Tax Treatment of Bequests when Donor Benefits are Discounted*

Robin Boadway
Queen's University
boadwayr@econ.queensu.ca

Katherine Cuff
McMaster University
cuffk@mcmaster.ca

July 2, 2014 – Draft

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

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 giving less than full social weight to 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 inheritance tax, and gives differential linear bequest tax credits to donor parents based on income. The nonlinear income taxes take standard forms. The inheritance tax and bequest tax credits are of indeterminate absolute level and together determine the effective price of net bequests. This effective price plays both an externality-correcting and redistributive role. The optimal effective price will depend on the social weight given to donor benefits and the share of donors of a given skill type.

Key Words: Bequest Tax, Inheritance, Optimal Income Tax
JEL: H21, H23, H24

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 intentional 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 2013). 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 labour supply of recipients, this reduces social welfare as long as labour 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 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 discounting 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 behaviour 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 fully 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

[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).

[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).

(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 discounts the utility parents obtain from bequests, thereby capturing both the extreme cases of 'double-counting' and of zero weight on the joy-of-giving term, as well the various intermediate cases. To be able to separate utility of bequests from utility of own consumption, we assume an additive form of utility function. Heterogeneity of bequest behaviour is captured by assuming that some parents give bequests, while the rest do not.[3]

The government has three policy instruments. It imposes nonlinear income taxes on both parents and children; it chooses a linear tax on inheritances; and, it allows a partial income tax credit on bequests. The latter reflects the fact that in the case when the government does not count the utility of bequests to donors, bequests simply reduce donors' consumption. An income tax credit on bequests 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 (2013), 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 assume the government puts zero social weight on the welfare of children and focus on the tax treatment of parents under two different informational environments, emphasizing both the externality-correcting and redistributive roles of bequest taxation. In Section 5, we incorporate the behaviour and taxation of children in the analysis by assuming the government puts positive social weight on the children's utility. Section 6 concludes.

[3]Brunner and Pech (2012a) incorporate a similar discounting of the utility of bequest, but assume that all individuals are donors.

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 $1 > d_2 > d_1 > 0$ 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.[4] We begin by characterizing the donor and non-donor parents' problems and then turn to the children's problem.

2.1 Parents' Behaviour

Household utility depends on a private consumption good x, on labour 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

[4] Brunner and Pech (2012a) allow parents to receive exogenous bequests from their parents, whose behaviour 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.

quasilinear 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 labour 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 donor 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 $f''(b) < 0$.

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 weight the utility of bequests to donors by less than one, and in the extreme to not count $f(b)$ in the social welfare function. If the utility function took a general 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 behaviour, 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 $b = (1 - t)B$,

$$x + \frac{b}{1 - t} = c + \tau \frac{b}{1 - t} \tag{2}$$

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 (2) 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, labour 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 (2), the problem of donor household with wage rate w can be written as follows, where we adopt b as the donor's choice variable:

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

The necessary condition for this problem is

$$f'(b) = \frac{1-\tau}{1-t} \tag{3}$$

whose solution yields the demand function for net bequests

$$b\Big(\frac{1-\tau}{1-t}\Big), \quad \text{with} \quad b'(\cdot) = \frac{1}{f''(b)} < 0. \tag{4}$$

Comparative statics of (3) yields:

$$b_\tau = \frac{-1}{1-t}b'(\cdot) > 0, \quad b_t = \frac{1-\tau}{(1-t)^2}b'(\cdot) < 0. \tag{5}$$

where subscripts refer to partial derivatives.

It will be useful in what follows to define the argument of (4) for a donor as follows:

$$p \equiv \frac{1-\tau}{1-t}. \tag{6}$$

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

$$\epsilon_p \equiv -\frac{pb'(p)}{b(p)} > 0 \tag{7}$$

where the inequality follows from (4).[5] If $\epsilon_p < 1$, the demand for net bequests is relatively inelastic and vice versa. Differentiating ϵ_p with respect to p yields:

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

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.

Stage 1: Choice of labour supply

When households choose labour 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:

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

The solution to this problem gives

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

which determines y. Thus, income or labour 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.

[5] Using (2), (3) and (5), the elasticity of net bequests with respect to the bequest tax credit and the inheritance tax can be written respectively as

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

2.2 Children's behaviour

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 and denote the nonlinear income tax schedule facing the children by $T_c(y)$. The budget constraint for a child with a donor parent of wage-type i is $x = y - T_c(y) + b(p_i)$, and with a non-donor parent is $x = y - T_c(y)$.

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

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

The first-order condition for this problem gives

$$h'\Big(\frac{y}{w_k}\Big) = (1 - T_c'(y))w_k$$

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

3 The Government Problem

The government chooses separate nonlinear income tax functions 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 $\overline{c}_k$ 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 behaviour 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.[6] We therefore treat p_i, $i = 1, 2$, as control variables for the government.

The government weights the utility of bequests to donors, $f(b)$, by some parameter $\delta \in [0, 1]$. From the point of view of the government, the utility of a donor household

[6]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.

is therefore given by $x + \delta f(b) - h(\ell)$. Consider a type$-i$ donor. Using the budget constraint (2) 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 using (6) as follows:

$$V^i(c, y, p_i; \delta) \equiv c - p_i b(p_i) + \delta f(b(p_i)) - h\Big(\frac{y}{w_i}\Big). \tag{9}$$

Differentiating this, we obtain

$$V_c^i = 1, \quad V_y^i = -\frac{h'(\ell)}{w_i}, \quad V_p^i = -b(p_i) - p_i b'(p_i) + \delta f'(b(p_i))b'(p_i), \quad V_\delta^i = f(b(p_i). \tag{10}$$

Using the expression determining the donor's optimal choice of bequest, (3), V_p^i can be rewritten as

$$V_p^i = -b(p_i) - (1-\delta)p_i b'(p_i) = -b(p_i)\left(1 - (1-\delta)\epsilon_p^i\right) \tag{11}$$

which is negative provided $\epsilon_p^i < 1/(1-\delta)$. 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

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

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:

$$R^{ki}(\bar{c}, \bar{y}, p_i) \equiv \bar{c} + b(p_i) - h\Big(\frac{\bar{y}}{w_k}\Big) \tag{12}$$

where

$$R_{\bar{c}}^{ki} = 1, \quad R_{\bar{y}}^{ki} = -h'(\ell)/w_k, \quad R_p^{ki} = b'(p_i) < 0.$$

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

$$r^k(\overline{c}, \overline{y}) = v^k(\overline{c}, \overline{y}) = \overline{c} - h\Big(\frac{\overline{y}}{w_k}\Big) \tag{13}$$

where

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

As there are no income effects on labour 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 and in utilities of children, subject to a budget constraint and incentive constraints on non-donor parents, donor parents, and children. Below we consider a special case of no social weight on the children's utility and therefore, assume the government discounts the children's utility by $\alpha \in \{0, 1\}$. Specifically, social welfare is:

$$\begin{aligned}
\mathcal{W} = & \sum_{i=1,2} n_i \Big(d_i W\left(V^i(c_i, y_i, p_i; \delta)\right) + (1-d_i) W\left(v^i(c_i, y_i)\right) \Big) \\
& + \; n_1 \pi d_1 \alpha W\left(R^{11}(\overline{c}_1, \overline{y}_1, p_1)\right) + \Big(n_1\pi(1-d_1) + n_2(1-\pi)(1-d_2)\Big) \alpha W\left(r^1(\overline{c}_1, \overline{y}_1)\right) \\
& + \; n_2 \pi d_2 \alpha W\left(R^{22}(\overline{c}_2, \overline{y}_2, p_2)\right) + \Big(n_2\pi(1-d_2) + n_1(1-\pi)(1-d_1)\Big) \alpha W\left(r^2(\overline{c}_2, \overline{y}_2)\right) \\
& + \; n_2(1-\pi) d_2 \alpha W\left(R^{12}(\overline{c}_1, \overline{y}_1, p_2)\right) + n_1(1-\pi) d_1 \alpha W\left(R^{21}(\overline{c}_2, \overline{y}_2, p_1)\right)
\end{aligned} \tag{14}$$

where $W'(\cdot) > 0 > W''(\cdot)$.

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

$$\sum_{i=1,2} \Big(n_i(y_i - c_i + t d_i B_i - \tau_i d_i B_i) + \Big(n_i \pi + n_{-i}(1-\pi) \Big) (\overline{y}_i - \overline{c}_i) \Big) = G,$$

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

$$\begin{aligned}
& n_1(y_1 - c_1) + \Big(n_1\pi + n_2(1-\pi) \Big) (\overline{y}_1 - \overline{c}_1) + n_1 d_1 (p_1 - 1) b(p_1) \\
& \quad n_2(y_2 - c_2) + \Big(n_2\pi + n_1(1-\pi) \Big) (\overline{y}_2 - \overline{c}_2) + n_2 d_2 (p_2 - 1) b(p_2) = G.
\end{aligned} \tag{15}$$

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]

$$c_2 - h\Big(\frac{y_2}{w_2}\Big) \geqslant c_1 - h\Big(\frac{y_1}{w_2}\Big). \tag{16}$$

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

$$c_2 - p_2 b(p_2) + f(b(p_2)) - h\Big(\frac{y_2}{w_2}\Big) \ \geqslant\ c_1 - p_1 b(p_1) + f(b(p_1)) - h\Big(\frac{y_1}{w_2}\Big). \tag{17}$$

Only one of the two incentive constraints for parents, (16) and (17), will generally be binding in a social optimum. If $\tau_2 > \tau_1$, so $p_1 > p_2$, (17) will be slack when (16) 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 $p_1 < p_2$, (16) will be slack when (17) 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

$$\bar{c}_2 - h\Big(\frac{\overline{y}_2}{w_2}\Big) \geqslant \bar{c}_1 - h\Big(\frac{\overline{y}_1}{w_2}\Big). \tag{18}$$

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

[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.

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

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 collects 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. It 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 we allow for less than full weight to be 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 behaviour.

We first 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. We then consider the case when the government cannot observe 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 parents. In a later section, we introduce the welfare of the inheritors.

With observable wage-types, neither incentive constraint will be binding. The government freely chooses c_i, y_i and p_i for $i = 1, 2$ to maximize (14) with zero weight on the utilities of children ($\alpha = 0$) subject to achieving a balanced budget covering the parents' lifetime, that is, subject to the budget constraint (15) excluding the income taxes raised from children. The first-order conditions are:

$$d_i W'(V^i) + (1 - d_i) W'(v^i) - \lambda = 0 \qquad i = 1, 2 \tag{19}$$

$$-d_i W'(V^i)\frac{h'(\ell_i)}{w_i} - (1-d_i)W'(v^i)\frac{h'(\ell_i)}{w_i} + \lambda = 0 \qquad i = 1,2 \tag{20}$$

$$-W'(V^i)\Big(b(p_i) + (1-\delta)p_i b'(p_i)\Big) + \lambda\Big((p_i - 1)b'(p_i) + b(p_i)\Big) = 0 \qquad i = 1,2 \tag{21}$$

Combining (19) and (20) for each wage-type, we obtain

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

Using (8), this implies that the marginal income tax rate for both wage-types is zero and the income tax is non-distorting, as expected.

From (19) for both wage types, we obtain:

$$d_1 W'(V^1) + (1-d_1)W'(v^1) = d_2 W'(V^2) + (1-d_2)W'(v^2) = \lambda. \tag{22}$$

Consider increasing the consumption of type-i parents by one unit. The revenue cost would be n_i and the social benefit would be the increase in social utility of both the donor and non-donor parent of that wage-type. The government optimally increases consumption of the type$-i$ parents until the marginal cost of doing so equals the marginal social benefit as given by (22) for the two wage-types. Consequently, the average marginal social utility of consumption of parents of a given wage-type (averaged over donor and non-donor parents of that wage-type) will be equal for the two wage-types, and equal to the shadow value of government revenue. This equity condition determines the optimal allocation of consumption across the four types of parents (donor/non-donor and high-/low-wage).

Consider now condition (21) determining the price of net bequests p_i. Government policy affecting the effective price of net bequests plays two distinct roles. The first role is to correct a social externality that arises as a result of the government giving a different weight to donors' utility of bequest than donors themselves. The second role is to redistribute between donors and non-donors. We highlight each of these roles in turn.

4.1.1 Social Externality-Correcting Role

Assume first that the government only cares about aggregate social utility, that is, $W'(\cdot) = 1$ and $W''(\cdot) = 0$, and therefore has no aversion to social utility inequality. The equity condition (22) implies that $\lambda = 1$ and condition (21) reduces to[9]

$$(\delta p_i - 1)b'(p_i) = 0 \qquad i = 1,2. \tag{23}$$

[9] This condition also holds if the government can condition the income tax on donor status.

If the government gave full weight to donors' benefits from leaving bequest so $\delta = 1$, then $p = 1$ in the optimum. The objectives of the donors and government would be perfectly aligned and the government would not tax bequests. With $\delta < 1$, donors are leaving larger bequests than is optimal from the point of view of the government and therefore, the government optimally taxes bequests, $p > 1$. Bequest taxation corrects for the externality that donors' bequest behaviour has on the social objective function. The optimal effective net price is $p_i = 1/\delta$, which is the same for both wage-types and is decreasing in the rate at which the government discounts the utility of bequests.

For sufficiently small social weights, bequest taxes may be so high as to reduce net bequests to zero. Define $\hat{p}$ such that $f'(0) = \hat{p}$ where $\hat{p} > 1$.[10] For any $p \geq \hat{p}$, donors do not leave any bequests. Evaluating the left-hand side of (23) at $p = \hat{p}$, the subsequently expression will be non-negative for any $\delta \leq 1/\hat{p}$. Therefore, for social weights greater than $1/f'(0)$, bequests are positive and taxed at a rate that is decreasing in the social weight. For social weights less than $1/f'(0)$, optimal bequests are zero.

4.1.2 Redistributive Role

Now suppose the government is averse to social utility inequality, so $W'(\cdot) > 0 > W''(\cdot)$. Starting from the laissez-faire ($p = 1$), the government would want to redistribute from donor to non-donor parents of a given wage-type if $V_i > v^i$. From (9), this will be the case whenever $\delta > b(1)/f(b(1))$. By the same token, redistribution from non-donor to donor parents would be desirable if $\delta < b(1)/f(b(1))$. If it happened to be the case that $\delta = b(1)/f(b(1))$, then the government would view donors and non-donors to be equally well-off in the laissez-faire.

Consider now the implications of this for optimal bequest taxation. Assume first that there is no behavioural response to taxation so that the amount of bequest is fixed at the laissez-faire level, denoted by $b(1)$, and $b'(p) = 0$. In this case, bequest taxation cannot alter the amount of bequest and therefore, there is no social externality-correcting role for bequest taxation. With no behavioural responses, the first-order condition on p_i, (21), implies that $W'(V^i) = \lambda$, and together with the equity condition, (22), the social marginal utility of donor and non-donor parents of a given wage type must be equal, $W'(V^i) = W'(v^i)$. This requires that $V^i = v^i$, which implies from

[10] The latter follows from the assumptions that donors leave positive bequests without government intervention and that the utility of bequest function is strictly concave.

the definition of the social utility of the donor parents, (9), that the optimal effective price on net bequests is $\delta f(b(1))/b(1)$ to ensure that the net social surplus of bequests is zero at $p = 1$ (i.e., $\delta f(1) - b(1) = 0$).

The intuition is as follows. Given no behavioural responses to taxation, the term $f(b(1))/b(1)$ is fixed and will be greater than one, given that donors leave positive bequests without government intervention. With zero weight on the utility of bequests ($\delta = 0$), p_i must be equal to zero. 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. Bequests are subsidized to fully compensate donors. With positive weight given to the utility of bequest, p_i may be positive or negative. For any $\delta > f(b(1))/b(1)$, bequests will optimally be taxed to redistribute from the donor to the non-donor parents and for any $\delta < f(b(1))/b(1)$, bequests will optimally be subsidized to redistribute from the non-donor to the donor parents. The optimal tax (subsidy) rate is increasing (decreasing) in δ. Bequest taxation redistributes between the donor to the non-donor parents to equalize their social utilities. Of course, although donors and non-donors have the same social utility when there is no behavioural response to taxation, donors receive the full utility of bequests so have higher private utility whenever the government discounts the utility of bequest.

Now consider what happens if bequests are endogenous and respond to bequest taxation. To eliminate the social externality-correcting role for bequest taxation, assume for expositional purposes that donors have the same preferences as the government, (9). The first-order condition on the donor's choice of bequest will be given by $\delta f'(b) - p = 0$. This condition yields $b(p, \delta)$ which is decreasing in p and increasing in δ. Provided $\delta f'(0) - p > 0$, donors leave a positive bequest and will be better off than non-donors. The first-order condition on the government's choice of p_i, (21), reduces to

$$[\lambda - W'(V^i)]b(p_i) + \lambda(p_i - 1)b'(p_i) = 0 \qquad i = 1, 2. \tag{24}$$

The equity condition (22) still applies, and since $V^i > v^i$, it follows that $W'(V^i) < \lambda$ and the first term is positive. Starting from the laissez-faire ($p_i = 1$) given no efficiency rationale for bequest taxation, the government will optimally tax bequests to redistribute from donor to non-donor parents. Any revenue raised by taxing bequests (increasing the effective price above unity) is shared equally among donor and non-donor parents and the social benefit of this redistribution is greater than the social loss, starting from the laissez-faire. This redistribution effect arises solely as a result of allowing for heterogeneity of bequest behaviour. If all parents are donors (as is generally assumed in the literature), then with double-counting of the benefit of bequests, there is no role for bequest taxation. Here we have shown that even with double-counting of the benefits of bequests ($\delta = 1$), bequest taxation will optimally be used to redistribute between individuals of differing bequest behaviour.

4.1.3 Tax or Subsidize Bequests?

Now consider the more general case when the government is averse to social utility inequality and donors' preferences do not align with the governments so both efficiency and redistributive arguments apply. For what follows, we assume that $\delta \in (1/f'(0), 1]$ so there are positive bequests in the optimum.[11] We can rewrite (21) as

$$\frac{W'(V^i)}{\lambda}(1-\delta) + \frac{\Big(1 - W'(V^i)/\lambda\Big)}{\epsilon_p^i} = \frac{p_i - 1}{p_i}. \tag{25}$$

Consider first the case when full weight is given to the utility of bequest, $\delta = 1$. In this special case, the objectives of the donor's and the government are perfectly aligned so there is no social externality-correcting role for bequest taxation. The first-term on the left-hand side is zero. The second-term on the left-hand side captures the redistributive role of bequest taxation. If all parents were donors, then the social utilities of parents would be equalized and would be equal to the marginal cost of public funds so this second term would also be zero and there would be no role for bequest taxation. The optimal effective price of net bequests would be unity. When parents differ in their bequest behaviour and with $\delta = 1$, donors will be better than non-donors from the government's point of view, $V^i > v^i$. It follows from the equity condition that $W'(V^i) < \lambda$. Consequently, the second-term on the left-hand side will be positive which implies that at the optimum $p_i > 1$ and bequests are taxed. The more responsive bequests are to the effective price, i.e., the larger ϵ_p^i, the lower the optimal amount of bequest taxation and the smaller the effective price. Given individuals exhibit different bequest behaviour, there is a role for bequest taxation even when the utility of bequest are given full weight in the social objective function.

Next, consider $\delta < 1$. The first-term of the left-hand side is positive and captures the social externality-correcting role of bequest taxation which gives the government an incentive to tax bequests ($p_i > 1$). As for the second-term, its sign depends on

[11] Evaluating the left-hand side of (21) at $p = \hat{p} \equiv f'(0)$ and using the equity condition (22) yields an expression that is non-negative for any $\delta \leq 1/f'(0)$ and therefore, for such values of the social weight the government will want to set a tax rate such that bequests are zero.

whether donors or non-donors are better-off from the government's point of view. Provided $V^i > v^i$, it follows from the equity condition, (22), that the sign will be positive. For small enough values of δ, however, donors will be viewed by the government as worse off relative to non-donor's, so $V^i < v^i$ and the second term on the left-hand side will be negative. Over this range of δ, bequests may be taxed or subsidized.[12]

Although the first-order condition on p_i may be satisfied at $V^i < v^i$, this will never be the globally optimal solution. To see this, note that if donor parents have lower social utility than non-donor parents of a given wage-type, then the net social benefit of bequest — or net social surplus — is negative, that is $\delta f(b(p)) - pb(p) < 0$. By eliminating bequests, the government unambiguously improves the social utility of donor parents and achieves full utility equality. Another way of stating this is that bequests will only be positive in the optimum when the net social surplus from bequests is positive. To see this, start with the case where p is set such that optimal bequests are zero, and $V^i = v^i$. Since there are no bequests, government revenue from bequest taxation is zero. Then, suppose the government eliminates the bequest tax, so $p = 1$. Government revenues from bequest are still zero. If at $p = 1$, there is a positive net social surplus from bequests, then donor parents will be strictly better off and non-donor parents are just as well off. This results in a Pareto improvement and the government will want to allow for positive bequests. If, instead, the net social surplus is negative at the laissez-faire, then donor parents will be strictly worse off. In this case, the government would want to ensure zero bequests. Therefore, for social weights less than $b(1)/f(b(1))$ bequests will optimally be taxed completely away so there will be no bequests in the optimum. For optima with positive bequests, the government will tax bequests and the social utility of donors will be at least as large as the social utility of non-donor parents of a given wage-type.

4.1.4 Differential Effective Prices of Net Bequests

In this 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. Bequest taxation serves that purpose and the case for imposing differential bequest taxes depends on the relative proportion of donors of each wage-

[12] It will, however, never be optimal to have full tax credits for bequests. To see this, evaluate the left-hand side of (21) at $p_i = 0$ which yields $(\lambda - W'(V^i))\bar{b} - \lambda b'(0) > 0$. The inequality follows from noting that donor parents are choosing $b^i = \bar{b}$, but face zero bequest expenditure so are at least as well of as non-donor parents for any δ and therefore, $W'(V^i) \leq \lambda$ from the equity condition, (22).

type. To see this, suppose $p_1 = p_2 = p > 0$. Then both types of donor parents are choosing the same bequest, and it follows from (21) that donors of both wage-type are equally well-off, $V^1 = V^2$. This in turn implies that non-donor parents are also equally well-off ($v^1 = v^2$). For the equity condition (22) to be satisfied, it must be that the share of donors of each wage-type is the same, $d_2 = d_1$. Thus, a uniform bequest tax policy is only optimal if there is the same proportion of donors of each wage-type. If $d_2 \neq d_1$, then social utilities for both donor and non-donor parents cannot be equalized across wage types given the expression for λ from (22). Therefore, it must be that the effective prices are differentiated by the wage-type of the donor. We can show that the government will optimally impose a lower effective price on those wage-types with a larger share of donors.

To first gain some intuition for this result, consider the following thought exercise. Start from the situation in which the government is imposing the same effective price which implies from the first-order condition on p_i, (21), that $V^1 = V^2$. Now consider increasing the effective price p_i to finance an increase in consumption to all type-i parents. The social gain of a marginal increase in p_i is the additional revenue it generates that benefits all type$-i$ parents and is given by λn_i. The social cost of a marginal increase in p_i is given by $n_i d_i W'(V^i)V_p^i$. Therefore, the increase in p_i needed to finance a uniform increase in c_i is given by $-d_i W'(V^i)V_p^i/\lambda$.[13] Given that $d_2 > d_1$ this expression will be larger for p_2 than for p_1, starting from $p_1 = p_2$. The larger the share of donor parents, the higher the social loss per unit of revenue gained and therefore, the government will optimally set $p_1 > p_2$.

This point can be verified by examining the first-order condition on p_i, (21), more closely. Rewriting (21),

$$\frac{W'(V^i)}{\lambda} - 1 = \frac{(\delta p_i - 1)b'(p_i)}{b(p_i) + (1-\delta)p_i b'(p_i)}. \tag{26}$$

Given $V^i > v^i$, $W'(V^i) < \lambda$ from (22) and the left-hand side of (26) will negative. Assuming $V_p^i < 0$, the denominator of the right-hand side of (26) be positive and since $b'(p_i) < 0$, it must be that $(\delta p_i - 1) > 0$ or $p_i > 1/\delta$. The optimal effective price is greater than the price that just corrects the social externality, because bequest taxation is also being used to redistribute from donor to non-donor parents. This

[13]Another way to state this is that the consumption needed to compensate for a marginal increase in p_i starting from $p_1 = p_2$ to keep social welfare constant is higher for the type$-i$ donor parent with the greater proportion of donors. Therefore, social welfare can be improved by increasing p_1 and reducing p_2 given that $d_2 > d_1$.

also implies that the right-hand side of (26) is decreasing in p_i.[14] Now consider the equity condition (22). It can be shown that in the optimum it must be that $W'(V^2) > W'(V^1)$.[15] Together with the above analysis, this implies that in the full-information optimum $p_1 > p_2 > 1$.

The following summarizes the results we have obtained in the full-information benchmark case when the heirs' welfare is not given any weight.

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 effective price of net bequests plays two distinct roles: it corrects for a social externality when the social weight given to the utility of bequest is less than unity and it redistributes between donor and non-donor parents for all values of the social weight. Both of these roles call for taxing bequests.*

3. *If the weight given to donors' utility of bequests is close enough to zero, the government will set the price of net bequests high enough to eliminate all bequests.*

[14]To see this, dropping subscripts

$$\frac{d}{dp}\Big(\frac{(\delta p-1)b'(p)}{b(p)+(1-\delta)pb'(p)}\Big) = \frac{1}{(b+(1-\delta)pb')^2}\Big((\delta p-1)[b''b-(2-\delta)(b')^2]+\delta pb'(b+(1-\delta)pb')\Big) < 0$$

where the inequality follows from $(\delta p_i - 1) > 0$.

[15]Assume $d_2 = 1 > d_1 > 0$. From (22), $W'(V^2) = d_1 W'(V^1) + (1-d_1)W'(v^1)$ which implies $V^1 > V^2 > v^1$. Now, consider decreasing d_2 slightly. By continuity, we still have $V^1 > V^2$. Assume next that $d_2 = d_1 < 1$. In this case, it follows from (21) and (22) that there will be uniform bequest taxation. Further, $V^1 = V^2 > v^1 = v^2$. Using the relationship that $V^i = v^i + \delta f(b(p_i)) - p_i b(p_i)$, consider a small increase in d_2. This will reduce the average social marginal utility of the high wage-types and (22) will no longer bind. For (22) to remain binding either v^2 must decrease or v^1 must increase. Either of these changes to keep (22) binding will reduce V^2 relative to V^1. This continues to hold for further increases in d_2. Therefore, when $d_2 > d_1$ it must be that $V^2 < V^1$.

4. In an interior solution with positive bequests,

(a) the net social surplus from bequests will be non-negative so that donors will be at least as well off (from the government's point of view) as non-donors parents of a given wage type,

(b) the optimal value of p_i will be greater than one and smaller for the wage-type with the highest proportion of donors, and

(c) the social utility of the donor parents with the highest proportion of donors will be lower than the social utility of donors parents of the other wage-type.

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 ensure that high-wage parents, regardless of donor status, are better off (from the government's perspective) than their low-wage counterparts.[16] The government maximizes (14) (with $\alpha = 0$) subject to government's budget constraint (2) (again excluding the income taxes raised from children) and the incentive constraints on both non-donor parents given by (16) and donor parents given by (17) with Lagrange multipliers γ and γ^d, respectively. The government cannot implement lump-sum redistribution between wage-types, leading to the bequest taxation being used both to redistribute between donors and non-donors and to redistribute between wage-types. The first-order conditions can be written as follows, using (10):

$$n_1 d_1 W'(V^1) + n_1(1-d_1)W'(v^1) - (\gamma + \gamma^d) - \lambda n_1 = 0 \tag{27}$$

$$-n_1 d_1 W'(V^1)\frac{h'(\ell_1)}{w_1} - n_1(1-d_1)W'(v^1)\frac{h'(\ell_1)}{w_1} + (\gamma+\gamma^d)\frac{h'(\hat{\ell}_2)}{w_2} + \lambda n_1 = 0 \tag{28}$$

$$n_2 d_2 W'(V^2) + n_2(1-d_2)W'(v^2) + (\gamma + \gamma^d) - \lambda n_2 = 0 \tag{29}$$

[16] For non-donor parents this follows directly from the incentive constraint (16) since high-wage mimickers have to supply less labour to earn y_1. If this constraint is binding, then $p_1 > p_2$ and it follows immediately that $V^2 > V^1$. If instead the incentive constraint on donor parents (17) binds, then $p_1 < p_2$. Adding and subtracting $\delta f(b(p_i))$ to each side of (17) yields, $V^2 + (1-\delta)f(b(p_2)) = \hat{V}^1 + (1-\delta f(b(p_1))$ where $\hat{V}^1$ is the social utility of a mimicking high-wage donor and will be greater than V^1 since the high-wage mimicker has to provide less labour. Given that $f(b)$ is decreasing in the price and since $p_1 < p_2$, it again follows that high-wage donor parents will have higher social utility than low-wage donor parents.

$$-n_2 d_2 W'(V^2)\frac{h'(\ell_2)}{w_2} - n_2(1-d_2)W'(v^2)\frac{h'(\ell_2)}{w_2} - (\gamma+\gamma^d)\frac{h'(\ell_2)}{w_2} + \lambda n_2 = 0 \quad (30)$$

$$-n_1 d_1 W'(V^1)(b(p_1)+(1-\delta)p_1 b'(p_1))+\lambda n_1 d_1\Big((p_1-1)b'(p_1)+b(p_1)\Big)+\gamma^d b(p_1) = 0 \quad (31)$$

$$-n_2 d_2 W'(V^2)(b(p_2)+(1-\delta)p_2 b'(p_2))+\lambda n_2 d_2\Big((p_2-1)b'(p_2)+b(p_2)\Big)-\gamma^d b(p_2) = 0 \quad (32)$$

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

Eqs. (27)–(30) characterize the optimal income tax structure. These results are standard. For example, from (29) and (30), 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 (27) and (29), 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):

$$n_1 d_1 W'(V^1) + n_1(1-d_1)W'(v^1) + n_2 d_2 W'(V^2) + n_2(1-d_2)W'(v^2) = \lambda \quad (33)$$

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

Note the difference between (33) and (22). 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 bequest taxation to redistribute from the non-donors to the donors and to correct the social externality. 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 bequest taxation.

Eqs. (31) and (32) 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, (16) or (17), will be binding depending on whether or not $p_1 < p_2$ in the optimum. In the full-information case where the incentive constraints did not apply, we found that $p_1 > p_2$ unambiguously. With imperfect information, the relative size of p_1 and p_2 is no longer clear. That is because they serve not only to redistribute from donors to non-donors and correct the social externality, but they also influence redistribution between wage-types, as the latter terms in (31) and (32) involving γ^d indicate. We can show, however, that the incentive constraint on donors must be binding at the optimum and therefore $p_1 < p_2$ when there is imperfect information.

To see this, suppose the incentive constraint on donors (17) is slack and therefore, $p_1 > p_2$. The first-order conditions on the effective prices (31) and (32) are the same as in the full-information case (see eq. (21)). Only the income tax rates are distorted. With positive net social surplus from bequests, donors parents will be better off than non-donor parents of a given wage-type and it follows from eq. (21) that $p_1 < p_2$, a contradiction.[17] Therefore, the incentive constraint on the donors must be binding in the imperfect information case and $p_1 < p_2$. Bequest taxation helps to redistribute between parents of differing wage-types when parents differ in their bequest behaviour.

When parents exhibit different bequest behaviour, it is optimal to impose a higher effective price of net bequests on high-wage donor parents. This is no longer the case if all parents were donors. In this special case, it was shown in the full information benchmark, that bequests should be taxed (at the same rate for both wage-types) to correct the social externality when $\delta < 1$, but otherwise not taxed if $\delta = 1$. These results still hold with imperfect information. To see the latter result, suppose first that $d_1 = d_2 = 1$ and $\delta = 1$. From the incentive constraint on these donor parents, (17), it follows that $V^2 > V^1$ and from (33), it then follows that $W'(V^2) < \lambda < W'(V^1)$. Solving for the multiplier on the donor parent incentive constraint, γ^d from eqs. (31) and (32) and setting equal yields a condition that can only be satisfied for $p_1 = p_2 = 1$.[18] Now suppose $\delta < 1$, but all parents continue to be donors. Following a similar exercise, it can be shown that the condition can only be satisfied for $p_i = 1/\delta$.[19] Given that all parents are donors, there is no role for bequest taxation

[17] With $V^i > v^i$ and $v^2 > v^1$, the social utility of the high-wage donors must be higher than the other three types of parents and from (33), $W'(V^2) < \lambda$. It is possible that either $W'(V^1)$ is greater or less than λ. In the latter case, the analysis is identical to the case with full information which implies that $p_1 < p_2$. In the former case, the first-order condition on p_i as given by (26) is satisfied at $p_2 > 1/\delta$ and $p_1 < 1/\delta$ which again leads to $p_1 < p_2$.

[18] Solving eqs. (31) and (32) for γ^d and setting equal yields: $n_1(W'(V^1) - \lambda) - \lambda n_1\big((p_1 - 1)b'(p_1)/b(p_1)\big) = n_2(\lambda - W'(V^2)) + \lambda n_2\big((p_2 - 1)b'(p_2)/b(p_2)\big)$. Substituting in for λ using (33) and noting that $n_1 + n_2 = 1$, the expression becomes $n_1(1 - n_1)(W'(V^1) - W'(V^2)) - \lambda n_1\big((p_1 - 1)b'(p_1)/b(p_1)\big) = n_2 n_1(W'(V^1) - W'(V^2)) + \lambda n_2\big((p_2 - 1)b'(p_2)/b(p_2)\big)$ which reduces to $-n_1\big((p_1 - 1)b'(p_1)/b(p_1)\big) = n_2\big((p_2 - 1)b'(p_2)/b(p_2)\big)$ and can only be satisfied at $p_1 = p_2 = 1$.

[19] Solving eqs. (31) and (32) for γ^d and setting equal yields: $n_1(W'(V^1) - \lambda) + n_1 W'(V^1)(1 - \delta)p_1 b'(p_1)/b(p_1) - \lambda n_1\big((p_1 - 1)b'(p_1)/b(p_1)\big) = n_2(\lambda - W'(V^2)) - n_2 W'(V^2)(1 - \delta)p_2 b'(p_2)/b(p_2) + \lambda n_2\big((p_2 - 1)b'(p_2)/b(p_2)\big)$. Again, substituting in for λ using (33) and noting that $n_1 + n_2 = 1$, the expression becomes $n_1\big(W'(V^1)(1 - \delta)p_1 - \lambda(p_1 - 1)\big)b'(p_1)/b(p_1) = -n_2\big(W'(V^2)(1 - \delta)p_2 - \lambda(p_2 -$

other than to correct the social externality that arises when the private utility of bequest differs from the social utility of bequest. The non-linear income tax is a sufficient policy instrument to redistribute between parents of differing wage-types when there is imperfect information. If, however, parents differ in their bequest behaviour, then it is optimal to tax the bequest of high-wage donors more than low-wage donors. This result holds regardless of whether there is a higher or lower proportion of high-wage donor parents provided $d_i < 1$ for $i = 1, 2$.

With full information and zero net bequests, the government could equalize the social utilities of donor and non-donor parents and therefore, for sufficiently small social weight on donors' utility of bequests δ, the government would want to induce net bequests to be zero. With imperfect information, this is no longer necessarily the case. By the incentive constraints, $V^2 > V^1$, even with zero net bequests. Therefore, the value of the social weight below which the government will want to drive net bequests to zero will now depend on the ratio of the marginal cost of public funds, λ, to the social marginal utility of the donor parent and maybe be positive or negative.[20]

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. *Without heterogeneous bequest behaviour, uniform bequest taxation can be used to correct the social externality arising when the social weight given to the utility of bequest is less than unity.*

3. *With heterogenous bequest behaviour, bequest taxation can also serve to redistribute between parents of differing wage-types. The optimal value of p_i will be higher for the high-wage donor parents.*

$1))b'(p_2)/b(p_2)$. This condition is satisfied at $p_i = 1/\delta$.

[20]Evaluating the left-hand side of (31) and (32) at $p = \hat{p}$ yields an expression that will be non-negative for any $\delta \leq \bar{\delta} \equiv (1 - \lambda/W'(V^i)) + (\lambda/W'(V^i))(1/f'(0))$ where from (33) with zero net bequests, $W'(V^2) < \lambda < W'(V^1)$.

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 with $\alpha = 1$. 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. The first-order conditions on c_i and y_i are given by (27)–(30). 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 (14), subject only to the resource constraint (2). The first-order conditions on c_i and y_i are (19) and (20) as in the previous full-information problem without zero social weight on the children's utility, which yield the equity condition (22) 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

$$\begin{aligned}\lambda &= \frac{n^{1d}}{n^1}\left[\left(\frac{n^{11}}{n^{1d}}\right)W'\left(R^{11}\right) + \left(1-\frac{n^{11}}{n^{1d}}\right)W'\left(R^{12}\right)\right] + \left(1-\frac{n^{1d}}{n^1}\right)W'\left(r^1\right)\\ &= \frac{n^{2d}}{n^2}\left[\left(\frac{n^{22}}{n^{2d}}\right)W'\left(R^{22}\right) + \left(1-\frac{n^{22}}{n^{2d}}\right)W'\left(R^{21}\right)\right] + \left(1-\frac{n^{2d}}{n^2}\right)W'\left(r^2\right) \qquad (34)\end{aligned}$$

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 $k \neq i = 1,2$. Analogous to (22) 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 (34) that

$$W'(r^i) > \lambda > \left(\frac{n^{kk}}{n^{kd}}\right) W'\left(R^{kk}\right) + \left(1 - \frac{n^{kk}}{n^{kd}}\right) W'\left(R^{ki}\right) \tag{35}$$

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

Consider now 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 for $i \neq k = 1,2$ as follows:

$$-W'\left(V^i\right)\Big(b(p_i) + (1-\delta)p_i b'(p_i)\Big) + \lambda\Big((p_i - 1)b'(p_i) + b(p_i)\Big)$$
$$+\Big[\pi W'\left(R^{ii}\right) + (1-\pi)W'\left(R^{ki}\right)\Big]b'(p_i) = 0 \tag{36}$$

Relative to the full-information case without children, where the condition (21) applies, there is an additional cost of increasing p_i since the bequests now benefit the children. This is reflected in the last term in (36), which is negative since $b'(p_i) < 0$. There is now a positive externality of bequests as a result of having a positive social weight on children's utility. Donors do not take into account the benefit of their bequests to their children, apart from the benefit $b(p_i)$ they obtain themselves.

To highlight how a positive social weight on children's utility changes the optimal taxation of bequest, consider again the pure efficiency setting in which the government has zero aversion to utility inequality so $W' = 1$ and $W'' = 0$ for both parents and children. The government cares only about aggregate social utilities. From the equity conditions (22) and (34), $\lambda = 1$ and the first-order conditions on p_i (36) reduce to:

$$\delta p_i b'(p_i) = 0 \tag{37}$$

Recall that with zero weight on the children's utility, the government would not distort the effective price of net bequest from the laissez-faire when $\delta = 1$. Here, starting from the laissez-faire, the government will want to fully subsidize bequests by setting $p_i = 0$ for any $\delta > 0$.[21] Unlike the case with no social weight on the children, the

[21] This corresponds with the case considered in the recent literature, such as Kaplow (2001), Farhi and Werning (2010) and Piketty and Saez (2013).

government will not want to reduce net bequests to zero unless it happened to be the case that $\delta = 0$.[22]

Now suppose the government is adverse to social utility inequality. With $\alpha = 0$, the government will never want to drive net bequests to zero provided it places a positive weight on the utility of bequest and bequests may optimally be taxed or subsidized.[23] To see the latter result, we can evaluate the left-hand side of (36) at $p_i = 1$ assuming $\delta > 0$ to obtain the following expression

$$\left(\lambda - W'\left(V^i\right)\right) - W'\left(V^i\right)(1-\delta)b'(1) + \left[\pi W'\left(R^{ii}\right) + (1-\pi)W'\left(R^{ki}\right)\right]b'(1) \quad (38)$$

The second and third term capture the social externalities. Starting from the laissez-faire, the government has an incentive to increase the effective price on net bequests, or tax bequests, given $\delta < 1$ and to reduce the effective price on net bequests, or subsidize bequests, given $\alpha = 1$. The first term reflects the redistributive role of bequest taxation. Without counting the welfare of children, the social utility of donors must be at least as large as the social utility of the non-donors in any optimum with positive bequests and from the equity condition on the parents (22) the first term will be positive. Bequests should be taxed to redistribute from donor to non-donor parents. With a positive weight on children's well-being in the social welfare function, this may no longer be the case. The social benefit to heirs may outweigh the social loss to parents if δ is low enough. In this case, there may be positive bequests even though the social utility of donors is lower than the social utility of non-donor parents. The sign of the first term will then be negative given the equity condition (22), calling for bequests to be subsidized as a means of redistributing from non-donor to donor parents.

Although bequests may be subsidized at the optimum, it is still the case that a zero effective price for net bequests is never optimal. To see this, suppose $p_i = 0$ then both type of donor parents leave $\bar{b}$, so $V^i > v^i$, $R^{11} = R^{12} > r^1$, $R^{22} = R^{21} > r^2$ and it follows from (34) that $W'(R^{kk}) < \lambda < W'(r^k)$ for $k = 1, 2$. Therefore, the left-hand side of (36) with $p_i = 0$ reduces to

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

[22] Evaluating (37) at $p = \hat{p}$ which is the price that ensures donors optimally choose zero net bequests yields an expression that is negative for any $\delta > 0$ and equal to zero for $\delta = 0$.

[23] Suppose $p = \hat{p}$ so net bequests are zero. From the two equity conditions, (22) and (34), the social utility of parents and of children are equalized across wage-types. The first-order condition on p_i, (36), evaluated at $p = \hat{p}$ reduces to $b'(\hat{p})\delta\hat{p}$ which is negative for any $\delta > 0$.

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 (36) at $\alpha = 0$. As shown above, the first two terms in (36) will be equal to zero at $p_1 > p_2$. Now consider the last terms in (36). With $p_1 > p_2$, low-wage donor parents are leaving smaller bequests than high-wage donors. Therefore, it must be that the average social marginal utility of high-wage donor heirs is lower than the average social marginal utility of low-wage donor heirs. Thus, starting from the optimum with $\alpha = 0$, the government will want to reduce both p_1 and p_2, but will have an incentive to reduce p_1 by a greater amount so as to increase the bequest received by low-wage donor heirs.[24]

The intuition behind this ambiguity is as follows. In the full-information case without heirs, we obtained that $p_1 > p_2$. In that case, bequest taxation served to redistribute between donor and non-donor parents of different wage-types. Given that a higher proportion of the high-skilled were donors compared with the low-skilled, a higher bequest tax on the low-skilled was relatively more effective at redistributing from donors to non-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 $p_1 < p_2$ in order to favor low-skilled inheritors counters the opposite desire to redistribute between non-donor and donor parents, and either influence can dominate.

A more formal demonstration that the relative sizes of p_1 and p_2 are ambiguous is as follows. Rewrite (36) as

$$\frac{W'(V^i)}{\lambda} - 1 = \frac{\Big(\delta p_i - 1 + (1/\lambda)\big(\pi W'\left(R^{ii}\right) + (1-\pi)W'\left(R^{ki}\right)\big)\Big)b'(p_i)}{b(p_i) + (1-\delta)p_i b'(p_i)}. \tag{39}$$

Suppose donors are better off than non-donors. It follows that the left-hand side of the above expression is negative which means given $V_p^i < 0$ that the denominator on the right-hand side is positive and the expression in the brackets in the numerator

[24]The term multiplying the average social marginal utility of the heirs of a type$-i$ donor parent is decreasing in p_i, so $b'(p_1) < b'(p_2) < 0$ at $p_1 > p_2$.

must be positive. Therefore, the right-hand side expression can again be shown to be decreasing in p. So if $p_1 = p_2$, then $V^1 = V^2$. Further if $V^2 > V^1$, then $W'(V^2) < W'(V^1)$ and $p_2 > p_1$ or vice versa. As in our previous analysis, if donors are better off than non-donors, we can again use the equity condition on parents to show $V^2 < V^1$ and thus $p_2 < p_1$ as without heirs. But, if donors are worse off than non-donors, the equity condition instead implies that $V^2 > V^1$ given $d_2 > d_1$ and how the right-hand side changes with p is now ambiguous. This latter case is now possible with a positive weight on the heirs.

5.2 Imperfect Information

Now assume the government cannot condition the consumption-income bundles on the type of individual so the incentive constraint (18) 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

$$\begin{aligned}\lambda = {} & \frac{n_1 \pi d_1}{n_1+n_2} W'\left(R^{11}\right) + \frac{n_1(1-\pi)d_1}{n_1+n_2} W'\left(R^{21}\right) \\ & + \frac{n_1\pi(1-d_1)+n_2(1-\pi)(1-d_2)}{n_1+n_2} W'\left(r^1\right) + \frac{n_2(1-\pi)d_2}{n_1+n_2} W'\left(R^{12}\right) \\ & + \frac{n_2\pi d_2}{n_1+n_2} W'\left(R^{22}\right) + \frac{n_2\pi(1-d_2)+n_1(1-\pi)(1-d_1)}{n_1+n_2} W'\left(r^2\right).\end{aligned}$$

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 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, bequest taxation 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 $p_2 < p_1$, 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 $p_{ik} = (1 - \tau_i)/(1 - t_k)$.[25] 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. To see this, note that the effective prices of net bequests are related as

[25] 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.

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.This holds regardless of the assumed information on the part of the government. 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? 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, (34), which can be rewritten as:

$$\lambda = \frac{n^{1d}}{n^1}\overline{W}'\left(R^1\right) + \left(1 - \frac{n^{1d}}{n^1}\right) W'\left(r^1\right) = \frac{n^{2d}}{n^2}\overline{W}'\left(R^2\right) + \left(1 - \frac{n^{2d}}{n^2}\right) W'\left(r^2\right) \quad (40)$$

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 $\overline{W}'\left(R^1\right) = \overline{W}'\left(R^2\right)$ 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. Given that $r^1 > r^2$, we have $R^{11} > R^{21}$ and $R^{12} > R^{22}$. Therefore, by increasing t_1 relative to t_2 the bequests received by children with parents of opposite skill type increases and inequality between recipient children of the same donor parent is reduced. 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.*

3. *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.*

4. *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.*

6 Concluding Comments

We have studied the optimal linear tax treatment of bequests when the government does not give full 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.

With zero welfare weight on children, the social utility of donors will always be higher than the social utility of donors. Otherwise, the net social surplus from bequests will be negative and the government will drive net bequests to zero. Bequest taxation plays two distinct roles. It attempts to correct for the social externality arising both from the government weighting the utility of bequest differently from donors and from donors not taking into account the benefit of bequests to inheriting children. At the same time, bequest taxation is used to redistribute between donor and non-donor parents and between inheriting and non-inheriting children. Thus, with zero welfare weight on children, taxing bequest will be both efficiency and equity-enhancing. With positive welfare weight on children, there will be an 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 with high social discounting of the joy-of-giving 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 2013). 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 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.

References

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

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

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

[4] 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.

[5] 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.

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

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

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

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

[10] 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.

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

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

[13] 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.

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

[15] 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.

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

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

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

[19] 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).

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

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

[22] 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).

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

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

[25] Piketty, T. and E. Saez, 'A Theory of Optimal Inheritance Taxation', *Econometrica* **81** (2013), 1851–1886.

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

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

Using the properties of R^{ki} and r^k in (12) and (13):

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

$$+ n_2(1-\pi)d_2W'\left(R^{12}\right) - \phi - \lambda(n_1\pi + n_2(1-\pi)) = 0 \qquad (A.1)$$

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

$$- n_2(1-\pi)d_2W'\left(R^{12}\right)\frac{h'(\bar{\ell}_1)}{w_1} + \phi\frac{h'(\widehat{\bar{\ell}_2})}{w_2} + \lambda(n_1\pi + n_2(1-\pi)) = 0 \qquad (A.2)$$

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

$$+ n_1(1-\pi)d_1W'\left(R^{21}\right) + \phi - \lambda(n_2\pi + n_1(1-\pi)) = 0 \qquad (A.3)$$

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

$$- n_1(1-\pi)d_1W'\left(R^{21}\right)\frac{h'(\bar{\ell}_2)}{w_2} - \phi\frac{h'(\bar{\ell}_2)}{w_2} + \lambda(n_2\pi + n_1(1-\pi)) = 0 \qquad (A.4)$$

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

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

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

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

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

B 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 (14) can be rewritten as:

$$\mathcal{W} = n_1d_1\Big(\pi W(V^1(c_1,y_1,p_{11}) + (1-\pi)W(V^1(c_1,y_1,p_{12}))\Big) + n_1(1-d_1)W(v^1(c_1,y_1))$$

$$+n_2d_2\Big(\pi W(V^2(c_2,y_2,p_{22}))+(1-\pi)W(V^2(c_2,y_2,p_{21}))\Big)+n_2(1-d_2)W(v^2(c_2,y_2)$$
$$+n_1d_1\pi W(R^{11}(\overline{c}_1,\overline{y}_1,p_{11}))+(n_1\pi(1-d_1)+n_2(1-\pi)(1-d_2))W(r^1(\overline{c}_1,\overline{y}_1))$$
$$+n_2d_2\pi W(R^{22}(\overline{c}_2,\overline{y}_2,p_{22}))+(n_2\pi(1-d_2)+n_1(1-\pi)(1-d_1))W(r^2(\overline{c}_2,\overline{y}_2))$$
$$+n_2d_2(1-\pi)W(R^{12}(\overline{c}_1,\overline{y}_1,p_{21}))+n_1(1-\pi)d_1W(R^{21}(\overline{c}_2,\overline{y}_2,p_{12}))$$

and the government's budget constraint (2) 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

$$\frac{\partial\mathcal{L}^*}{\partial t_2}=n_1d_1(1-\pi)\left[W'(V^1)\frac{\partial V^1}{\partial p_{12}}+W'(R^{21})\frac{\partial R^{21}}{\partial p_{12}}+\lambda\left[(p_{12}-1)b'(p_{12})+b(p_{12})\right]\right]\frac{\partial p_{12}}{\partial t_2}$$
$$+n_2d_2\pi\left[W'(V^2)\frac{\partial V^2}{\partial p_{22}}+W'(R^{22})\frac{\partial R^{22}}{\partial p_{22}}+\lambda\left[(p_{22}-1)b'(p_{22})+b(p_{22})\right]\right]\frac{\partial p_{22}}{\partial t_2}$$

With a uniform inheritance tax, $t_1=t_2$, so $p_{11}=p_{12}=p_1$ and $p_{22}=p_{21}=p_2$. Using $V_p^i=-b(p_i)-(1-\delta)p_ib'(p_i)$ and $R_p^{ki}=b'(p_i)$, the above can be rewritten as:

$$\frac{\partial\mathcal{L}^*}{\partial t_2}=n_1d_1(1-\pi)\Big[W'(V^1)(-b(p_1)-(1-\delta)p_1b'(p_1))+W'(R^{21})b'(p_1)$$
$$+\lambda\left[(p_1-1)b'(p_1)+b(p_1)\right]\Big]\frac{(1-\tau_1)}{(1-t)^2}+n_2d_2\pi\Big[W'(V^2)(-b(p_2)-(1-\delta)p_2b'(p_2))$$
$$+W'(R^{22})b'(p_2)+\lambda\left[(p_2-1)b'(p_2)+b(p_2)\right]\Big]\frac{(1-\tau_2)}{(1-t)^2}$$

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

$$\frac{\partial\mathcal{L}^*}{\partial t_2}=n_1d_1(1-\pi)\left[W'(R^{21})b'(p_1)-\left[\pi W'\left(R^{11}\right)+(1-\pi)W'\left(R^{21}\right)\right]b'(p_1)\right]\frac{(1-\tau_1)}{(1-t)^2}$$
$$+n_2d_2\pi\left[W'(R^{22})b'(p_2)-\left[\pi W'\left(R^{22}\right)+(1-\pi)W'\left(R^{12}\right)\right]b'(p_2)\right]\frac{(1-\tau_2)}{(1-t)^2}$$

which reduces to

$$\frac{\partial\mathcal{L}^*}{\partial t_2}=n_1d_1(1-\pi)\pi\left[W'(R^{21})-W'(R^{11})\right]b'(p_1)\frac{(1-\tau_1)}{(1-t)^2}$$
$$+n_2d_2\pi(1-\pi)\left[W'\left(R^{22}\right)-W'\left(R^{12}\right)\right]b'(p_2)\frac{(1-\tau_2)}{(1-t)^2}$$

C 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 (36) can be written:

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

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

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

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 (36) becomes:

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

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 $E(\overline{W}'(R^1)) = E(\overline{W}'(R^2))$ and Case 1 holds.