0 (that is, informational considerations have only an insignificant impact on welfare), the second-best allocation is close to the first-best one. If instead $\lambda$ is very large, it would be close to the laissez-faire allocation.

Turning to the $L$ families, the following expression provides the respective values of the MRS between children’s consumption (that is of $u'(c_{L1})/u'(c_{L2})$), for the second-best, the first-best and the laissez-faire solutions:

$$\frac{1 - \lambda \gamma_2}{1 - \lambda \gamma_1} \left( = \frac{\gamma_2(1 + T'_{L1})}{\gamma_1(1 + T'_{L2})} \right) < 1 < \frac{\gamma_2}{\gamma_1}. \quad (17)$$

This expression is in sharp contrast with (16), its counterpart obtained for $H$ type families. It implies that poor (unable) children will have a higher after-transfer income than rich (able) children ($c_{L1} > c_{L2}$). In other words, the induced transfer scheme reverses the initial ranking. This is even more noticeable as parents put a higher weight on their rich child. Once again, both corrective and incentive effects are at work, but it is clear that the ‘overshooting’ (excess redistribution between children) can only be explained by incentive considerations. This intuition can be confirmed by a formal inspection of the incentive constraint 5 which shows that a $1$ increase in $b_{L1}$ is more effective in relaxing this constraint than the same increase in $b_{L2}$.

As it can be expected from this argument, (17) implies $T'_{L1} < T'_{L2}$. The sign of these marginal tax rates is, however, ambiguous. This does not come as a surprise and it is perfectly in line with the intuition and results presented in the previous sub-section (case of $\gamma_1 = \gamma_2 < 1$).

A summary of the formal results obtained in the different cases for marginal rates of substitution and marginal tax rates is provided in Table 1.

3.4. Type of taxation and relevance of equal sharing

To conclude the section, let us also summarize the results in words and discuss their implication as far as type of taxation (estate vs. bequest) and relevance of sharing rules are concerned. First, when $\gamma_1 = \gamma_2 = 1$, the marginal tax is 0 for a type $H$’s family and positive for a type $L$’s family. Because the same marginal tax rate is applied to transfers to children of both types the optimal tax can be applied to the overall estate (the sum of bequests); consequently, a non-linear (progressive) estate taxation suffices. Second, when $\gamma_1 = \gamma_2 < 1$, bequest is marginally subsidized for type $H$’s families; it is not clear whether it is subsidized or taxed for type $L$’s families. Once again, estate taxation is sufficient to implement the solution.

---

5 Recall that the transfer scheme preserves the ranking between $H$ children.
Table 1
Summary of the results

<table>
<thead>
<tr>
<th>Across generations</th>
<th>Between children</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u'(c_{i1})/iu'(y)$</td>
<td>$u'(c_{i2})/iu'(y_{i})$</td>
</tr>
</tbody>
</table>

(a) Marginal rates of substitution ($y_1 \geq y_2 \geq 1$)

<table>
<thead>
<tr>
<th>Type</th>
<th>First-best</th>
<th>Second-best</th>
</tr>
</thead>
<tbody>
<tr>
<td>Laisser-faire</td>
<td>$\frac{1}{y_1} \geq 1$</td>
<td>$\frac{1}{y_1} \geq 1$</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>$1 + \lambda y_1$</td>
</tr>
<tr>
<td></td>
<td>$1 + \lambda y_1 \leq 1$ for $H$;</td>
<td>$1 + \lambda y_1 &gt; 1$ for $H$;</td>
</tr>
<tr>
<td></td>
<td>$\frac{1 - \lambda y_1}{1 - \lambda y_1} \geq 1$ for $L$.</td>
<td>$\frac{1 - \lambda y_1}{1 - \lambda y_1} &lt; 1$ for $L$.</td>
</tr>
</tbody>
</table>

(b) Marginal tax rates

| $\gamma_1 = \gamma_2 = 1$ | $T_{i1} = T_{i2} = 0$ | $T_{i1} = T_{i2} = 0$ |
| $\gamma_1 < y_1 < 1$ | $\gamma_1 / y_1 < T_{i1} = T_{i2} < 0$ | $T_{i1} = T_{i2} \geq 0$ |
| $\gamma_1 < y_1 < 1$ | $T_{i1} < T_{i2} < 0$ | $T_{i1} < T_{i2}$; $T_{i3} \geq 0$ |

Third and last, when $\gamma_1 < y_2 < 1$, taxation favors low ability children even though they receive less weight in their parents preferences. In this case, an estate tax is not sufficient; the tax schedule must depend on the individual values of $b$ to allow for a differentiation of tax rates between the children of a same parent. In general, a non-separable tax function depending on the individual bequests is then called for.\textsuperscript{21}

Finally, it can be observed that equal sharing does not arise in any of the considered cases. In the first two cases, the policy does not interfere with the parent’s decision on how to share their estate between children. In the third case, there are ‘corrective’ measures in that the implementing tax function implies different (marginal) tax rates for transfers towards the different children. In either of the cases, a strict equal sharing rule would thus decrease welfare.\textsuperscript{22}

---

\textsuperscript{21}Within the (admittedly restrictive) two-types setting considered here a separable function (i.e., an inheritance tax system) may in fact be sufficient to implement the optimal solution. This is true, for instance, if the largest bequest of a $L$ parent is smaller than the smallest bequest of a $H$ parent.

\textsuperscript{22}Our solution is of course not incompatible with alternative, less stringent sharing rules. The second-best problem yields an (constrained) Pareto-efficient allocation which can be implemented in a variety of ways. A non-linear tax function is one option on which we have concentrated. Alternatively, a combination of taxes and direct regulations can achieve the same outcome. In that sense, one can argue that appropriately designed sharing rules are not harmful; however, they do not add anything to the sole instrument of taxation.
4. The Cinderella effect: Equal sharing.

So far, we have ignored the incentive constraints (10), ensuring that parents do not find it beneficial to switch bequests between their children. In most of the considered cases, this is of no relevance; recall that when \( \gamma_1 \geq \gamma_2 \), condition (10) is automatically satisfied. However, when \( \gamma_1 < \gamma_2 < 1 \), this is no longer necessarily true. In that case there is a potential conflict between the government and the parents as far as the targeting of bequests is concerned. This conflict appears most strikingly for \( L \) type families (see Section 3.3) where the optimal solution implies that the bias in favor of the more productive child is to be reversed. One can then expect that when \( \gamma_1 \) is sufficiently low \( \delta \) does become binding and this is likely to have a significant impact on the solution pattern. In particular, one can ask whether this can lead to equal sharing.

A detailed study of all the possible cases that can arise when (10) is binding is a tedious exercise which would go beyond the scope of the current paper. We shall therefore concentrate on the specific (and from our perspective most interesting) issue of equal sharing and illustrate its emergence by considering a somewhat extreme case.

Let us thus adopt the drastic assumption of zero altruism toward the lower ability child, that is \( \gamma_1 = 0 \). G. Brenner (1985b) explains `... the restrictive, egalitarian inheritance customs during the English Middle Ages (....) by the mistreatment of both stepchildren by step-parents and younger children by their older brothers and sisters.' (p. 96). The prime example of this situation is the Cinderella story. When both the demographic and the societal evolution made such situation less likely, England moved towards unrestricted bequeathing whereas most continental Europe maintained restrictive equal estate-sharing. Why such a contrasting evolution? To answer that question is outside the realm of this paper. For our purpose, it suffices to show the link between mandatory equal sharing and the lack of altruism towards some children.

Assuming non negative bequests, the \textit{laisser-faire} solution is now \( c_{i1} = w_i \) and \( c_{i2} = y_i \) if \( \gamma_2 = 1 \). The first-best solution is unchanged. One easily checks that with \( \gamma_1 = 0 \), the constraints (10) are equivalent to \( b_{H2} \geq b_{H1} \) and \( b_{H2} \geq b_{H1} \). In words, there is no possibility to implement a solution in which the low-ability child receives a strictly larger bequest than the high-ability child. With these added constraints, the second-best problem can then be rewritten as the maximization of

---

\footnote{For \( H \) types parents one might be tempted to think that such a problem cannot arise. This is because the optimal solution respects the parents ranking in that case (see Section 3.3). One has to keep, however, in mind that throughout Section 3 we have also assumed that the non negativity constraint on bequest is not binding. Now, when \( \gamma_1 \) is sufficiently small, the solution may no longer be interior, and our results do not necessarily apply. In that case, (10) may well be binding for \( H \) type families too; our discussion below (based on \( \gamma_1 = 0 \)) provides an illustration of such a possibility.}

\footnote{Recall that \( y_i \) denotes parents` consumption.}
(1) with respect to $T$ and $b_{ij}, i = L, H; j = 1, 2$ subject to the standard incentive constraint (9) and the sharing constraints $b_{L2} \geq b_{L1}$ and $b_{H2} \geq b_{H1}$.

It follows directly from the results in Section 3.3 that (with $\gamma_i = 0$) the sharing constraint for $L$ type parents is necessarily binding. As to the $H$ type family, it may or may not be binding.

When the sharing constraints are taken into account, implementation of the solution then requires equal sharing rules. Specifically, two types of situations can arise. Either ‘strict’ equal sharing has to be imposed for all bequest levels (in the case where the sharing constraint binds in $H$ type families), or (otherwise) ‘limited’ equal sharing applies. In the latter case, ‘small’ estates are subject to an equal sharing constraint, while (a limited extent of) unequal sharing is tolerated for ‘large’ bequests.

5. Conclusion

In this paper, we have studied the optimal design of a tax policy towards bequests. Bequests are motivated by parental altruism, which can vary between children and the tax policy is determined by a utilitarian social planner. The optimal tax system so derived is a compromise between intergenerational equity within each family and intragenerational equity across families of different wealth.

When planner and parents weight the weight the children in the same way, an estate tax, that is a tax based on aggregate bequests suffices. When they adopt different weights, then one needs to use a progressive tax formula that depends on individual bequests. Finally, when there is a possibility of the parent disinheriting their less productive child, the government may find it optimal to impose a tax schedule which implies equal sharing along with a progressive estate tax.

Our basic model assumes that parents’ income is exogenous and unobservable so that an income tax is not available. The variant presented in the appendix considers an alternative information structure under which income is endogenous and observable; a (general) income tax is then also available, but restricted by the unobservability of earning ability (and labor supply). We show that the original problem and the variant have a similar structure and yield the same properties for

---

25Recall that the solution in Section 3.3 implies $b_{L2} > b_{L1}$ thus violating the constraint.

26We have seen in Section 3.3 that the solution corrects the parent’s bias without reversing it. It can be easily shown that, depending on the parameter values, one can then have either $b_{H2} > b_{H1}$ or $b_{H2} < b_{H1}$.

27Mathematically speaking, a non-linear tax function can still do the job. However, it would imply infinite marginal tax rates for unequal bequests, a solution which appears hard to accept on economic grounds.

28In either case, there is no equalization of consumption levels between parent’s and children. In other words, tradeoffs between generation are distorted compared to the fist best solution.
the bequest tax (same type of distortions and same signs for the marginal bequest tax rates). Consequently, the bequest tax has to be part of the redistributive tax arsenal, at least as long as the income tax is also restricted by informational asymmetries.

How do these findings fit the reality of bequeathing behavior and that of existing legal and fiscal institutions? Starting with the institutions, one can say that the estate tax with unrestricted bequests such as prevailing in the US and the UK is consistent with our result when parents weight their children’s utility equally. As to the Continental regime of inheritance tax and mandatory equal estate division, it corresponds to the admittedly particular Cinderella effect case. As to actual behavior, one observes two types of convergence in advanced countries. First, regardless of the regime prevailing, the return of inheritance or estate taxation is very low (less than 1/2 percent of total revenue) in spite of high statutory taxes. In other words, tax reporting and compliance are lacking. This explains why British economists use to call estate taxation a tax on sudden death. Our paper does not deal with the issues of evasion and avoidance.

The other convergence is towards what we would call imperfect equal sharing. In the USA, where estate division is unrestricted, converging evidence indicates equal sharing in about 60% of cases. In France, in about 8% of cases, estates are unequally divided but the increasing importance of inter vivos gifts makes it possible to bypass the equal division rule.

These real life features should be taken into account in future research. In particular, in our paper, we assume perfect observability of bequests and we do not consider the possibility of unplanned bequests which could make the equal division case stronger. Furthermore, even within our model of planned bequests, equal division can result if children are not different, which is often the case.

Acknowledgements

We would like to thank Jim Davies, Lisa Grazzini, Patrick Legros, P.L. Sacco, and the participants of the International Seminar on Public Economics on Bequest and Wealth taxation, Liège, May 18–20, 1998, for their observations and suggestions. We also very grateful to the referees and the editor, Jim Poterba, for their insightful and constructive comments.

\[\text{29}^\text{This is the most widely quoted figure. In fact, the percentage ranges from 22\% in Tomes (1981) to 87\% in Menchick (1988).}\]

\[\text{30}^\text{For a survey of the evidence on estate sharing in the US and in France, see Arrondel et al. (1997).}\]
Appendix A

Variant: bequest and income taxation

Let us now briefly sketch a variant of the model, characterized by a different information structure, which allows us to consider income taxation as an additional instrument.

Assume that parents differ in earning abilities \(v_i\) with \(v_H > v_L\). Let \(l_i\) denote labor supply and \(I_i = \omega l_i\) (before tax) income. In the absence of taxation, the utility of parent \(i = H, L\) is given by:

\[
u_l^i = \frac{I_i}{v_i} - b_{i1} - b_{i2}\]

where \(h(l_i/\omega_i) = h(l^i)\) is the monetary disutility of labor. Following the optimal income tax literature, we assume that \(I_i\) is observable while \(v_i\) and \(l_i\) are private information; wealth transfers, \(b_{Hj}\) and \(b_{Lj}\), on the other hand continue to be observable (subject to the same restrictions as before).

The optimal allocation is obtained by maximizing welfare, redefined as:

\[
W = u\left(I_H - h\left(\frac{l_H}{\omega_H}\right) - b_{H1} - b_{H2} - T\right) + u(I_H) + u(w_2 + b_{H2})
\]

\[
+ u\left(I_L - h\left(\frac{l_L}{\omega_L}\right) - b_{L1} - b_{L2} + T\right) + u(I_L) + u(w_2 + b_{L2}).
\]

(A.2)

with respect to \(I_i, b_{ij}\) and \(T\). Observe that (A.2) replaces (1) from which it differs only in the terms pertaining to the parents’ (own) utility. Furthermore, we have two additional instruments, namely \(I_H\) and \(I_L\).

The downward self-selection constraint (i.e. the counterpart to (9)) is now given by:

\[
\begin{align*}
u_l^i & = \frac{I_i}{v_i} - b_{i1} - b_{i2} - T + u\left(I_i - h\left(\frac{l_i}{\omega_i}\right) - b_{i1} - b_{i2} + T\right) + u(I_i) + u(w_2 + b_{i2}) \\
\geq & u\left(I_i - h\left(\frac{l_i}{\omega_i}\right) - b_{i1} - b_{i2} + T\right) + u(I_i) + u(w_2 + b_{i2}).
\end{align*}
\]

(A.3)

Like in Section 3 we ignore (10) (which remains unchanged).

Let us denote the solution to this problem by \(I_{H}^*, b_{Hj}^*\) and \(T^*\). Observe that \(b_{Hj}^*\) and \(T^*\) solve the first-order conditions to this problem given \(I_{Hj}^*\) and \(I_{Lj}^*\). We can then define:
Substituting (A.4) and (A.5) into (A.1) and (A.3) yields the reformulated problem (for the determination of \( b^* \) and \( T^* \), which is exactly identical to the problem studied in Section 3 (and specified by (1) and (9) except that \( W_H \) in the right-hand-side of (9) has to be replaced by
\[ \tilde{W}_H = I_H^* - h \left( \frac{I_H^*}{\omega_H} \right), \]
but this has no impact on the results.\(^{31}\)

To sum up, while the two problems are not perfectly equivalent, it can be easily checked that all the properties of the marginal taxes on bequests (and hence the characterizations of the distortions) derived in Section 3 also obtain for the modified problem. The are thus robust to the introduction of a (general) income tax on the parents’ income.\(^{32}\)

References


\(^{31}\)And like in the original problem \( W_s \) are not directly observable (because they depend on \( \omega \)).

\(^{32}\)The separation of \( T \) into income and bequest taxes is essentially arbitrary here (except that the marginal properties must be satisfied so that we effectively need both instruments). This is because we have a single dimension of adverse selection, like in a standard optimal income tax problem, but the space of observables is of larger dimension here. To obtain a well-defined implementation problem, one would have to consider an additional dimension of adverse selection, allowing for instance parents to differ in ability and altruism (weights attached to children). While this is an interesting extension, it raises technical problems which go beyond the scope of the current paper (see Cremer et al. (2000) for the analysis of a two-dimensional optimal tax problem).