Tax Treatment of Bequests when Donor Benefits do not Count

by

Robin Boadway, Queen's University[†]
<boadwayr@econ.queensu.ca>
and
Katherine Cuff, McMaster University
<cuffk@mcmaster.ca>

25 February 2014

Abstract: Recent literature on bequest taxation has made a case for subsidizing bequests on efficiency grounds, with the rate of subsidy declining with the size of the bequest to equalize opportunities among inheritors. These results rely heavily on the fact that bequests benefit both the donor and the recipient, creating an externality that is not taken into account by donors. The double-counting of the benefits of bequests is questionable on normative grounds. We study the consequences for bequest taxation of not counting the benefit to donors. We analyze a simple model of parents and children with different skills where parents differ in their preferences for bequests. The government implements nonlinear income taxes on both parents and children as well as a linear bequest tax, and gives differential linear bequest tax credits to donor parents. The nonlinear income taxes take standard forms. The bequest tax and tax credits are of indeterminate absolute level. Bequest tax credit rates are positive and may be higher for high-wage parents.

Key Words: bequest tax, inheritance, optimal income tax

JEL Classification: H21, H23, H24

[†] **Acknowledgments:** Earlier versions of this paper were presented at the Fiscal Studies Workshop on Taxation of Inheritance and Wealth at Oslo University, the International Workshop on Indirect Taxation at Sciences Po, Paris, and the Workshop on Public Economics at Erasmus University Rotterdam. Comments of participant are appreciated. Financial support from the Social Sciences and Humanities Research Council of Canada and the Canada Research Chair Programme are gratefully acknowledged.

1 Introduction

Bequests represent both a consequence and a determinant of economic inequality. As such, the taxation of bequests and inheritances is a potentially important component of redistribution policy. Despite the fact that revenues from bequest taxation constitute a small or non-existent proportion of total tax revenues in most countries, optimal bequest taxation is a lively and contentious area of research and policy consideration. For example, the Mirrlees Review (2011), echoing the Meade Report (1978), proposed a cumulative lifetime tax on inheritances as a way to mitigate large differences in economic opportunities at birth. And, there has been a resurgence of interest in the optimal taxation of wealth transfers, recently surveyed in Cremer and Pestieau (2011) and Kopczuk (2013).

The optimal tax treatment of bequests remains contentious, and depends on the bequest motive, the responsiveness of bequests to taxation and the normative underpinnings for bequest taxation. On the one hand, bequests can be unintentional or accidental, resulting from wealth accumulated for lifetime purposes, and left unspent at death. The purpose can be to self-insure against an uncertain length of life or uncertain health or care expenses, or wealth may be accumulated solely as an end in itself. In either case, taxing an estate consisting of wealth that has simply been left over has no disincentive effect and could be taxed at a very high rate. On the other hand, bequests may be intentional. They may reflect transfers purposely given to one's heirs (or to charity) either for altruistic motives or from some satisfaction from giving, typically called the joy-of-giving but possibly including morally felt obligations. In either case, taxing them presumably has some discouraging effect. Intentional bequests may also take the form of payment to one's heirs for services, so-called strategic bequests. These are in principle no different from any other market transaction and could treated as such for taxation purposes.

The normative basis for the taxation of bequests is controversial. From the point of view of recipients, it seems clear that inheritances represent windfall sources of income that not only provide benefits to them, but do so in a very unequal way. Taxation serves both an equalizing objective and contributes to equality of opportunity. How to treat the donors is another matter, at least if the bequests are intentional. Those who, like Kaplow (1998, 2001, 2008), adopt a strict welfarist stance argue on revealed preference grounds that donors must have received a benefit from giving the bequest, and that benefit should 'count' in the social welfare criterion used by the government. Others, like Hammond (1987) and Mirrlees (2007, 2011), argue that this gives rise to double-counting. Since the benefit to the recipients from bequests is already counted in social welfare, it is unnecessary to include them again in the guise of benefits to the donors.

The recent literature on the optimal taxation of bequests has tended to take the welfarist position and double-count the benefits of bequests (Farhi and Werning 2010, 2013; Brunner and Pech 2012; Kopczuk 2012, 2013; Piketty and Saez 2012). The typical findings are fairly intuitive. Suppose individuals differ in wage rates, and the government redistributes using an optimal income tax. The question is what tax treatment of bequests and inheritances should accompany the income tax. From the point of view of the donor, a bequest is comparable to the purchase of a good. If preferences are weakly separable and if the government implements a optimal nonlinear tax, the Atkinson and Stiglitz (1976) theorem applies. Differential taxation of bequests should not be applied as a device for improving the redistributive capacity of the income tax. (Of course, if the tax mix consists of a uniform commodity tax plus a nonlinear income tax, the commodity tax should apply to bequests given.)

This simple outcome is complicated when the benefit of inheritances to recipients is taken into consideration. For one thing, the value of inheritances to the recipients is a social benefit that is not taken into account by the donors, who attach weight only to the benefits the donations give to themselves. There is thus an externality that calls for a subsidy on bequests. Second, the receipt of inheritances of different sizes yields unequal benefits to the recipients, which can be addressed by making the tax on inheritances progressive. Finally, to the extent that inheritances are correlated with the wage rate of recipients, that constitutes a further argument for taxing them as an indirect way of taxing high-wage persons. The first two influences lead Farhi and Werning (2010) to the finding that bequest taxes should be negative, but progressive. The third increases the optimal tax on bequests, possibly enough to make it positive (Brunner and Pech 2012a). Kopczuk (2013) further argues that if inheritances reduce the labor supply of recipients, this reduces social welfare as long as labor income taxes are positive. This leads to another argument for taxing bequests, making it ambiguous whether bequests should be taxed or subsidized.

These results rely on the double-counting of the benefits of bequests. Although the welfarist arguments leading to double-counting are on the surface persuasive, especially the revealed-preference argument, there are compelling arguments against double-counting. The benefit one obtains from the utility of one's heirs should apply in principle to any form of interdependent utility whether revealed through transfers or not. In a family, presumably each member values the well-being of other members, but there is no suggestion that this multiplicity of utilities be counted in social welfare. The same applies more generally to feelings of altruism or even avarice toward fellow citizens.[1] Some have noted the analogy

[1] There has been some analysis of the consequences of interdependent utilities for optimal tax-transfer policy. Examples include Archibald and Donaldson (1976), Oswald (1983) and Frank (2008).

between saving for one's own future and saving for heirs. As Mirrlees (2007) argues, we would not consider counting the utility we obtain now when saving for our future self's consumption. Some might regard the role of government redistribution as reflecting the altruism of the rich for the poor, and internalizing the free-riding from private donations.[2] There is no suggestion that in this case the rich taxpayers' altruism should be counted as social welfare. Finally, intentional bequests may not give utility to donors at all. They may represent voluntary transfers done out of a sense of obligation, making the donors worse off.

The purpose of this paper is to explore the consequences for optimal bequest taxation of neglecting the benefits of bequests to the donors. In the literature, this has been done by analyzing the tax treatment of bequests when the benefits of donors are excluded from social welfare, or 'laundered out' to use the terminology of Cremer and Pestieau (2006, 2011). The literature has focused on the implications for bequest taxation, and has not studied how the tax treatment of donors is affected. This is important because, as Kopczuk (2013) emphasizes, bequest behavior is very heterogeneous. Failure to take account of that heterogeneity entails discriminating in favor of non-donors, who are not forgoing any consumption by leaving bequests, relative to donors (McCaffery 1994; Mankiw 2006). Models used in the existing literature to study the effect of laundering out on bequest taxation typically assume that all households have the same preferences for bequests (Cremer and Pestieau 2006, 2011; Mirrlees, 2007; Brunner and Pech 2012b). The treatment of donors relative to non-donors then does not arise.

In this paper, we study the implications of not counting the benefits of bequests to donors in a simple model with three sources of heterogeneity. Individuals differ in their wage rates, in their preferences for bequests and in the inheritances they receive. We study optimal nonlinear income taxation, the income tax treatment of bequests, and the taxation of inheritances. Our model closely follows the recent literature by adopting some important simplifications to facilitate the analysis and the interpretation of results. As in Farhi and Werning (2010), Brunner and Pech (2012a) and Kopczuk (2013), we focus on two generations, parents and children. Each generation has an exogenous distribution of wage rates, and there may be some correlation between parental and child wage rates. Social welfare includes the sum of social utilities of parents and children, where the social utility of parents excludes the utility they obtain from bequests. To be able to separate utility of bequests from utility of own consumption, we assume an additive form of utility function. Heterogeneity of bequest behavior is captured by assuming that some parents

[2] Classic references to redistribution as altruism of the rich for the poor may be found in Hochman and Rodgers (1969), Thurow (1971) and Pauly (1973).

give bequests, while the rest do not.

The government has three policy instruments. It imposes a nonlinear income tax on both parents and children; it chooses a linear tax on inheritances; and, it allows an income tax credit on bequests. The latter reflects the fact that the government does not count the utility of bequests to donors. Instead, bequests simply reduce donors' consumption. A bequest tax credit allows non-donors and donors to be treated relatively comparably by the income tax system. We assume that, apart from the bequest tax credit, the income tax is not conditioned on bequests. Also, the inheritance tax is not related to either parental or child incomes. These are obviously strong assumptions, albeit similar to those used in Farhi and Werning (2010), Brunner and Pech (2012a) and Piketty and Saez (2012), but they simplify the analysis considerably.

We begin by outlining the decision problems of the parents and children and then characterize the government's problem in Section 3. In Section 4, we focus on the tax treatment of parents under two different informational environments, emphasizing the differences between donor and non-donors by assuming the government puts zero social weight on the welfare of children. In Section 6, we incorporate the behavior and taxation of children in the analysis by assuming the government puts positive social weight on the children's utility. Section 7 concludes.

2 Household Problem

The simplest model is chosen to illustrate the effects. There are two wage-types of parents, n_1 persons with wage w_1 and n_2 with wage w_2, where $w_2 > w_1$ and $n_1 + n_2 = 1$. A proportion d_i of type$-i$'s are donors, so $1 - d_i$ are non-donors, where we suppose $d_2 > d_1$ for concreteness. Donor and non-donors differ in their preferences for bequests, for example, because of their views of the role of government in making transfers.

Following Farhi and Werning (2010), we assume each parent has one child. Each child is endowed with either high or low skills, but not necessarily the same as his parent. In particular, the probability that a child is the same skill type as his parent is given by $\pi \in (1/2, 1]$, where we assume that the same probability π applies to both high- and low-skilled parents. Thus, the skill-type of a child is positively correlated with the parent's type. All children with parents of a given skill-type are equally likely to have a donor parent. Therefore, the probability of having a donor parent is simply given by the proportion of donor parents of a given skill-type, d_i. For children of non-donor parents, the skill-type of their parents is irrelevant. Consequently, there will be six distinct child-types: each of the two skill-types of children can receive bequests of b^1, b^2 or zero. We can denote a bequest-receiving child's type by the characteristics (w_k, b^i) for $k, i = 1, 2$, where k is the child's skill-type and i is a donor parent's skill-type. To simplify matters, we focus on only two generations, and assume that the children do not make bequests.[3] We begin by characterizing the donor and non-donor parents' problems and then turn to the children's problem.

2.1 Parents' Behavior

Household utility depends on a private consumption good x, on labor supply ℓ, and in the case of donors, on net-of-tax bequests b, where we drop subscripts for the time being and whenever it causes no confusion. Suppose the inheritance tax is proportional at the rate t. Then, net bequests can be written $b = (1-t)B$, where B is the actual (gross) bequest. Preferences of donors and non-donors take the following quasi-linear forms, respectively:

$$U(x, b, \ell) = x + f(b) - h(\ell), \quad u(x, \ell) = x - h(\ell) \tag{1}$$

where $f(b)$ is the utility of bequests (joy-of-giving) applying to donors, and $h(\ell)$ is the disutility of labor supply. The latter satisfies $h'(\ell) > 0$ and $h''(\ell) > 0$. We assume that the utility-of-bequest function satisfies $f''(b) < 0$, with $f'''(b) \leqslant 0$ and $f'(\bar{b}) = 0$ for some $\bar{b} > 0$. That is, there is some maximal value of net bequests b that contributes to the donor's utility: the $f(b)$ curve is an inverted U-shape with a peak at $b = \bar{b}$. We also assume that $f'(0) > 1$. This ensures that in the laissez-faire, where the price of bequests is unity, the parents will make positive bequests. An example of a functional form satisfying these properties is the quadratic form $f(b) = 2\bar{b}b - b^2$ with $\bar{b} > 1/2$, where $f'(b) = 2(\bar{b} - b)$ so $f'(\bar{b}) = 0$, $f'(0) > 1$, and $f''(b) < 0$. For future reference, define the elasticity of the marginal utility of bequests as follows:

$$\varepsilon_b \equiv -\frac{bf''(b)}{f'(b)} > 0. \tag{2}$$

Our assumption that $f'''(b) \leqslant 0$ ensures that ε_b is increasing in b.

Three points should be noted about the donor's utility function in (1). First, the additively separable form of $U(x, b, \ell)$ allows the government to disregard the utility of bequests to donors by simply not counting $f(b)$. If the utility function took a general

[3] Brunner and Pech (2012a) allow parents to receive exogenous bequests from their parents, whose behavior is not explicitly modeled. Adding this would entail adding a bequest tax and credit system applying to parental inheritances, which would complicate matters without adding a great deal of insight. We could also allow the children to leave bequests, which would add another level of bequest tax and credit policies.

non-separable form, it would not be obvious how to specify the net-of-bequests utility of the donors. Second, we assume that donors get utility from the after-tax bequest b. That is, they care about what ends up in the hand of their heirs, rather than the amount they leave before-tax, which is B. This implies that a tax levied on inheritances will affect their behavior, which would not be the case if B were the argument of $f(\cdot)$. Finally, since x absorbs all income effects, we should include a non-negativity constraint on x. In what follows, we assume it is never binding, so that choices of b and ℓ are always in the interior.

A households whose wage rate is w earns $y = w\ell$, and pays income tax $T(y)$. In addition, donors obtain a tax credit at the rate τ on their gross bequests. The government can observe a household's income y and its bequests B, in which case it can condition the tax credit τ on the level of bequests as well as on income. For simplicity, we assume that the tax credit rate is proportional, but may differ by income. That is, τ can take values τ_1 and τ_2 for the two wage-types. The budget constraint for a donor is

$$x + B = y - T(y) + \tau B \equiv c + \tau B$$

or, using $b = (1-t)B$,

$$x + \frac{b}{1-t} = c + \tau \frac{b}{1-t} \tag{3}$$

where c is after-tax income, or disposable income, before the tax credit. Note that the income tax function $T(y)$ does not depend on bequests, although the bequest tax credit does. This simplifies the analysis, although it may be restrictive. Note also that the budget constraint (3) assumes that the inheritance tax is paid by the recipient.

It is useful to disaggregate the donor's problem artificially into two stages. In the first stage, labor is chosen, which determines income $y = w\ell$ and disposable income $c = y - T(y)$. In the second stage, the donor chooses how to divide after-tax income between x and B, or $b/(1-t)$. Begin with the second stage.

Stage 2: Donor's allocation of disposable income

Given y and c, and using the utility function (1) and the budget constraint (3), the problem of donor household with wage rate w can be written as follows, where we adopt b as the donor's choice variable:

$$\max_{\{b\}} \; c - \frac{1-\tau}{1-t} b + f(b) - h\left(\frac{y}{w}\right).$$

The necessary condition for this problem is

$$f'(b) = \frac{1-\tau}{1-t} \tag{4}$$

whose solution yields the demand function for net bequests

$$b\Big(\frac{1-\tau}{1-t}\Big), \quad \text{with} \quad b'(\cdot) = \frac{1}{f''(b)} < 0. \tag{5}$$

Comparative statics of (4) yields:[4]

$$b_\tau = \frac{-1}{1-t}b'(\cdot) > 0, \quad b_t = \frac{1-\tau}{(1-t)^2}b'(\cdot) < 0. \tag{6}$$

where subscripts refer to partial derivatives.

It will be useful in what follows to define the argument of (5) for a donor as follows:

$$p \equiv \frac{1-\tau}{1-t}. \tag{7}$$

This can be interpreted as the effective price of net bequests to the donor. We can define the price elasticity of net bequests as

$$\varepsilon_p \equiv -\frac{pb'(p)}{b(p)} = \frac{1}{\varepsilon_b} > 0$$

where the last equality follows from (2), (4) and (5). If $\varepsilon_p > 1$, the demand for net bequests is relatively elastic and vice versa. Differentiating ε_p with respect to p yields:

$$\frac{d\varepsilon_p}{dp} = -\frac{b'(p)}{b(p)} - \frac{pb''(p)}{b(p)} + \frac{p(b'(p))^2}{(b(p))^2} > 0.$$

Note that if $\tau = 1$ so $p = 0$, we have $b(0) = \bar{b}$. Also, if $t = 1$, then $b = (1-t)B = 0$ for any B, so the utility of bequests is $f(0) = 0$. The donor faces the cost $(1-\tau)B$ of leaving a bequest, so none will be left, and we would be in a no-bequest world. Note, however, that if gross bequests are fixed so the donor does not make any bequest decision, then bequests can be fully taxed without consequences. Net bequests, of course, will then be zero.

---

[4] Using (2), (4) and (6), the elasticity of net bequests with respect to the bequest tax credit and the inheritance tax can be written respectively as

$$\frac{\tau b_\tau}{b} = \frac{\tau}{1-\tau}\frac{1}{\varepsilon_b} > 0 \quad \text{and} \quad \frac{tb_t}{b} = -\frac{t}{1-t}\frac{1}{\varepsilon_b} < 0.$$

Stage 1: Choice of labor supply

When households choose labor supply, they anticipate the choice of net bequests b in the following stage. Given a tax function $T(y)$, the type$-w$ households solve the following problem:

$$\max_{\{y\}} \; c - pb(p) + f(b(p)) - h\Big(\frac{y}{w}\Big) \quad \text{s.t.} \quad c = y - T(y).$$

The solution to this problem gives

$$h'\Big(\frac{y}{w}\Big) = (1 - T'(y))w \tag{8}$$

which determines y. Thus, income or labor supply is not affected by the inheritance tax and bequest tax credit. The same result (8) on income applies to both donor and non-donor parents.

In what follows, we assume that the bequest tax credit rate varies with the skill level of the household, while the inheritance tax is constant. Therefore, we can write the price of bequests as $p_i = (1 - \tau_i)/(1 - t)$ for $i = 1, 2$. Later we consider the consequences of the government differentiating the inheritance tax according to the wage-type of the recipient child.

2.2 Children's Behavior

Children's preferences are the same as non-donor parents' as given in (1). Like the parents, children face a nonlinear income tax schedule. We assume that the government applies separate income tax schedules to the parents and the children. The budget constraint for a child with a donor parent of wage-type i is $x = y - T(y) + b(p_i)$, and with a non-donor parent is $x = y - T(y)$.

Given a tax function $T(y)$, a type$-k$ child with a type$-i$ donor parent solves the following problem:

$$\max_{\{y\}} \; y - T(y) + b(p_i) - h\Big(\frac{y}{w_k}\Big).$$

The first-order condition for this problem gives

$$h'\Big(\frac{y}{w_k}\Big) = (1 - T'(y))w_k$$

which determines income y_k. As there are no income effects on labor supply, whether a child has a donor parent or not does not affect their labor supply decision.

We now turn to the government's problem.

3 The Government Problem

The government chooses nonlinear income tax functions $T(y)$ for both the parents and the children, bequest tax credit rates τ_1 and τ_2 for the two parent wage-types, and an inheritance tax rate t. As usual, we solve this using the direct approach, letting the government choose consumption and income for both the two parent wage-types, denoted by c_i and y_i, and the two children wage-types, denoted by $\bar{c}_k$ and $\overline{y}_k$, as well as the inheritance tax t and bequest tax credits τ_i, for $i, k = 1, 2$. In fact, since household behavior and all budget constraints depend only on the net price of bequests $p_i = (1-\tau_i)/(1-t)$, the inheritance tax t is redundant and could be set at any rate including zero.[5] We therefore treat $p_i,\ i = 1, 2$, as control variables for the government.

The government does not count the utility of bequests to donors, $f(b)$. From the point of view of the government, the utility of a donor household is therefore given by $x - h(\ell)$. Consider a type$-i$ donor. Using the budget constraint (3) and assuming the donor chooses b_i optimally, the government's value of utility of a type$-i$ donor, referred to hereafter as the social utility of a type$-i$ donor, can be defined as follows:

$$V^i\Big(c, y, \frac{1-\tau_i}{1-t}\Big) \equiv c - \frac{1-\tau_i}{1-t} b\Big(\frac{1-\tau_i}{1-t}\Big) - h\Big(\frac{y}{w_i}\Big)$$

or equivalently, using (7), by

$$V^i(c, y, p_i) \equiv c - p_i b(p_i) - h\Big(\frac{y}{w_i}\Big).$$

Differentiating this, we obtain

$$V_c^i = 1, \quad V_y^i = -\frac{h'(\ell)}{w_i}, \quad V_p^i = -b(p_i) - p_i b'(p_i). \tag{9}$$

Given p_i, the slope of an indifference curve in $c, y-$space is

$$\left.\frac{dc}{dy}\right|_{V^i, p_i} = -\frac{V_y^i}{V_c^i} = \frac{h'(\ell)}{w_i}.$$

For the type$-i$ non-donor parent, utility from the government's point of view is given simply by $v^i(c, y) \equiv c - h(y/w_i)$, where $v_c^i = 1$ and $v_y^i = h'(\ell)/w_i$. Indifference curves have the slope

$$\left.\frac{dc}{dy}\right|_{v^i} = -\frac{v_y}{v_c} = \frac{h'(\ell)}{w_i}.$$

[5] In principle, the inheritance tax rate could also differ by the type of the donor, but this would not give the government any more degrees of freedom, and would add unnecessary complexity. Later we consider the consequences of allowing the inheritance tax rate to depend on the wage-type of the recipient child.

Thus, the donor and non-donor parents have the same indifference curves in $c, y-$space, so they cannot be separated by the income tax. (Recall that we have assumed that the income tax is not contingent on bequests.) Since $-w_i dc/dy|_{V^i,p_i}$ is decreasing in w_i, the Spence-Mirrlees conditions are satisfied, so although donor parents cannot be separated from non-donor parents by a nonlinear income tax, high- and low-wage types can be separated.

In the case of heirs, social utility and private utility are identical. The value of utility of a type$-k$ child with a type$-i$ donor parent is defined as follows:

$$R^{ki}(\overline{c}, \overline{y}, p_i) \equiv \overline{c} + b(p_i) - h\Big(\frac{\overline{y}}{w_k}\Big) \tag{10}$$

where

$$R^{ki}_{\overline{c}} = 1, \quad R^{ki}_{\overline{y}} = -h'(\ell)/w_k, \quad R^{ki}_p = b'(p_i) < 0.$$

For a type$-k$ child with non-donor parents, utility is

$$r^k(\overline{c}, \overline{y}) = v^k(\overline{c}, \overline{y}) = \overline{c} - h\Big(\frac{\overline{y}}{w_k}\Big) \tag{11}$$

where

$$r^k_{\overline{c}} = 1, \quad r^k_{\overline{y}} = -h'(\ell)/w_k < 0.$$

As there are no income effects on labor supply, the income tax schedule cannot separate the children with inheritances from those without, but can separate high-wage children from low-wage children.

The government maximizes an additive and strictly concave social welfare function in social utilities of parents (which have been purged of the utility of bequests to donors) and in utilities of children, subject to a budget constraint and incentive constraints on non-donor parents, donor parents, and children. Following Farhi and Werning (2010), the government discounts the children's utility by $\alpha \in [0, 1]$. Specifically, social welfare is:

$$\begin{aligned}
\mathcal{W} = \sum_{i=1,2} n_i \Big(d_i W\left(V^i(c_i, y_i, p_i)\right) + (1-d_i) W\left(v^i(c_i, y_i)\right) \Big) \\
+ n_1 \pi d_1 \alpha W\left(R^{11}(\overline{c}_1, \overline{y}_1, p_1)\right) + \Big(n_1 \pi (1-d_1) + n_2(1-\pi)(1-d_2)\Big) \alpha W\left(r^1(\overline{c}_1, \overline{y}_1)\right) \\
+ n_2 \pi d_2 \alpha W\left(R^{22}(\overline{c}_2, \overline{y}_2, p_2)\right) + \Big(n_2 \pi (1-d_2) + n_1(1-\pi)(1-d_1)\Big) \alpha W\left(r^2(\overline{c}_2, \overline{y}_2)\right) \\
+ n_2(1-\pi) d_2 \alpha W\left(R^{12}(\overline{c}_1, \overline{y}_1, p_2)\right) + n_1(1-\pi) d_1 \alpha W\left(R^{21}(\overline{c}_2, \overline{y}_2, p_1)\right)
\end{aligned} \tag{12}$$

where $W'(\cdot) > 0 > W''(\cdot)$.[6]

The government faces an intertemporal budget constraint covering both the parents' and the children's generations. This implies that the government can make intergenerational transfers implicitly through the income tax. Therefore, the inheritance tax t, or more accurately the price of bequests p_i, is not needed as an instrument for making transfers between parents and children, allowing us to focus on other roles for the inheritance tax. Assuming for simplicity that the interest rate is zero, the government's budget constraint is

$$\sum \Big(n_i(y_i - c_i + tB_i - \tau_i B_i) + \Big(n_i\pi + n_{-i}(1-\pi) \Big) (\overline{y}_i - \overline{c}_i) \Big) = G,$$

where G is the given level of government expenditures. Using $B_i = b_i/(1-t)$ and $p_i = (1-\tau_i)/(1-t)$, this can be written as

$$n_1(y_1 - c_1) + \Big(n_1\pi + n_2(1-\pi) \Big) (\overline{y}_1 - \overline{c}_1) + n_1 d_1 (p_1 - 1) b(p_1)$$

$$n_2(y_2 - c_2) + \Big(n_2\pi + n_1(1-\pi) \Big) (\overline{y}_2 - \overline{c}_2) + n_2 d_2 (p_2 - 1) b(p_2) = G. \tag{13}$$

The incentive constraints faced by the government hinge on what the government, or its income tax authority, observes. In the case of non-donor parents, the incentive constraint is the standard one:[7]

$$c_2 - h\Big(\frac{y_2}{w_2}\Big) \geqslant c_1 - h\Big(\frac{y_1}{w_2}\Big). \tag{14}$$

The donors' incentive constraint is more subtle. Donors of a given wage-type will choose the same consumption-income bundle as their non-donor counterparts, given their quasilinear preferences, but their bequests depend on the bequest tax credit they receive. Although we are assuming for simplicity that the income tax structure is not conditioned on bequests, nonetheless the bequest tax credit depends on one's income. That implies that if a type-2 donor mimics a type-1 donor, the bequest tax credit is τ_1 rather than τ_2. Given the separability assumption on the joy-of-giving function, the mimicker would choose to leave $b(p_1)$. Therefore, in order for a type-2 donor to mimic a type-1, not only would c_1 and y_1

[6] If there were no aversion to inequality, so $W''(\cdot) = 0$, government policy would rely solely on efficiency since the marginal utility of consumption is fixed at unity here. Since there are no externalities associated with bequests in our model, there would be no role for government policy.

[7] The government can observe who is a parent and who is a child, so separate incentive constraints apply to each. As mentioned, the government can aggregate its budget constraint over both parents and children, so the income tax system in effect tags parents and children.

be chosen, but so would the bequest of a type-1 person, $b(p_1)$. The incentive constraint for donors can then be written:

$$c_2 - p_2 b(p_2) + f(b(p_2)) - h\Big(\frac{y_2}{w_2}\Big) \;\geqslant\; c_1 - p_1 b(p_1) + f(b(p_1)) - h\Big(\frac{y_1}{w_2}\Big). \tag{15}$$

Only one of the two incentive constraints for parents, (14) and (15), will generally be binding in a social optimum. If $\tau_2 > \tau_1$, so $p_1 > p_2$, (15) will be slack when (14) is binding since $f(b(p_i)) - p_i b(p_i)$ is decreasing in p_i.[8] Alternatively, if $\tau_1 > \tau_2$, so $p_2 > p_1$, (14) will be slack when (15) is binding for the same reason. Only in the unlikely event that $p_1 = p_2$ in the optimum, would both constraints on the parents be binding at the same time. In the problems we consider, at least one of the constraints will be binding given the government's redistributive motive.

The incentive constraint for the children takes the standard form

$$\bar{c}_2 - h\Big(\frac{\bar{y}_2}{w_2}\Big) \geqslant \bar{c}_1 - h\Big(\frac{\bar{y}_1}{w_2}\Big). \tag{16}$$

since the bequest a child receives only depends on the parent's wage-type and not on any characteristic of the child. This incentive constraint will always be binding.

4 Zero Social Weight on Children's Utility

We begin by assuming a myopic government that acts solely in the interest of the parents, and must balance a budget covering the parents' lifetime. The government can however choose the inheritance tax rate and collect the tax revenue when the bequest is made. This case serves a useful pedagogical purpose of focusing on how the bequest tax regime treats donor versus non-donor parents. This case is also comparable with the base case in Farhi and Werning (2010) where children are passive recipients of bequests and no weight is given to their utility in the government's objective function. (Of course, the child's utility is taken into account indirectly through the parent's altruism.) The difference is that in our model no weight is given to the utility of bequests to donors, and we highlight the implications of this for the tax treatment of the parents when there is heterogenous bequest behavior.

The Lagrangian expression for the government in this case is

$$\mathcal{L} = \sum_{i=1,2} n_i\Big(d_i W(V^i(c_i, y_i, p_i)) + (1-d_i)W(v^i(c_i, y_i))\Big) + \gamma\left(c_2 - h\Big(\frac{y_2}{w_2}\Big) - c_1 + h\Big(\frac{y_1}{w_2}\Big)\right)$$

[8] Given that individuals optimize over the choice of net bequests, the derivative of this expression with respect to p_i reduces to $-b(p_i) < 0$.

$$+\gamma^d\left(c_2 - p_2 b(p_2) + f(b(p_2)) - h\left(\frac{y_2}{w_2}\right) - c_1 + p_1 b(p_1) - f(b(p_1)) + h\left(\frac{y_1}{w_2}\right)\right)$$

$$+\lambda\left(n_1\left(y_1 - c_1 + d_1(p_1 - 1)b(p_1)\right) + n_2\left(y_2 - c_2 + d_2(p_2 - 1)b(p_2)\right) - G\right) \tag{17}$$

where either $\gamma = 0$ or $\gamma^d = 0$, depending on whether in the optimum $p_2 > p_1$ or $p_1 > p_2$, respectively.

The first-order conditions can be written as follows, using (9):

$$n_1 d_1 W'(V^1) + n_1(1 - d_1)W'(v^1) - (\gamma + \gamma^d) - \lambda n_1 = 0 \tag{18}$$

$$-n_1 d_1 W'(V^1)\frac{h'(\ell_1)}{w_1} - n_1(1 - d_1)W'(v^1)\frac{h'(\ell_1)}{w_1} + (\gamma + \gamma^d)\frac{h'(\hat{\ell}_2)}{w_2} + \lambda n_1 = 0 \tag{19}$$

$$n_2 d_2 W'(V^2) + n_2(1 - d_2)W'(v^2) + (\gamma + \gamma^d) - \lambda n_2 = 0 \tag{20}$$

$$-n_2 d_2 W'(V^2)\frac{h'(\ell_2)}{w_2} - n_2(1 - d_2)W'(v^2)\frac{h'(\ell_2)}{w_2} - (\gamma + \gamma^d)\frac{h'(\ell_2)}{w_2} + \lambda n_2 = 0 \tag{21}$$

$$-n_1 d_1 W'(V^1)(b(p_1) + p_1 b'(p_1)) + \lambda n_1 d_1\left((p_1 - 1)b'(p_1) + b(p_1)\right) + \gamma^d b(p_1) = 0 \tag{22}$$

$$-n_2 d_2 W'(V^2)(b(p_2) + p_2 b'(p_2)) + \lambda n_2 d_2\left((p_2 - 1)b'(p_2) + b(p_2)\right) - \gamma^d b(p_2) = 0 \tag{23}$$

Before turning to the general characterization of the solution to this problem, it is useful to consider the optimum achieved when the government can observe wage-types. This will provide insight into the optimal effective prices of net bequests, and therefore bequest tax credit rates for the two wage-types.

4.1 Full Information Benchmark

Suppose the government knows the wage-type of each worker, and whether they are donors. For comparison purposes, we assume the government is restricted to using the same policy instruments as in the imperfect information case, that is, a nonlinear income tax system that is not conditioned on donor status, as well as an inheritance tax and bequest tax credits. Given the above discussion, the optimal choices of τ_i and t can be subsumed in the optimal choice of the price of net bequests p_i. The outcome will not be first-best, but it will serve to clarify the efficiency and equity effects of proportional bequest taxation in a setting with redistributive income taxes. Of course, the equity effects we are considering in this section concern only the donors. In a later section, we introduce the welfare of the inheritors.

Neither incentive constraint will be binding, so the government freely chooses c_i, y_i and p_i for $i = 1, 2$ to maximize (17) with $\gamma = \gamma^d = 0$. The first-order conditions (18)–(23) reduce to:

$$d_i W'(V^i) + (1 - d_i) W'(v^i) - \lambda = 0 \qquad i = 1, 2 \qquad (24)$$

$$-d_i W'(V^i) \frac{h'(\ell_i)}{w_i} - (1 - d_i) W'(v^i) \frac{h'(\ell_i)}{w_i} + \lambda = 0 \qquad i = 1, 2 \qquad (25)$$

$$-W'(V^i)\Big(b(p_i) + p_i b'(p_i)\Big) + \lambda\Big((p_i - 1) b'(p_i) + b(p_i)\Big) = 0 \qquad i = 1, 2 \qquad (26)$$

Combining (24) and (25) for each wage-type, we obtain

$$\frac{h'(\ell)}{w_i} = 1.$$

This implies from (8) that the marginal income tax rate for both wage-types is zero, so the income tax is non-distorting, as expected.

From (24) for each wage type, we obtain:

$$d_1 W'(V^1) + (1 - d_1) W'(v^1) = d_2 W'(V^2) + (1 - d_2) W'(v^2) = \lambda. \qquad (27)$$

This is an equity condition that says that the average marginal social utility of consumption for donors and non-donors is equated for the two wage-types, and is equal to the shadow value of government revenue. As we see below, the marginal social utility of donors and non-donors is typically not equalized because of the behavioral effects induced by proportional bequest taxation, which is the policy instrument intended to compensate donors for their loss of consumption due to bequests. From (27) and recalling our assumption that $d_2 > d_1$, we can show that $W'(V^1) > W'(V^2)$.[9] This condition will be useful below in characterizing optimal bequest tax credits.

Consider now condition (26) determining the price of net bequests p_i. Note first that if there is no behavioral response to taxation, for example, if the inheritance tax and bequest tax credits were both lump-sum, then $b'(p_i) = 0$ and $W'(V^i) = \lambda$. This implies

---

[9] Assume $d_2 = 1 > d_1 > 0$. Then from (27), $W'(V^2) = d_1 W'(V^1) + (1 - d_1) W'(v^1)$ which implies that $V^1 < V^2 < v^1$. Consider decreasing d_2 slightly. By continuity, we still have $V^2 > V^1$. Assume next that $d_2 = d_1$. In this case, there will be a uniform tax credit and from (27) $V^1 = V^2 < v^1 = v^2$. Using the relationship that $V^i = v^i - p_i b(p_i)$, consider a small increase in d_2. This will reduce the average marginal social utility of the high wage-types and (23) will no longer bind. For (27) to remain binding either τ_2 or v^2 must increase (p_2 decrease) or either τ_1 or v^1 must decrease (p_1 increase). Any of these changes to keep (27) binding will increase V^2 relative to V^1. This continues to hold for further increases in d_2.

from (27) that $W'(V^i) = W'(v^i)$, which requires that $p_i = 0$. That is, full tax credit is given for bequests ($\tau_i = 1$), reflecting the fact that the government treats bequests simply as a reduction in donors' consumption with no offsetting social benefit. Of course, although donors and non-donors have the same social utility, donors also receive the utility of bequests so have higher private utility.

Once we take into account donors' behavioral responses, full tax credits for bequests will not be given. To see this, evaluate the first-order condition (26) at $p_i = 0$, or $\tau_i = 1$. With $p_i = 0$ for $i = 1, 2$, both types of donor parents will choose $b^i = \bar{b}$. Furthermore, $V^1 = v^1$ and $V^2 = v^2$ and it follows from (27) that $\lambda = W'(V^i) = W'(v^i)$ for $i = 1, 2$. Therefore, the lefthand side of (26) reduces to $-\lambda b'(0) > 0$.[10] This implies that the government would want to increase p_i above zero, that is, reduce τ_i below unity. The intuition for this is as follows. Starting at $\tau_i = 1$, a small reduction in τ_i will cause bequests to fall. The government will save revenue because the bequest tax credit is reduced. At the same time, less tax revenue will be collected through the inheritance tax. Since $\tau_i = 1 > t$, government revenue will increase on balance, which is the only thing the government cares about.

To determine the conditions characterizing optimal bequest tax credits, note that donor and non-donor parents of a given wage-type only differ by the bequest they leave. In particular, $V^i < v^i$ for any $\tau_i < 1$. It then follows from (27) that $W'(V^i) > \lambda > W'(v^i)$. Using the definitions of ε_p and ε_b, condition (26) can be rewritten as:

$$\Big(\frac{W'(V^i)}{\lambda} - 1\Big)\left(\frac{1-\varepsilon_p^i}{\varepsilon_p^i}\right) = \Big(\frac{W'(V^i)}{\lambda} - 1\Big)(\varepsilon_b^i - 1) = \frac{1}{p_i} \qquad i = 1, 2. \tag{28}$$

Given that $W'(V^i) > \lambda$, this can be satisfied only if $\varepsilon_b^i > 1$. Given our assumption that ε_b^i is increasing in b^i (as in the simple quadratic form for $f(b)$ discussed earlier), this requires that b^i has to be not too far below $\bar{b}$.

From (28), we can also deduce that the government must impose differential tax credits. Suppose $p_1 = p_2 > 0$, then both types of donor parents are choosing the same bequest and $\varepsilon_b^1 = \varepsilon_b^2$ since the elasticity only depends on $f(b^i)$ which is the same for both donor types. It follows from (28) that $V^1 = V^2$ which implies that $v^1 = v^2$ given that the

---

[10] If the government could condition the income tax on donor status as well as wage-type, then the government would set $\lambda = W'(V^i) = W'(v^i)$ for $i = 1, 2$. A bequest tax credit would only generate the negative revenue effect as identified above as a result of individuals' behavioral responses and consequently would not be used. An income tax conditioned on both donor status and wage-type is a sufficient policy instrument to redistribute between wage-types and between donors and non-donors.

tax credits are the same. This cannot be satisfied if $d_2 > d_1$ given the expression for λ from (27). Therefore, it must be that the tax credits are differentiated by the wage-type of the donor. A uniform tax credit is only optimal if there is the same proportion of donors of each wage-type. In the full-information case, the government can redistribute between wage-types using the income tax. However, the income tax does not redistribute between donors and non-donors. The bequest tax credit serves that purpose, but only imperfectly since it encourages bequests so increases the cost to the government. The case for redistributing between donors and non-donors and imposing differential bequest tax credits depends on the relative proportion of donors of a given wage-type.

We can be more precise about the relationship between p_1 and p_2 given the assumptions we have made. Recall from (27) that $W'(V^1) > W'(V^2)$, which followed from the assumption that $d_2 > d_1$. That implies from (26) that

$$\frac{b'(p_1)}{b(p_1) + p_1 b'(p_1)} < \frac{b'(p_2)}{b(p_2) + p_2 b'(p_2)}.$$

It is straightforward to show that the ratio in this inequality is decreasing in p_i.[11] Therefore, in the full-information optimum, $p_1 > p_2$ (or $\tau_2 > \tau_1$).

We are unable to say whether p_i is greater than or less than unity. This ambiguity arises because the government does not put any weight on the benefit to the donors of leaving a bequest. To see this, suppose on the contrary that the donor's full utility $U(\cdot)$ as given by (1), which includes the benefit of bequests $f(b)$, enters the government's objective. Then, we would write $U^i(c, y, p_i) = c - p_i b(p_i) + f(b(p_i)) - h(y/w_i)$, so $U^i_p = -b(p_i) - p_i b'(p_i) + f'(b(p_i))b'(p_i) = -b(p_i)$, since $f'(b(p_i)) = p_i$. Thus, $U^i_p < V^i_p$ in (9). The first-order condition on p_i in the full-information case could then be written as

$$\Big(\lambda - W'(U^i)\Big) b(p_i) = \lambda(1 - p_i) b'(p_i) > 0 \qquad i = 1, 2$$

where the inequality sign follows from (27) and the fact that $U^i > u^i$ if the parent donor is optimally choosing to leave a positive bequest. Therefore, given that $b'(p_i) < 0$, the government will set $p_i > 1$ ($\tau_i < 0$ if $t = 0$) for both types of donor parents. That is,

[11] Dropping the subscripts, differentiate the ratio with respect to p:

$$\frac{d}{dp}\left(\frac{b'(p)}{b(p) + p b'(p)}\right) = \frac{b''}{b + pb'} - \frac{b'}{(b + pb')^2}(2b' + pb'') = \frac{bb'' - 2(b')^2}{(b + pb')^2} < 0$$

since by (5), $b'' = -f'''(b) b'(p)/(f''(b))^2 \leqslant 0$ given that $b'(p) < 0$ and $f'''(b) \leqslant 0$.

the government wants to redistribute from the donor parents to the non-donor parents.[12] Ignoring the utility of bequests $f(b)$ changes the direction of redistribution. The ranking of the donor's utility (from the point of view of the government) relative to the non-donor's utility reverses, that is $W'(V^i) > \lambda > W'(v^i)$. The government wants to redistribute from non-donor parents to donor parents. This reversal alone would have $p_i < 1$, but then there is the addition of the $-W'(V^i)p_i b'(p_i)$ term in (26) that does not appear in the above and that pushes p_i up, making it ambiguous whether p_i will be greater or less than one. This term reflects the government's sole concern with the expenditure on net bequests and not the benefit to donors of leaving a bequest. Consequently, the behavioral effect on bequest expenditure of an increase in the effective price of bequests is viewed by the government as a positive social marginal benefit. If the government also cared about $f(b)$, then by the envelope theorem this positive effect would be completely offset by the behavioral effect on the joy-of-giving term of an increase in the effective price of bequest. Therefore, by ignoring $f(b)$ the government may want p_i to be greater or less than unity.

The following summarizes the results we have obtained in the full-information benchmark case when the heirs are not taken into consideration.

Summary of Results in the Full-Information Benchmark

When the government can observe the wage-types of donors and non-donors, and imposes a nonlinear income tax that is not conditioned on donor status along with a less than 100 percent inheritance tax and a wage-type-specific bequest tax credit, the following results apply.

1. *The optimal solution determines the effective price of net bequests, p_i, so the absolute size of the inheritance tax and the bequest tax credit are indeterminate and are set to satisfy $p_i = (1 - \tau_i)/(1 - t)$ for each type$-i$ person. In particular, the inheritance tax is redundant so could be set to zero.*
2. *The optimal value of p_i will be positive, but can be greater or less than one. This implies that the bequest tax credits are less than 100 percent for both wage-types of donors.*
3. *Donors of each wage-type will be worse-off (from the government's point of view) than non-donors of both wage-types since $W'(V^i) > \lambda > W'(v^i)$ for $i = 1, 2$.*
4. *Among donors, social utility V^i will be highest for the wage-type with the highest*

[12] Note that in the case where the government is not averse to inequality so $W''(U^i) = 0$, then the above expression implies $p_i = 1$.

proportion of donors, and they will obtain a higher bequest tax credit than the other wage-type. The ranking of non-donor social utilities v^i is ambiguous.

The intuition behind these results is as follows. With full information, the government can redistribute between wage-types using non-distorting (lump-sum) taxes and transfers. However, it cannot distinguish donors from non-donors in each skill class. Since the utility of bequests is ignored by the government, it would like to redistribute from non-donors to donors and uses the only instruments at its disposal, which are bequest tax credits. Full compensation of donors is not optimal since bequests are responsive to the bequest tax credit. The choice of bequest tax credit rates favors the high-skilled persons since a higher proportion of them are donors so there is a greater redistributive benefit obtained per unit of revenue lost.

4.2 Imperfect Information

Now suppose that the government cannot observe wage-types. Incentive constraints for the donors and non-donors are now relevant, and the government's full Lagrangian expression (17) applies. The government cannot implement lump-sum redistribution between wage-types, leading to the bequest tax credit being used both to redistribute between donors and non-donors and to redistribute between wage-types.

Eqs. (18)–(21) characterize the optimal income tax structure. These results are standard. For example, from (20) and (21), we obtain $h'(\ell_2)/w_2 = 1$, which implies a zero marginal tax rate at the top. The marginal tax rate for low-wage workers will be less than unity. From (18) and (20), we obtain a standard equity condition that the weighted average of the marginal social utility of consumption for all parents (donors and non-donors) is equal to the marginal cost of raising an additional unit of tax revenue λ, where the weights are given by the population share of each of the four types of parents (high or low wage and donor or non-donor):

$$n_1 d_1 W'(V^1) + n_1(1-d_1)W'(v^1) + n_2 d_2 W'(V^2) + n_2(1-d_2)W'(v^2) = \lambda \qquad (29)$$

with the weights summing to unity since $n_1 + n_2 = 1$.

Note the difference between (29) and (27). With full information the government is able to equate the average social marginal utility of consumption of the high- and low-wage types using the nonlinear income tax, leaving the bequest tax credit to redistribute from the non-donors to the donors. With imperfect information, that targeting of policies is no longer possible. The government is constrained in redistributing between wage-types, and that has implications for the role of the bequest tax credit. As in the full-information case, $V^i < v^i$ for $i = 1, 2$ and consequently, $W'(V^i) > W'(v^i)$, but unlike the full-information case the ranking of all of the social marginal utilities relative to λ is not clear. We can, however, say from (18) and (20) that $W'(V^1) > \lambda > W'(v^2)$ given that either $\gamma > 0$ or $\gamma^d > 0$ and from the incentive constraint (10) we know that $v^2 > v^1$ or $W'(v^2) < W'(v^1)$ since high-wage mimickers have to supply less labor to earn y_1. We come back to these rankings below.

Eqs. (22) and (23) determine the optimal prices of net bequests p_i. These equations differ from the full-information benchmark by the terms involving the donor incentive constraints. Recall from above that either the non-donor's or the donor's incentive constraint, (14) or (15), will be binding depending on whether or not $p_1 < p_2$ (i.e., $\tau_1 > \tau_2$) in the optimum. In the full-information case where the incentive constraints did not apply, we found that $p_1 > p_2$, or $\tau_2 > \tau_1$, unambiguously. With imperfect information, the relative size of τ_1 and τ_2 is no longer clear. That is because they serve not only to redistribute from non-donors to donors, but they also influence redistribution between wage-types, as the latter terms in (22) and (23) involving γ^d indicate. Depending on the relative importance of bequest tax credits for redistributing between wage-types as opposed to between donor types, that is, whether incentive constraint (15) binds or (14) binds, τ_1 could be larger or smaller than τ_2 in the optimum. We illustrate each case in turn.

Case 1: Incentive Constraint on Donors Binding

With incentive constraint (14) slack and incentive constraint (15) binding, it must be that $\tau_1 > \tau_2$ or $p_1 < p_2$ and low-wage donors will leave a larger bequest than high-wage donors, $b(p_1) > b(p_2)$. High-wage donors' utility net of the joy-of-giving, however, will be higher than low-wage donors as in the full-information outcome. From the binding incentive constraint (15), $V^2 + f(b^2) > V^1 + f(b^1) > V^1 + f(b^2)$ where the first inequality follows since high-wage mimickers have to supply less labor to earn y_1 and the second inequality follows from $b(p_1) > b(p_2)$. Unlike the full-information outcome, the result that $V^2 > V^1$ does not follow from the assumption that there is a greater share of high-wage donors, $d_2 > d_1$. In other words, assuming $d_1 < d_2$ would not affect the ranking of donors' utility from the point of view of the government with imperfect information, but would reverse the ranking of V^1 and V^2 with full information.

In this case, it is not possible to rank V^2 relative to v^1. To see this, we can combine conditions (22) and (26) and conditions (24) and (27) by solving for γ^d to obtain:

$$d_iW'(V^i)+(1-d_i)W'(v^i)-\lambda = d_i[W'(V^i)-\lambda](b(p_i)+p_ib'(p_i))+\lambda d_ib'(p_i) \quad i = 1, 2. \quad (30)$$

Eq. (30) states that the effect on the government's Lagrangian of a marginal increase in c_i must be the same as a marginal increase in p_i for both wage types. For the low-wage parent, the above expressions are positive and for the high-wage parent the expressions are negative. Therefore, we know that $W'(v^1) > \lambda$, but $W'(V^2)$ could be greater or less than λ. High wage-type donors (net of joy-of-giving) could be better-off or worse-off than low-wage non-donors. Recall, in the full-information case donor parents were worse-off from the point of view of the government than non-donor parents of each wage-type.

Case 2: Incentive Constraint on Non-Donors Binding

When the incentive constraint on donors (15) is slack, it must be that $\tau_2 > \tau_1$ or $p_2 < p_1$. The optimal bequest tax credit conditions (22) and (23) are the same as in the full-information case (see eq. (26) or (28).) Only the income tax rates are distorted. As $W'(V^1) > \lambda$ from (18), it follows from (28) that $\varepsilon_b^1 > 1$. Since ε_b is increasing in b and $b(p_2) > b(p_1)$, it follows that $\varepsilon_b^2 > \varepsilon_b^1 > 1$. Therefore, from (28) it must also be the case that $W'(V^2) > \lambda$. Finally, taking the difference of (22) and (23) with $\gamma^d = 0$ and using the above observations it follows (as in the full-information case) that $V^2 > V^1$ and both donor types (net of joy-of-giving) are worse off than the non-donor parents.

Summary of Results with Imperfect Information

When the government cannot observe the wage-types of donors and non-donors, and imposes a nonlinear income tax that is not conditioned on donor status along with a less-than-100 percent inheritance tax and a wage-type-specific bequest tax credit, the following results apply.

1. *As in the case of full information, the optimal solution determines the effective price of net bequests, p_i, so the absolute size of the inheritance tax and the bequest tax credit are indeterminate and are set to satisfy $p_i = (1 - \tau_i)/(1 - t)$ for each type–i person.*

2. *The optimal value of p_i will be positive and the bequest tax credits are less than 100 percent for both wage-types of donors.*

3. *The high-wage parents will be better off than low-wage parents regardless of donor status. But the high-wage donor parents may receive a higher or lower bequest tax credit than the low-wage donors.*

5 Positive Social Weight on the Children's Utility

Now suppose the government puts positive social weight on children's utilities. The social welfare function, budget constraint and incentive constraints are as in Section 3. The government chooses consumption-income bundles for the two types of parents, c_i, y_i, and for the two types of children, $\bar{c}_k, \bar{y}_k$, and the net price of bequests for the donor parents, p_i. The first-order conditions on c_i and y_i are given by (18)–(21). The first-order conditions on $\overline{c}_k$, $\overline{y}_k$ are comparable, and are given in the Appendix, along with those for p_i. As in the case without a positive social weight on children's utility, it is useful to begin with the case where the government can observe wage-types.

5.1 Full-Information Benchmark

Suppose the government can observe wage-types, so the incentive constraints do not apply. The government chooses consumption-income bundles for the two types of parents, c_i, y_i, and for the two types of children, $\overline{c}_k, \overline{y}_k$, and the net price of bequests for the donor parents, p_i, to maximize the sum of social utilities (12), subject only to the resource constraint (13). The first-order conditions on c_i and y_i are (24) and (25) as in the previous full-information problem without zero social weight on the children's utility, which yield the equity condition (27) and the zero-marginal income tax rate condition, $h'(\ell)/w_i = 1$. Combining the first-order conditions for $\overline{c}_k$ and $\overline{y}_k$ when the incentive constraint is not binding (eqs. $(A.1)$–$(A.4)$ in the Appendix with $\phi = 0$) yields

$$\frac{h'(\overline{\ell})}{w_k} = 1, \quad k = 1, 2.$$

This confirms that marginal income tax rates for the children are also zero when the incentive constraints are not binding.

Next, consider the equity condition applying to the children. From the first-order conditions on $\overline{c}_k$ (eqs. $(A.1)$ and $(A.3)$ with $\phi = 0$), we have

$$\begin{aligned}\frac{\lambda}{\alpha} &= \frac{n^{1d}}{n^1}\left[\left(\frac{n^{11}}{n^{1d}}\right)W'\left(R^{11}\right) + \left(1 - \frac{n^{11}}{n^{1d}}\right)W'\left(R^{12}\right)\right] + \left(1 - \frac{n^{1d}}{n^1}\right)W'\left(r^1\right) \\ &= \frac{n^{2d}}{n^2}\left[\left(\frac{n^{22}}{n^{2d}}\right)W'\left(R^{22}\right) + \left(1 - \frac{n^{22}}{n^{2d}}\right)W'\left(R^{21}\right)\right] + \left(1 - \frac{n^{2d}}{n^2}\right)W'\left(r^2\right) \qquad (31)\end{aligned}$$

where $n^k = n_k\pi + n_i(1-\pi)$ is the number of type$-k$ children, $n^{kd} = n_k\pi d_k + n_i(1-\pi)d_i$ is the number of type$-k$ children who receive a bequest, and $n^{kk} = n_k\pi d_k$ is the number of type$-k$ children with a type$-k$ donor parent where $k \neq i = 1, 2$. As with (27) above, this equity condition requires that the average marginal social utility of consumption be equal for high-wage and low-wage children receiving three different amounts of bequests, $b(p_1), b(p_2)$ and zero. With positive bequests, $R^{kk} > r^k$ and $R^{ki} > r^k$ so $W'(R^{kk}) < W'(r^k)$ and $W'(R^{ki}) < W'(r^k)$ which implies from (31) that

$$W'(r^i) > \frac{\lambda}{\alpha} > \left(\frac{n^{kk}}{n^{kd}}\right)W'\left(R^{kk}\right) + \left(1 - \frac{n^{kk}}{n^{kd}}\right)W'\left(R^{ki}\right)$$

and $\lambda/\alpha > \min\left[W'\left(R^{kk}\right), W'\left(R^{ki}\right)\right]$.

Finally, consider the choice of p_i for $i = 1, 2$. The first-order conditions on the p_i's with the donor's incentive constraint not binding (eqs. $(A.5)$ and $(A.6)$ in the Appendix with $\gamma^d = 0$) can be written as follows:

$$-\left[W'\left(V^i\right) - \lambda\right]\left(b(p_i) + p_i b'(p_i)\right) - \lambda b'(p_i) + \left[\pi\alpha W'\left(R^{ii}\right) + (1-\pi)\alpha W'\left(R^{ki}\right)\right] b'(p_i) = 0 \qquad i \neq k = 1, 2 \qquad (32)$$

Relative to the full-information case without children, where the condition (26) applies, there is an additional cost of increasing p_i since the bequests benefit the children. This is reflected in the last term in (32), which is negative since $b'(p_i) < 0$.

To interpret the implications of (32) for optimal policies, begin with the case there is no behavioral response, so $b'(p_i) = 0$. Eq. (32) implies that $W'(V^i) = \lambda$ and it follows from the equity condition for the parents given by (27) that the government would ensure that $W'(V^i) = W'(v_i)$ which implies $p_i = 0$. A zero effective price for net bequests is no longer optimal, however, if there are donors' behavioral responses. To see this, suppose $p_i = 0$ which implies that $V^i = v^i$ and from (27), $W'(V^i) = \lambda$. Further, with $p^i = 0$ both type of donor parents leave $\bar{b}$ so $R^{11} = R^{12} > r^1$, $R^{22} = R^{21} > r^2$ and it follows from (31) that $\alpha W'(R^{kk}) < \lambda < \alpha W'(r^k)$ for $k = 1, 2$. Therefore, the left-hand side of (32) with $p_i = 0$ reduces to

$$-\lambda b'(0) + \left[\pi_i \alpha W'\left(R^{ii}\right) + (1-\pi_i)\alpha W'\left(R^{kk}\right)\right] b'(0) > 0$$

so the government optimally chooses $p_i > 0$, that is, $\tau_i < 1$.

Consider now the pattern of net bequest prices, p_1 and p_2. It turns out that, unlike in the full-information case with zero social weight on the heirs, once we allow for a positive social weight on the heirs the relative magnitude of p_1 and p_2 is ambiguous. To gain some intuition, suppose we were to evaluate (32) at $p_1 = p_2$. Ignoring the last term in (32), which is the effect of p_i on the social utilities of the heirs, we would conclude that the government would want to increase p_1 relative to p_2 as shown in the full information outcome with zero social weight on the heirs. Now consider how the last term in (32) affects this conclusion. With $p_1 = p_2$, the relative magnitude of this last term (which is negative) between the two conditions on p_1 and p_2 in (32) depends on whether:

$$\pi W'\left(R^{22}\right) + (1-\pi) W'\left(R^{12}\right) \gtrless \pi W'\left(R^{11}\right) + (1-\pi) W'\left(R^{21}\right)$$

Since with $p_1 = p_2$ all heirs receive the same bequest, so $R^{11} = R^{12}$ and $R^{21} = R^{22}$ and we can rewrite the above as

$$W'\left(R^{11}\right) \gtrless W'\left(R^{22}\right).$$

Suppose it were the case that low-wage heirs are better off than high-wage heirs, so $R^{11} > R^{22}$ (or, equivalently $W'(R^{11}) < W'(R^{22})$). In this case, the weighted average of the social marginal utilities of the low-wage heirs will be less than that of the high-wage heirs. Consequently, starting from $p_1 = p_2$ the government has an additional incentive to reduce p_2 relative to p_1. Suppose instead the low-wage heirs are worse off starting at $p_1 = p_2$. Now the government has an incentive to increase p_2 relative to p_1 since the low-wage heirs have a higher social marginal utility and the net effect on the relative magnitudes of optimal effective prices of net bequests is ambiguous. A more formal demonstration that the relative sizes of p_1 and p_2 are ambiguous is given in the Appendix.

The intuition behind this ambiguity is as follows. Recall that in the full-information case without heirs, we obtained that $p_1 > p_2$, or $\tau_2 > \tau_1$. In that case, the bequest tax credit served to redistribute between non-donor and donor parents of different wage-types. Given that a higher proportion of the high-skilled were donors compared with the low-skilled, a bequest tax credit to the high-skilled was relatively more effective at redistributing from non-donors to donors. Once heirs are taken into account, redistribution between heirs receiving bequests and those not receiving bequests becomes relevant, given that the income tax system does not distinguish between them. Since skills are positively correlated between parents and children ($\pi > 1/2$) and bequests are correlated with parental skills ($d_2 > d_1$), a higher proportion of high-skilled children receive bequests than of low-skilled children. This favors giving a higher bequest tax credit to low-skilled donors to encourage them to give bequests. This desire to set $\tau_1 > \tau_2$ in order to favor low-skilled inheritors counters the opposite desire to redistribute between non-donor and donor parents, and either influence can dominate.

5.2 Imperfect Information

Now assume the government cannot condition the consumption-income bundles on the type of individual so the incentive constraint (16) for the children is binding. As explained earlier, either the incentive constraint on the donors is binding and the incentive constraint on the non-donors is slack or the opposite holds. From the first-order conditions $(A.1)$ and $(A.3)$ on $\bar{c}_k$ in the Appendix, we have

$$\frac{\lambda}{\alpha} = \frac{n_1 \pi d_1}{n_1+n_2} W'\left(R^{11}\right) + \frac{n_1(1-\pi) d_1}{n_1+n_2} W'\left(R^{21}\right) + \frac{n_1 \pi (1-d_1) + n_2(1-\pi)(1-d_2)}{n_1+n_2} W'\left(r^1\right)$$

$$+ \frac{n_2(1-\pi) d_2}{n_1+n_2} W'\left(R^{12}\right) + \frac{n_2 \pi d_2}{n_1+n_2} W'\left(R^{22}\right) + \frac{n_2 \pi (1-d_2) + n_1(1-\pi)(1-d_1)}{n_1+n_2} W'\left(r^2\right).$$

With full information, the government can equate the average social marginal utility of consumption of high- and low-wage type children using the non-linear income tax, leaving the inheritance tax credit to redistribute from those children receiving an inheritance to those who do not. As in the previous case with zero social weight on the heirs, imperfect information constrains the government's ability to redistribute between children of different wage-types, which has implications for optimal policy.

From the binding incentive constraint on the children, we know that $r^2 > r^1$ and consequently, $R^{21} > R^{11}$ and $R^{22} > R^{12}$. The average social marginal utility of heirs receiving a bequest from a low-wage donor parent is greater than the average social marginal utility of heirs receiving a request from a high-wage donor. The relative magnitude of the effective prices on net bequest continues to depend up whether the incentive constraint on the donors is binding or not. In the former case, it must be that $p_1 < p_2$ and in the latter case, $p_1 > p_2$. This continues to hold.

Intuitively, the bequest tax credit now has to take account of several sources of inequality. There is the inequality between the donor and non-donor parents and between the recipient and non-recipient children as in the full information case. The former calls for $\tau_2 > \tau_1$, and the latter the opposite. Then, there is inequality between the high- and low-skilled for both parents and children arising because of imperfect information and the incentive constraints in the nonlinear income tax systems. Both of these call for $\tau_2 < \tau_1$: this favors the type-1 parents relative to the type-2's directly, and also favours the type 1 children relative to the type 2's indirectly by reducing inheritances more on average for the latter. On balance, it is ambiguous whether τ_2 should be higher or lower than τ_1, but relative to the full information case, τ_2 would tend to be reduced relative to τ_1.

5.3 Inheritance Taxes Conditional on Child's Wage-type

The preceding analysis shows that the absolute value of an inheritance tax which applies to all children, regardless of wage-type, is indeterminate, given that the government can differentiate the bequest tax credit by wage-type of donor parent. Only the effective price of net bequests conditioned on donor-type, that is, p_i, is determined at the optimal solution. How does this result change if the inheritance tax can be conditioned on the child's wage-type?

The government will now have four policy instruments, a bequest tax credit applying to each type of donor parent and an inheritance tax applying to each type of donee child. Assuming that donor parents know the child's wage-type when making their bequest, household behaviour and all budget constraints will again depend only on the net price of bequest, which for a donor of type-i with a child of type-k is now given by $p_{ik} = (1 -$

$\tau_i)/(1-t_k)$.[13] There will be four different effective prices of bequests, but like the previous case in which the uniform inheritance tax is redundant, these four prices cannot be chosen independently. In other words, given that household behaviour depends on the effective price which is a ratio of two of the four policy instruments, one of the policy instrument is redundant. The government's optimal choice of three of the policy instruments will ensure the fourth instrument is also optimal. This holds regardless of the assumed information on the part of the government.

To see this, consider for simplicity the case with full information. The government's problem is to maximize (12) subject to (13), with each type of donor parent facing an effective net price of bequest which depends on the type of their child. Define the Lagrangian of the government's problem as a function of these effective prices, $\mathcal{L}(p_{11}, p_{12}, p_{21}, p_{22})$ where $p_{ik} = (1-\tau_i)/(1-t_k)$ for $i, k = 1, 2$. The first-order conditions on the policy instruments $\{\tau_1, \tau_2, t_1, t_2\}$ can be expressed, respectively, as follows:

$$\frac{\partial \mathcal{L}}{\partial p_{11}} \frac{-1}{1-t_1} + \frac{\partial \mathcal{L}}{\partial p_{12}} \frac{-1}{1-t_2} = 0 \tag{33}$$

$$\frac{\partial \mathcal{L}}{\partial p_{21}} \frac{-1}{1-t_1} + \frac{\partial \mathcal{L}}{\partial p_{22}} \frac{-1}{1-t_2} = 0 \tag{34}$$

$$\frac{\partial \mathcal{L}}{\partial p_{11}} \frac{-(1-\tau_1)}{(1-t_1)^2} + \frac{\partial \mathcal{L}}{\partial p_{21}} \frac{-(1-\tau_2)}{(1-t_1)^2} = 0 \tag{35}$$

$$\frac{\partial \mathcal{L}}{\partial p_{12}} \frac{-(1-\tau_1)}{(1-t_2)^2} + \frac{\partial \mathcal{L}}{\partial p_{22}} \frac{-(1-\tau_2)}{(1-t_2)^2} = 0 \tag{36}$$

With less than 100% inheritance taxes, substituting any of the two above conditions into one of the remaining two conditions yields the last condition.[14] Therefore, the government can choose any three of $(t_1, t_2, \tau_1, \tau_2)$, or equivalently and three of $(p_{11}, p_{12}, p_{21}, p_{22})$.[15]

---

[13] If, instead, bequests were made without knowledge of a child's type, then donor parents would maximize expected utility over the possible types of children in making their bequest decisions. We assume such uncertainty is resolved prior to making the bequest decision.

[14] Recall that if inheritances are completely taxed away then donor parents would optimally chose not to leave any bequests.

[15] Another way of seeing this is to note that the effective prices of net bequests are related as follows:

$$p_{11}p_{22} = \frac{1-\tau_1}{1-t_1} \cdot \frac{1-\tau_2}{1-t_2} = p_{12}p_{21}$$

so only three of the prices can be chosen independently.

Allowing the government to differentiate the inheritance tax by the wage-type of the receiving child gives the government only one more degree of freedom relative to the case of a single inheritance tax.

What can we say about the ranking of the effective prices when the inheritance tax can be conditioned on child type? Recall that with a uniform inheritance tax t and zero social weight on children, $p_1 > p_2$ in the optimum as the government wants to redistribute from non-donor to donor parents (of a given type) and there are a greater share of type-2 donor parents. As there is a positive correlation between parent-child types, there is a greater share of type-2 children who receive bequests. Therefore, when caring about the children the government would want to push down p_1 relative to p_2 to redistribute from those children (of a given type) who receive an inheritance to those who do not. The final ranking of the two effective prices is ambiguous.

Now consider the possibility that the tax faced by the child can be conditioned on their type. First, note that the proportion of type-2 children receiving a bequest is greater than the proportion of type-1 children, that is,[16]

$$\frac{\pi n_2 d_2 + (1-\pi) n_1 d_1}{\pi n_2 + (1-\pi) n_1} > \frac{\pi n_1 d_1 + (1-\pi) n_2 d_2}{\pi n_1 + (1-\pi) n_2} \tag{37}$$

One might be tempted to argue that this would call for a higher inheritance tax on type-2's relative to type-1's, that is $t_2 > t_1$. However, that intuition turns out not to be correct as the following analysis suggests.

Consider the full-information case where the government knows the wage-types of both the parents and the children. Suppose we start in the outcome considered above in which $t_1 = t_2 = t$ and p_1 and p_2 are chosen optimally. Applying the Envelope Theorem to the value function for that problem, we show in the Appendix that social welfare will increase with an incremental increase in t_2 (holding t_1 constant) if $r^2 > r^1$, and vice versa.

To determine the relationship between r^1 and r^2, consider the equity condition for the children, (31), which can be rewritten as:

$$\frac{\lambda}{\alpha} = \frac{n^{1d}}{n^1}\overline{W}'\left(R^1\right) + \left(1 - \frac{n^{1d}}{n^1}\right) W'\left(r^1\right) = \frac{n^{2d}}{n^2}\overline{W}'\left(R^2\right) + \left(1 - \frac{n^{2d}}{n^2}\right) W'\left(r^2\right) \tag{38}$$

where $\overline{W}'\left(R^k\right)$ is the average marginal utility of type$-k$ recipients. We know that $\overline{W}'\left(R^k\right) < W'\left(r^k\right)$ for $k = 1, 2$, but the relative size of $\overline{W}'\left(R^1\right)$ and $\overline{W}'\left(R^2\right)$ is ambiguous. However, as we show in the Appendix, $\overline{W}'\left(R^1\right) < \overline{W}'\left(R^2\right)$ if $\pi = 1$ and

[16] This follows from cross-multiplying and using $\pi > 1/2$. If $\pi = 1/2$, then parent and child types are independent and (37) would be an equality.

$\overline{W}'\left(R^1\right) = \overline{W}'\left(R^2\right)$ if $\pi = 1/2$. So, we expect that $\overline{W}'\left(R^1\right) < \overline{W}'\left(R^2\right)$ for $\pi > 1/2$. Given that, the following result is apparent from (37) and (38). Given that $\overline{W}'\left(R^1\right) \leqslant \overline{W}'\left(R^2\right)$, $W'\left(r^1\right) < W'\left(r^2\right)$, so $r^1 > r^2$. This implies from above that t^2 should be reduced relative to t_1.

This ambiguity is not resolved under imperfect information. The government is constrained in redistributing from type-2 to type-1 parents and children, and this will tend to reduce τ_2 relative to τ_1, and increase t_2 relative to t_1.

The following summarizes the results in the case with positive social weight on children's utility.

Summary of Results with Positive Social Weight on Children's Utility

When the government can impose a separate nonlinear income tax on each generation financed by a consolidated budget constraint as well as an inheritance tax and a wage-specific bequest tax credit, the following results apply.

1. *The optimal solution determines the effective price of net bequests, p_i, so the absolute size of the inheritance tax and the bequest tax credit are indeterminate and are set to satisfy $p_i = (1-\tau_i)/(1-t)$ for each type–i person.*

2. *The optimal value of p_i will be positive, but can be greater or less than unity. This implies that the bequest tax credits are less than 100 percent for both wage-types of donors.*

3. *Donors of a given wage-type will be worse-off (from the government's point of view) than non-donors and inheritors will be better off than non-inheritors of a given wage type.*

4. *When the government cannot observe the wage-types of parents and children, high-wage parents will be better off than low-wage parents regardless of donor status and high-wage children will be better off than low-wage children regardless of inheritance status.*

5. *The relative size of the effective price of net bequests of the two wage-types of donors is ambiguous, and reflects the conflicting roles of the effective price in redistributing from non-donor to donor parents and from inheriting to non-inheriting children.*

6. *Conditioning the inheritance tax on the wage-type of the child gives the government an additional degree of freedom. In the full-information benchmark, the inheritance tax should be higher for low-skilled children.*

7 Concluding Comments

We have studied the optimal linear tax treatment of bequests when the government does not give welfare weight to the donors' utility of bequests. There are many cogent reasons for taking this normative perspective. Perhaps the most persuasive one is that voluntary transfers to one's heirs are in principle analogous to government redistribution based on the altruistic preferences of well-off taxpayers, and there is apparently no support for counting the benefits to the taxpayers from such redistribution. The policy consequences of discounting the benefits of bequests are significant. Most important, some form of credit needs to be given to those parents who choose to leave a bequest to compensate them for their loss of personal consumption. Such compensation affects the decision to bequeath, leading to a standard equity-efficiency trade-off. We have explored this in a simple setting in which there are both donor and non-donor parents, and therefore both donee and non-donee children. Our results are in sharp contrast to recent analyses that double count the social welfare benefits of bequests, once to the donors and a second time to the donees (Kaplow 1988; Farhi and Werning 2010; Brunner and Pech 2012a,b; Piketty and Saez 2012). In this approach, the externality of bequests leads to an argument for subsidizing them, tempered by the desire to redistribution among donnees in different circumstances.

The framework of our analysis is restrictive, although similar to recent literature on bequest taxation. By restricting bequest policies to linear tax and tax credit instruments we are able to uncover the various factors that are relevant for policy. It is clear that compensating donors for the costs of their voluntary transfers leads to complicated policy prescriptions even in a simple setting. In that sense our analysis is exploratory. Ideally one might want to allow the government to implement nonlinear inheritance taxation alongside nonlinear income taxation, and that would be a next useful step.

Appendix

First-order conditions on $\bar{c}_k$ and $\bar{y}_k$

Using the properties of R^{ki} and r^k in (28) and (29):

$$n_1\pi d_1\alpha W'\left(R^{11}\right) + \Big(n_1\pi(1-d_1) + n_2(1-\pi)(1-d_2)\Big)\alpha W'\left(r^1\right)$$

$$+ n_2(1-\pi)d_2\alpha W'\left(R^{12}\right) - \phi - \lambda(n_1\pi + n_2(1-\pi)) = 0 \tag{A.1}$$

$$-n_1\pi d_1\alpha W'\left(R^{11}\right)\frac{h'(\bar{\ell}_1)}{w_1} - \Big(n_1\pi(1-d_1) + n_2(1-\pi)(1-d_2)\Big)\alpha W'\left(r^1\right)\frac{h'(\bar{\ell}_1)}{w_1}$$

$$- n_2(1-\pi)d_2\alpha W'\left(R^{12}\right)\frac{h'(\bar{\ell}_1)}{w_1} + \phi\frac{h'(\widehat{\bar{\ell}_2})}{w_2} + \lambda(n_1\pi + n_2(1-\pi)) = 0 \tag{A.2}$$

$$n_2\pi d_2\alpha W'\left(R^{22}\right) + \Big(n_2\pi(1-d_2) + n_1(1-\pi)(1-d_1)\Big)\alpha W'\left(r^2\right)$$

$$+ n_1(1-\pi)d_1\alpha W'\left(R^{21}\right) + \phi - \lambda(n_2\pi + n_1(1-\pi)) = 0 \tag{A.3}$$

$$-n_2\pi d_2\alpha W'\left(R^{22}\right)\frac{h'(\bar{\ell}_2)}{w_2} - \Big(n_2\pi(1-d_2) - n_1(1-\pi)(1-d_1)\Big)\alpha W'\left(r^2\right)\frac{h'(\bar{\ell}_2)}{w_2}$$

$$- n_1(1-\pi)d_1\alpha W'\left(R^{21}\right)\frac{h'(\bar{\ell}_2)}{w_2} - \phi\frac{h'(\bar{\ell}_2)}{w_2} + \lambda(n_2\pi + n_1(1-\pi)) = 0 \tag{A.4}$$

where ϕ is the multiplier on the incentive constraint for the children. The first-order conditions on p_i are:

$$-n_1d_1W'(V^1)(b(p_1) + p_1b'(p_1)) + n_1d_1\Big[\pi\alpha W'\left(R^{11}\right) + (1-\pi)\alpha W'\left(R^{21}\right)\Big]b'(p_1)$$

$$+\lambda n_1d_1\Big((p_1-1)b'(p_1) + b(p_1)\Big) + \gamma^d b(p_1) = 0 \tag{A.5}$$

$$-n_2d_2W'(V^2)(b(p_2) + p_2b'(p_2)) + n_2d_2\Big[\pi\alpha W'\left(R^{22}\right) + (1-\pi)\alpha W'\left(R^{12}\right)\Big]b'(p_2)$$

$$+\lambda n_2d_2\Big((p_2-1)b'(p_2) + b(p_2)\Big) - \gamma^d b(p_2) = 0 \tag{A.6}$$

Ambiguity of p_1 relative to p_2

Using $W'(V^1) > W'(V^2)$, (34) implies

$$\Big(\lambda - \pi\alpha W'\left(R^{11}\right) - (1-\pi)\alpha W'\left(R^{21}\right)\Big)\frac{b'(p_1)}{b(p_1) + p_1b'(p_1)} \;<$$

$$\Big(\lambda - \pi\alpha W'\left(R^{22}\right) - (1-\pi)\alpha W'\left(R^{12}\right)\Big)\frac{b'(p_2)}{b(p_2) + p_2b'(p_2)} \tag{A.7}$$

Since $R^{11} > r^1, R^{12} > r^1, R^{22} > r^2, R^{21} > r^2$, the terms in the brackets are negative from (33). The weights on $W'(r^1)$ and $W'(r^2)$ in (33) may be written:

$$1 - \frac{n_1 \pi d_1 + n_2(1-\pi)d_2}{n_1\pi + n_2(1-\pi)} \gtrless 1 - \frac{n_2 \pi d_2 + n_1(1-\pi)d_1}{n_2\pi + n_1(1-\pi)}, \text{ or } (2\pi - 1)d_2 \gtrless (2\pi - 1)d_1$$

Since $\pi > 1/2$, and $d_2 > d_1$, this implies that the weight on $W'(r^1)$ is greater than that on $W'(r^2)$. Therefore, given that $W'(r^i) > W'(R^{i1}) > W'(R^{i2})$, in order to satisfy (33), $W'(r^1)$, $W'(R^{11})$ and $W'(R^{12})$ would all have to be reduced relative to $W'(r^2)$, $W'(R^{21})$ and $W'(R^{22})$.

The relative sizes of p_1 and p_2 depends on the relative sizes of the terms in brackets in $(A.7)$, or

$$\pi W'\left(R^{22}\right) + (1-\pi)W'\left(R^{12}\right) \gtrless \pi W'\left(R^{11}\right) + (1-\pi)W'\left(R^{21}\right)$$

that is,

$$\pi\Big(W'\left(R^{22}\right) - W'\left(R^{12}\right) + W'\left(R^{21}\right) - W'\left(R^{11}\right)\Big) \gtrless W'\left(R^{21}\right) - W'\left(R^{12}\right)$$

Since both the lefthand and the righthand sides are negative, the direction of the inequality is ambiguous. Therefore, the relative sizes of p_1 and p_2 are ambiguous.

Size of t^1 relative to t^2

Define $p_{ik} = (1-\tau_i)/(1-t_k)$ where i is the wage-type of the parent and k is the wage-type of the child, and R^{ki} as the utility of a type$-k$ child with a type$-i$ donor parent. The government's objective (12) can be rewritten as:

$$\begin{aligned}
\mathcal{W} = n_1 d_1 \Big(\pi W(V^1(c_1, y_1, p_{11}) + (1-\pi)W(V^1(c_1, y_1, p_{12}))\Big) + n_1(1-d_1)W(v^1(c_1, y_1)) \\
+ n_2 d_2 \Big(\pi W(V^2(c_2, y_2, p_{22})) + (1-\pi)W(V^2(c_2, y_2, p_{21}))\Big) + n_2(1-d_2)W(v^2(c_2, y_2) \\
+ n_1 d_1 \pi \alpha W(R^{11}(\overline{c}_1, \overline{y}_1, p_{11})) + (n_1\pi(1-d_1) + n_2(1-\pi)(1-d_2))\alpha W(r^1(\overline{c}_1, \overline{y}_1)) \\
+ n_2 d_2 \pi \alpha W(R^{22}(\overline{c}_2, \overline{y}_2, p_{22})) + (n_2\pi(1-d_2) + n_1(1-\pi)(1-d_1))\alpha W(r^2(\overline{c}_2, \overline{y}_2)) \\
+ n_2 d_2 (1-\pi)\alpha W(R^{12}(\overline{c}_1, \overline{y}_1, p_{21})) + n_1(1-\pi)d_1 \alpha W(R^{21}(\overline{c}_2, \overline{y}_2, p_{12}))
\end{aligned}$$

and the government's budget constraint (13) can be rewritten as (with corresponding multiplier λ)

$$n_1(y_1 - c_1) + (n_1\pi + n_2(1-\pi))(\overline{y}_1 - \overline{c}_1) + n_2(y_2 - c_2) + (n_2\pi + n_1(1-\pi))(\overline{y}_2 - \overline{c}_2)$$

$$+n_1\pi d_1(p_{11}-1)b(p_{11})+n_1(1-\pi)d_1(p_{12}-1)b(p_{12})$$

$$+n_2\pi d_2(p_{22}-1)b(p_{22})+n_2(1-\pi)d_2(p_{21}-1)b(p_{21})=0.$$

Now, consider differentiating the government's maximized Lagrangian with respect to t_2 and evaluating the resulting expression at $t_2=t_1$. We have

$$\frac{\partial\mathcal{L}^*}{\partial t_2}=n_1d_1(1-\pi)\left[W'(V^1)\frac{\partial V^1}{\partial p_{12}}+\alpha W'(R^{21})\frac{\partial R^{21}}{\partial p_{12}}+\lambda\left[(p_{12}-1)b'(p_{12})+b(p_{12})\right]\right]\frac{\partial p_{12}}{\partial t_2}$$

$$+n_2d_2\pi\left[W'(V^2)\frac{\partial V^2}{\partial p_{22}}+\alpha W'(R^{22})\frac{\partial R^{22}}{\partial p_{22}}+\lambda\left[(p_{22}-1)b'(p_{22})+b(p_{22})\right]\right]\frac{\partial p_{22}}{\partial t_2}$$

With a uniform inheritance tax, $t_1=t_2$, so $p_{11}=p_{12}=p_1$ and $p_{22}=p_{21}=p_2$. Using $V_p^i=-b(p_i)-p_ib'(p_i)$ and $R_p^{ki}=b'(p_i)$, the above can be rewritten as:

$$\frac{\partial\mathcal{L}^*}{\partial t_2}=n_1d_1(1-\pi)\Big[W'(V^1)(-b(p_1)-p_1b'(p_1))+\alpha W'(R^{21})b'(p_1)$$

$$+\lambda\left[(p_1-1)b'(p_1)+b(p_1)\right]\Big]\frac{(1-\tau_1)}{(1-t)^2}+n_2d_2\pi\Big[W'(V^2)(-b(p_2)-p_2b'(p_2))$$

$$+\alpha W'(R^{22})b'(p_2)+\lambda\left[(p_2-1)b'(p_2)+b(p_2)\right]\Big]\frac{(1-\tau_2)}{(1-t)^2}$$

Substituting in the first-order conditions on p_1 and p_2 with $\gamma^d=0$ given by $(A.5)$ and $(A.6)$ yields.

$$\frac{\partial\mathcal{L}^*}{\partial t_2}=n_1d_1(1-\pi)\left[\alpha W'(R^{21})b'(p_1)-\left[\pi\alpha W'\left(R^{11}\right)+(1-\pi)\alpha W'\left(R^{21}\right)\right]b'(p_1)\right]\frac{(1-\tau_1)}{(1-t)^2}$$

$$+n_2d_2\pi\left[\alpha W'(R^{22})b'(p_2)-\left[\pi\alpha W'\left(R^{22}\right)+(1-\pi)\alpha W'\left(R^{12}\right)\right]b'(p_2)\right]\frac{(1-\tau_2)}{(1-t)^2}$$

which reduces to

$$\frac{\partial\mathcal{L}^*}{\partial t_2}=n_1d_1(1-\pi)\pi\alpha\left[W'(R^{21})-W'(R^{11})\right]b'(p_1)\frac{(1-\tau_1)}{(1-t)^2}$$

$$+n_2d_2\pi\alpha(1-\pi)\left[W'\left(R^{22}\right)-W'\left(R^{12}\right)\right]b'(p_2)\frac{(1-\tau_2)}{(1-t)^2}$$

Size of $\overline{W}'\left(R^1\right)$ relative to $\overline{W}'\left(R^2\right)$

Suppose first that $\pi = 1$ so parent and child wage-types are perfectly correlated. The equity condition (31) can be written:

$$\frac{\lambda}{\alpha} = d_1 W'\left(r^1 + b(p_1)\right) + (1-d_1)W'\left(r^1\right) = d_2 W'\left(r^2 + b(p_2)\right) + (1-d_2)W'\left(r^2\right) \quad (31')$$

where $R^i = r^i + b(p_i) > r^i$ so $W'(R^i) < \lambda/\alpha < W'(r^i)$. The first-order conditions on p_i becomes:

$$-\left[W'\left(V^i\right) - \lambda\right]\left(b(p_i) + p_i b'(p_i)\right) - \lambda b'(p_i) + \alpha W'\left(R^i\right) b'(p_i) = 0 \qquad i = 1, 2.$$

The first-order conditions can be satisfied at $p_1 = p_2$ provided $V^1 = V^2$ and $R^1 = R^2$ which implies that $r^1 = r^2$ given $p_1 = p_2$. The equity condition can be satisfied for the children in this case only if $d_1 = d_2$ (as is the case with the parents). Now consider a small increase in d_2. The righthand side of (31′) becomes less than the lefthand side (since $W'(R^2) < W'(r^2)$). Given d_1, the derivative of the lefthand side with respect to r^1 is $d_1 W''(r^1+b(p_1))+(1-d_1)W''(r^1) < 0$ and with respect to p_1 is $d_1 W''(r^1+b(p_1))b'(p_1) > 0$. Therefore, to satisfy the equity condition with $d_2 > d_1$ need either r^1 to go up or p_1 to go down which increases R^1. (Note as well that if $d_2 = 1 > d_1$, then $W'(R^1) < W'(R^2) < W'(r^1)$ and by continuity this would hold for a reduction in $d_2 < 1$. Therefore, given $d_2 > d_1$ it must be that $R^1 > R^2$ and $W'(R^1) < W'(R^2)$.

Suppose now that $\pi = 1/2$ so parent and child wage-types are independent. The equity condition (31) becomes:

$$\frac{\lambda}{\alpha} = \frac{n_1 d_1 + n_2 d_2}{n_1 + n_2}\left[\left(\frac{n_1 d_1}{n_1 d_1 + n_2 d_2}\right) W'\left(R^{11}\right) + \left(\frac{n_2 d_2}{n_1 d_1 + n_2 d_2}\right) W'\left(R^{12}\right)\right]$$
$$+ \left(\frac{n_1(1-d_1) + n_2(1-d_2)}{n_1 + n_2}\right) W'\left(r^1\right)$$
$$= \frac{n_1 d_1 + n_2 d_2}{n_1 + n_2}\left[\left(\frac{n_2 d_2}{n_1 d_1 + n_2 d_2}\right) W'\left(R^{22}\right) + \left(\frac{n_1 d_1}{n_1 d_1 + n_2 d_2}\right) W'\left(R^{21}\right)\right]$$
$$+ \left(\frac{n_1(1-d_1) + n_2(1-d_2)}{n_1 + n_2}\right) W'\left(r^2\right)$$

In this case, the weights on $E(\overline{W}'(R^i))$ and $W'(r^i)$ are the same. Therefore, we have the following three possible cases

1. $W'(r^1) = W'(r^2)$ and $E(\overline{W}'(R^1)) = E(\overline{W}'(R^2))$
2. $W'(r^1) < W'(r^2)$ and $E(\overline{W}'(R^1)) > E(\overline{W}'(R^2))$

3. $W'(r^1) > W'(r^2)$ and $E(\overline{W}'(R^1)) < E(\overline{W}'(R^2))$ where we can write

$$E(\overline{W}'(R^1))(n_1d_1 + n_2d_2) = n_1d_1W'(R^{11}) + n_2d_2W'(R^{12})$$

$$E(\overline{W}'(R^2))(n_1d_1 + n_2d_2) = n_2d_2W'(R^{22}) + n_1d_1W'(R^{21})$$

Consider Case 2. $r^1 > r^2$ and therefore, $R^{11} > R^{21}$ and $R^{22} < R^{12}$ which implies from the above that $E(\overline{W}'(R^1)) < E(\overline{W}'(R^2))$ which contradicts the equity condition. Likewise, for Case 3. Therefore, with $\pi = 1/2$ it must be that $r^1 = r^2$ and $E(\overline{W}'(R^1)) = E(\overline{W}'(R^2))$.

References

Archibald, C. and D. Donaldson (1976), 'Non-Paternalism and Externalities', *Canadian Journal of Economics* **9**, 492–507.

Atkinson, A.B. and J.E. Stiglitz (1976), 'The Design of Tax Structure: Direct vs. Indirect Taxation', *Journal of Public Economics* **6**, 55–75.

Brunner, J.K. and S. Pech (2012a), 'Optimal Taxation of Bequests in a Model with Initial Wealth', *Scandinavian Journal of Economics* **114**, 1368-92.

Brunner, J.K. and S. Pech (2012b), 'Optimal Taxation of Wealth Transfers When Bequests are Motivated by Joy of Giving', *B.E. Journal of Economic Analysis and Policy* **2**, article 8.

Cremer, H. and P. Pestieau (2006), 'Wealth Transfer Taxation: A Survey of the Theoretical Literature', in L-A. Gérard-Varet, S-C. Colm and J. Mercier Ythier (eds.), *Handbook of the Economics of Giving, Reciprocity and Altruism*, Volume 2 (Amsterdam: North-Holland), 1107–34.

Cremer, H. and P. Pestieau (2011), 'The Tax Treatment of Intergenerational Wealth Transfer', *CESifo Economic Studies* **57**, 365-401.

Farhi, E. and I. Werning (2010), 'Progressive Estate Taxation', *Quarterly Journal of Economics* **125**, 635–73.

Farhi, E. and I. Werning (2013), 'Estate Taxation with Altruism Heterogeneity', *American Economic Review: Papers and Proceedings* **103**, 489-95.

Frank, R.H. (2008), 'Should Public Policy Respond to Positional Externalities?', *Journal of Public Economics* **92**, 1777–86.

Hammond, P.J. (1987), 'Altruism', in John Eatwell, Murray Milgate and Peter Newman (eds.), *The New Palgrave: A Dictionary of Economics*, First Edition (Basingstoke: Palgrave Macmillan, 1987), 85–7.

Hochman, H.M. and J.D. Rodgers (1969), 'Pareto Optimal Redistribution', *American Economic Review* **59**, 542–57.

Kaplow, L. (1988), 'Tax Policy and Gifts', *American Economic Review Papers and Proceedings* **88** (1998), 283–88.

Kaplow, L. (2001), 'A Framework for Assessing Estate and Gift Taxation', in W.G. Gale, J.R. Hines and J. Slemrod (eds.), *Rethinking Estate and Gift Taxation* (Washington: Brookings Institution), 164–204.

Kaplow, L. (2008), *The Theory of Taxation and Public Economics*, (Princeton: Princeton University Press).

Kopczuk, W. (2013), 'Taxation of Intergenerational Transfers and Wealth', in A. Auerbach, R. Chetty, M. Feldstein, and E. Saez (eds.), *Handbook of Public Economics*, Volume 5 (Amsterdam: Elsevier), 329–90.

Kopczuk, W. (2013), 'Incentive Effects of Inheritances and Optimal Estate Taxation', *American Economic Review: Papers and Proceedings* **103**, 472-77.

Mankiw, G.N. (2006), 'The Estate Tax Debate', Greg Mankiws Blog, June.

McCaffery, E.J. (1994), 'Uneasy Case for Wealth Transfer Taxation', *Yale Law Journal* **104**, 283–365.

Meade Report (1978), *The Structure and Reform of Direct Taxation*, Report of a Committee chaired by Professor J. E. Meade (London: George Allen and Unwin).

Mirrlees, J.A., 'Taxation of Gifts and Bequests', slides for a talk at the Centenary of James Meade Conference, 2007.

Mirrlees, J.A., Optimal Taxation of Saving and Inheritance', slides for a talk at the University of St. Gallen, 2011.

Mirrlees, J.A., S. Adam, T. Besley, R. Blundell, S. Bond, R. Chote, M. Gammie, P. Johnson, G. Myles and J. Poterba (2011), *Tax by Design: the Mirrlees Review* (Oxford: Oxford University Press).

Oswald, A. (1983), 'Altruism, Jealousy and the Theory of Optimal Non-linear Taxation', *Journal of Public Economics* **20** (1983), 77–87.

Pauly, M.V., 'Income Distribution as a Local Public Good', *Journal of Public Economics* **2** (1973), 35–58.

Piketty, T. and E. Saez, 'A Theory of Optimal Inheritance Taxation', CEPR Discussion Paper No. 9241, 2012.

Thurow, L.S., 'The Income Distribution as a Pure Public Good', *Quarterly Journal of Economics* **85** (1971), 327–36.