

Tax treatment of bequests when donor benefits are discounted

Robin Boadway¹ · Katherine Cuff²

Published online: 11 June 2015
© Springer Science+Business Media New York 2015

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 analyse 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 externality-correcting and redistributive roles. The optimal effective price will depend on the social weight given to donor benefits and the share of donors of a given skill type.

Keywords Bequest tax · Inheritance · Optimal income tax

JEL Classification H21 · H23 · H24

✉ Katherine Cuff
cuffk@mcmaster.ca

Robin Boadway
boadwayr@econ.queensu.ca

¹ Department of Economics, Queen's University, Kingston, ON K7L 3N6, Canada

² Department of Economics, McMaster University, Hamilton, ON L8S 4M4, Canada

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 not different from any other market transaction and could be treated as such for taxation purposes.

The normative basis for the taxation of bequests is controversial. From the point of view of recipients, inheritances represent windfall sources of income that not only provide benefits to them, but do so in a very unequal way. Taxation both serves an equalizing objective and contributes to equality of opportunity. How to treat the donors is another matter, at least if 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 recipients from bequests is already counted in social welfare, it is unnecessary to include it again in the guise of benefits to 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 2012a; 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 an 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 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 towards 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 analysing 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

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 \(1973\)](#).

[Kopczuk \(2013\)](#) emphasizes, bequest behaviour is very heterogeneous. Failure to take into account of that heterogeneity entails discriminating in favour 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 \(2012a\)](#) 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 Sect. 3. In Sect. 4, we consider the social optimum when the government can observe wage-types of both parents and children, and highlight the externality-correcting and redistributive roles of bequest taxation by restricting the government to use the same policy instruments as in the imperfect-information case. In Sect. 5, we consider the imperfect-information environment. Section 6 discusses some extensions, and Sect. 7 concludes.

³ [Brunner and Pech \(2012b\)](#) 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 that 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 quasilinear forms, respectively:

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

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 $f'''(b) \leq 0$ 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

⁴ Brunner and Pech (2012a, b) allow parents to receive exogenous bequests from their parents, whose behaviour is not explicitly modelled. 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.

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 $f''(b) < 0$.

Four 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 weigh the utility of bequests to donors by less than unity 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 $f(\cdot)$. Third, the quasilinear form of utility implies that there will be no income effect on bequests. While this simplifies things greatly, it is restrictive. Our assumption that $d_2 > d_1$ partly mitigates the absence of income effects by ensuring that the average level of bequests is higher for type-2's than type-1's. A main advantage of ignoring income effects is that we can clearly highlight the different roles that bequest taxation plays in correcting the social externality resulting from the government discounting the utility of bequest and redistributing between donor and non-donor parents and inheriting and non-inheriting children. In Sect. 6, we discuss how allowing for income effects does not change these potential roles for bequest taxation, but the problem becomes considerably more complicated. Finally, since x absorbs all income effects, we should include a non-negativity constraint on x . In what follows, we assume that it is never binding, so that choices of b and ℓ are always in the interior.

A household 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 $b = (1 - t)B$,

$$x + \frac{b}{1-t} = c + \tau \frac{b}{1-t} \quad (2)$$

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 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 disposable 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:

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

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

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

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). 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.

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:

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

The solution to this problem gives

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

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.

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:

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

The first-order condition for this problem gives

$$h'\left(\frac{y}{w_k}\right) = (1 - T'_c(y))w_k$$

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. We consider the implications of income effects for bequest taxation in Sect. 6.

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 problem using the direct approach, letting the government choose consumption and income for the two parent wage-types, denoted by c_i and y_i , and the two children wage-types, denoted by \bar{c}_k and \bar{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.⁵ 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 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:

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

Differentiating this, we obtain

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

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

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

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 $v_y^i = -h'(\ell_i)/w_i$. Indifference curves have the slope

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

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. This is a consequence of utility being quasilinear and simplifies matters considerably although we discuss the implications of relaxing this assumption in Sect. 6. (Recall that we assume 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.

⁵ In principle, the inheritance tax rate could also differ by the type of the donor, but this would not give the government any more degrees of freedom and would add unnecessary complexity.

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:

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

where

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

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

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

where

$$r_{\bar{c}}^k = 1, \quad r_{\bar{y}}^k = -h'(\bar{\ell}_k)/w_k < 0.$$

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:

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

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 transfer 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

$$\sum_{i=1,2} \left(n_i(y_i - c_i + td_i B_i - \tau_i d_i B_i) + (n_i \pi + n_{-i}(1 - \pi))(\bar{y}_i - \bar{c}_i) \right) = G,$$

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

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

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:

$$c_2 - h\left(\frac{y_2}{w_2}\right) \geq c_1 - h\left(\frac{y_1}{w_2}\right). \tag{16}$$

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 as:

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

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 .⁶ 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

$$\bar{c}_2 - h\left(\frac{\bar{y}_2}{w_2}\right) \geq \bar{c}_1 - h\left(\frac{\bar{y}_1}{w_2}\right) \tag{18}$$

⁶ Given that individuals optimize over the choice of net bequests, the derivative of this expression with respect to p_i reduces to $-b(p_i) < 0$.

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.

The government maximizes social welfare in (14) subject to the budget and incentive constraints, (15)–(18). The first-order conditions are listed in the “Appendix”. We first consider the social optimum when the government can observe wage-types of both parents and children. 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 turn to the imperfect-information case when the government cannot observe wage-types.

4 Full-information benchmark

Suppose the government knows the wage-types of all parents and children and whether a parent is a donor. For comparison purposes, we assume that the government is restricted to using the same policy instruments as in the imperfect-information case, that is, nonlinear income tax systems on parents and children that are not conditioned on donor or donee 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 since the income tax is not contingent on donor or donee status, but it will serve to clarify the efficiency and equity effects of proportional bequest taxation in a setting with redistributive income taxes.

With observable wage-types, no incentive constraints will be binding. The government freely chooses c_i , y_i , \bar{c}_i , \bar{y}_i and p_i for $i = 1, 2$ to maximize (14) subject to the budget constraint (15). The first-order conditions in the “Appendix” apply with $\gamma = \gamma^d = \phi = 0$. The conditions on c_i , y_i , \bar{c}_k and \bar{y}_k yield

$$\frac{h'(\ell_i)}{w_i} = 1, \quad \frac{h'(\bar{\ell}_k)}{w_k} = 1, \quad i, k = 1, 2 \quad (19)$$

which together with (8) imply that the marginal income tax rate for all wage-types is zero and the income tax is non-distorting, as expected. We also obtain results on the equity conditions characterizing optimal redistribution among types and the optimal prices of net bequests.

Consider first the parents. From the first-order conditions on c_i for both wage-types, we obtain:

$$d_1 W'(V^1) + (1 - d_1)W'(v^1) = d_2 W'(V^2) + (1 - d_2)W'(v^2) = \lambda. \quad (20)$$

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 (20) 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. Next, consider the equity condition for the children. From the first-order conditions on \bar{c}_k , we have

$$\lambda = \frac{n^{kd}}{n^k} \left[\left(\frac{n^{kk}}{n^{kd}} \right) W' \left(R^{kk} \right) + \left(1 - \frac{n^{kk}}{n^{kd}} \right) W' \left(R^{ki} \right) \right] + \left(1 - \frac{n^{kd}}{n^k} \right) W' \left(r^k \right)$$

$$k = 1, 2, \tag{21}$$

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 (20) 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.

Together these equity conditions determine the optimal allocation of consumption across the four types of parents (donor/non-donor and high-/low-wage) and the six types of children (high-/low-wage and high-wage/low-wage/non-donor parent). It is worth highlighting at this point that if all parents were donors ($d_1 = d_2 = 1$), then from (20) and (21), there would be full equalization of parents' social utilities and children's utilities. With heterogenous bequest behaviour, $1 > d_2 > d_1$, this is no longer the case. Further, it follows from (20) that if $V^i > v^i$, then $W'(V^i) < \lambda$, and the converse holds.

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 can be written for $i = 1, 2, k \neq i$ as follows:

$$- W' \left(V^i \right) \left(b(p_i) + (1 - \delta) p_i b'(p_i) \right) + \lambda \left((p_i - 1) b'(p_i) + b(p_i) \right)$$

$$+ \alpha \left(\pi W' \left(R^{ii} \right) + (1 - \pi) W' \left(R^{ki} \right) \right) b'(p_i) = 0. \tag{22}$$

The effective price of net bequests plays three distinct roles in government policy. The first role is to correct social externalities. One externality arises as a result of the government giving a different weight to donors' utility of bequests than donors themselves (the 'laundering out' of the benefits of bequest), and another arises as a result of the government giving a different weight to children's utility than donors themselves. The second role is to redistribute between donor and non-donor parents. The final role is to redistribute among children who do and do not receive bequests. Both of these two latter roles arise from differing bequest behaviour. (Recall that the income tax system can do neither of the latter two.) We highlight each of these roles in turn.

4.1 Social externality-correcting role

To focus on this 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 conditions (20) and (21) imply that $\lambda = 1$, and condition (22) reduces to

$$[\delta p_i + \alpha - 1] b'(p_i) = 0 \tag{23}$$

Starting from the laissez-faire, if the government gives full weight to children’s utility ($\alpha = 1$), then the government will want to fully subsidize bequests by setting $p_i = 0$ for any $\delta > 0$. This corresponds with the case considered in the recent literature, such as [Kaplow \(2001\)](#), [Farhi and Werning \(2010\)](#) and [Piketty and Saez \(2013\)](#). In this case, only if the government completely discounted the utility of bequests so $\delta = 0$, would the government reduce net bequests to zero.⁷ Discounting the donor benefits of bequests does not change the efficiency rationale for the subsidization of bequest when the government is not averse to utility inequality provided children’s utility is given full weight in the government’s objective.

Using a comparable base case as 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 ($\alpha = 0$), then for discount factors δ greater than $1/f'(0)$, optimal bequests are positive and the government optimally taxes bequests by setting $p_i = 1/\delta \geq 1$. For discount factors less than $1/f'(0)$, optimal bequests are zero. In this special case with $\alpha = 0$, donors are leaving bequests larger than is optimal from the point of view of the government and the government taxes bequests. The more the government discounts the utility of bequests, the smaller is δ , so the higher the tax rate on bequests to correct this social externality.

Now suppose the government is averse to social utility inequality, so $W'(\cdot) > 0 > W''(\cdot)$. Bequest taxation will continue to serve an externality-correcting role, but will also now serve a role in redistributing between donors and non-donors and among children who do and do not receive bequests.

4.2 Redistribution between donors and non-donors

To consider the role of p_i in distributing between donors and non-donors, begin first with the case where $\alpha = 0$ to suppress redistribution between recipient and non-recipient children for the time being and assume positive bequests in the optimum.⁸ It is useful to consider first whether the bequest tax p_i should be positive or negative, and then whether p_1 should be greater or less than p_2 .

4.2.1 Tax or subsidize bequests?

We can rewrite (22) as

$$\frac{W'(V^i)}{\lambda}(1 - \delta) + \frac{1 - W'(V^i)/\lambda}{\epsilon_p^i} = \frac{p_i - 1}{p_i}. \tag{24}$$

Consider first the case when full weight is given to the utility of bequest, $\delta = 1$, so there is no social externality-correcting role for bequest taxation. The first term on the

⁷ Define \hat{p} such that $f'(0) = \hat{p}$ where $\hat{p} > 1$ since donors leave positive bequests without government intervention and the utility of bequest function is strictly concave. For any $p_i \geq \hat{p}$, donors do not leave any bequests. Evaluating (23) at $p = \hat{p}$ yields an expression that is negative for any $\delta > 0$ and equal to zero for $\delta = 0$.

⁸ That is, the government’s discount rate is assumed to be larger than $1/f'(0)$.

left-hand side is zero. The second term on the left-hand side captures the redistributive role of bequest taxation. When parents differ in their bequest behaviour and with $\delta = 1$, donors will be better off than non-donors from the government's point of view, $V^i > v^i$. It follows from the equity condition (20) 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 the ϵ_p^i , the lower the optimal amount of bequest taxation and the smaller the effective price. This redistribution effect arises solely as a result of heterogeneous bequest behaviour. If all parents are donors (as is generally assumed in the literature), then the social marginal utilities of parents would be identical and equal to the marginal cost of public funds, so this second term would also be zero and the optimal effective price of net bequests would be unity. Here we have shown that it will be optimal to use bequest taxation to redistribute between individuals of differing bequest behaviour.

Next, consider how discounting the utility of bequests affects the role of the effective price in redistribution between non-donor and donor parents. The discount rate δ affects the government's ordering of social utilities between donor and non-donor parents of a given wage-type and ultimately the direction of redistribution. The sign of the second term of (24) depends on this ordering. From the equity condition, (20), the term will be positive (negative) if donors are viewed as better (worse) off than non-donors. In the case where $V^i < v^i$, the net social benefit of bequest—or net social surplus—is negative, that is $\delta f(b(p_i)) - p_i b(p_i) < 0$, and the government wants to redistribute from non-donor to donor parents. By eliminating bequests, the government unambiguously improves the social utility of donor parents and achieves full utility equality. In other words, the sign of the second term of (24) can never be negative in an optimum. Bequests will only be positive in the optimum when the net social surplus from bequests is positive.

More formally, start with the case where p_i 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 $p_i = 1$, and government revenues are still zero. If at $p_i = 1$, 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 is 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, and the government would want to ensure zero bequests. Therefore, if the government sufficiently discounts the utility of bequests (i.e. for social weights less than $b(1)/f(b(1))$), then bequests will optimally be taxed away. For larger social weights, bequests will optimally be positive and donors will be better off than non-donors. The government will tax bequests to correct the social externality and to redistribute from donor to non-donor parents of a given wage-type.

4.2.2 Differential effective prices of net bequests?

Having established that the bequest tax should be positive if bequests are positive in the optimum, it remains to consider how p_i should differ between type-1 and type-2 donor

parents. 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-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 (22) and $\alpha = 0$ that donors of both wage-types 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 (20) 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.

Next, we can show that the government should impose a lower effective price on those wage-types with a larger share of donors. Rewrite the first-order condition on p_i , (22), with $\alpha = 0$ as follows,

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

Given $V^i > v^i$, $W'(V^i) < \lambda$ from (20) and the left-hand side of (25) will be negative. Assuming $V_p^i < 0$, the denominator of the right-hand side of (25) will be positive and since $b'(p_i) < 0$, it must be that $p_i > 1/\delta$. 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 also implies that the right-hand side of (25) is decreasing in p_i .⁹ Now consider the equity condition (20). It can be shown that in the optimum $W'(V^2) > W'(V^1)$.¹⁰ It then follows from (25) that $p_1 > p_2$.

To 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 (22) 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 $n_i d_i W'(V^i) V_p^i$ (still assuming $\alpha = 0$). Therefore, the increase in p_i needed to finance a uniform

⁹ To see this, dropping subscripts

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

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

¹⁰ Assume $d_2 = 1 > d_1 > 0$. From (20), $W'(V^2) = d_1 W'(V^1) + (1 - d_1)W'(v^1)$ 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 (22) and (20) 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 (20) will no longer bind. For (20) to remain binding either v^2 must decrease or v^1 must increase. Either of these changes to keep (20) 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 $V^2 < V^1$.

increase in c_i is given by $-d_i W'(V^i) V_p^i / \lambda$. 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 $p_1 > p_2$.¹¹

The upshot is by restricting attention to the roles of bequest taxation to correct for the discounting of the utility of bequest and to redistribute between donor and non-donor parents, we obtain $p_1 > p_2 > 1$ in the full-information optimum. Let us now extend the analysis to include the effect of p_i on the children.

4.3 Redistribution between recipient and non-recipient children

4.3.1 Tax or subsidize bequests?

Now let $\alpha = 1$ so the utility of children count. Relative to the full-information case with $\alpha = 0$, there is an additional cost of increasing p_i since the bequests now benefit the children. This is reflected in the last term in (22), 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 account of the benefit of their bequests to their children, apart from their own benefit $f(b(p_i))$. This gives the government an incentive to subsidize bequests as discussed above and previously shown in the literature. Consequently, the effective price of net bequests may be greater or less than unity. The effective price of net bequest will always be positive though given the government is averse to social utility inequality, that is, bequests will not be fully subsidized. The revenue gain of reducing the bequest tax credit slightly starting from $\tau_i = 1$ will always be greater than the social loss to donor parents and children receiving bequests and reflects the government's incentive to redistribute from donor to non-donor parents and inheriting children to non-inheriting children.¹²

4.3.2 Differential effective prices of net bequests?

Consider now the pattern of net bequest prices, p_1 and p_2 . 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 low-skilled children. This favours 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 favour low-skilled inheritors counters the opposite desire to redistribute between non-donor and donor parents, and either influence can dominate. To see this more

¹¹ 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 $d_2 > d_1$.

¹² Formally, suppose $p_i = 0$, then both types 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 (21) that $W'(R^{kk}) < \lambda < W'(r^k)$ for $k = 1, 2$. The left-hand side of (22) at $p_i = 0$ reduces to $[\lambda - W'(V^i)]\bar{b} - \lambda b'(0) + [\pi_i W'(R^{ii}) + (1 - \pi_i) W'(R^{kk})] b'(0) > 0$, so the government optimally chooses $p_i > 0$.

formally, suppose we evaluate (22) at $\alpha = 0$. As shown above, the first two terms in (22) will be equal to zero at $p_1 > p_2$. Now consider the last terms in (22). With $p_1 > p_2$, low-wage donor parents are leaving smaller bequests than high-wage donors. Therefore, the average social marginal utility of high-wage donor heirs must be lower than that 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.¹³

To see that the relative sizes of p_1 and p_2 are ambiguous with $\alpha = 1$, rewrite (22) as

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

Suppose the social utility of donors is higher than the social utility of 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 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. 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 $p_2 < p_1$. With a positive weight on children’s well-being in the social welfare function, it may be that the social utility of donors is less than the social utility of the non-donors. The social benefit to heirs of bequests 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. With positive bequests, the equity condition implies that $V^2 > V^1$ given $d_2 > d_1$. How the right-hand side of (26) changes with p is now ambiguous, and therefore, p_1 may be greater or less than p_2 .

Summary of Results in the Full-Information Benchmark

When the government can observe the wage-types of parents and children and impose nonlinear income taxes on each that are not conditioned on donor status along with a less than 100% 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, $p_i = (1 - \tau_i)/(1 - t)$ for each type- i person, so the absolute size of the inheritance tax and the bequest tax credit are indeterminate.
2. The effective price of net bequests plays three distinct roles: it corrects for a social externality when the social weight given to either the utility of bequests or children’s utility differs from the donors’ utility; it redistributes between donor and non-donor parents and between recipient and non-recipient children.
3. In an optimum with positive bequests, if no social welfare weight is given to children ($\alpha = 0$),

¹³ 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 $p_1 > p_2$.

- (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-donor parents of a given wage-type, and
 - (b) the optimal value of p_i will be greater than one and smaller for the wage-type with the highest proportion of donors.
4. Giving children full social weight favours making $p_2 > p_1$, so given that $d_2 > d_1$ the relative size of p_1 and p_2 will be ambiguous; the optimal value of p_i will be positive, but can be greater or less than unity.

5 Imperfect information

Now suppose that the government cannot observe wage-types. Incentive constraints for non-donor and donor parents, (16) and (17), and for children, (18), now apply. These ensure that high-wage parents, regardless of donor status, are better off (from the government’s perspective) than their low-wage counterparts, and similarly for high-wage children.¹⁴ The government cannot implement lump-sum redistribution between wage-types, so bequest taxation is used to redistribute not only between donor and non-donor parents and inheritor and non-inheritor children, but also between wage-types.

The first-order conditions on c_i, y_i, \bar{c}_i and \bar{y}_i in the Appendix 1 characterize the optimal income tax structure. These results are standard: $h'(\ell_2)/w_2 = h'(\bar{\ell}_2)/w_2 = 1, h'(\ell_1)/w_1 < 1, h'(\bar{\ell}_1)/w_1 < 1$, which imply zero marginal tax rates at the top and positive ones at the bottom.

From the first-order conditions on c_1 and c_2 , 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) and sum to unity since $n_1 + n_2 = 1$:

$$n_1 d_1 W'(V^1) + n_1 (1 - d_1) W'(v^1) + n_2 d_2 W'(V^2) + n_2 (1 - d_2) W'(v^2) = \lambda. \tag{27}$$

Similarly, from the first-order conditions on \bar{c}_1 and \bar{c}_2 , we obtain:

$$\begin{aligned} \lambda = & n^{11} W'(R^{11}) + (n^{1d} - n^{11}) W'(R^{12}) + (n^1 - n^{1d}) W'(r^1) \\ & + n^{22} W'(R^{22}) + (n^{2d} - n^{22}) W'(R^{21}) + (n^2 - n^{2d}) W'(r^2) \end{aligned} \tag{28}$$

¹⁴ 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 . For donor parents, if (16) is binding, then $p_1 > p_2$ and it follows that $V^2 > V^1$. If instead (17) binds, then $p_1 < p_2$. Adding and subtracting $\delta f(b(p_i))$ to each side of (17) yield, $V^2 + (1 - \delta)f(b(p_2)) = \hat{V}^1 + (1 - \delta)f(b(p_1))$ 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 p_i and $p_1 < p_2$, it follows that $V^2 > V^1$.

with the weights also summing to unity. Note the differences between (27) and (20) and between (28) and (21). 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 as an admittedly blunt device to redistribute from non-donors to donors and recipients and non-recipients and to correct the social externality. With imperfect information, even 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 given heterogenous bequest behaviour.

The first-order conditions on p_1 and p_2 in the Appendix 1 determine the optimal prices of net bequests. These differ from (22) in 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. As in the full-information case, if the government is not averse to social utility inequality ($W' = 1$) or if all parents are donors ($d_1 = d_2 = 1$) then there is no rationale for bequest taxation other than to correct the social externalities. Under either of these scenarios, the equity conditions (27) and (28) together with the first-order conditions on the net effective prices imply that there will be a uniform effective price on net bequest given by (23) as in the full-information benchmark. Differential bequest taxation will be optimal only when there is heterogeneity of bequest behaviour and some aversion to social inequality. Unlike the full-information case where the incentive constraints did not apply and the effective price of bequests is used to correct the social externality and to redistribute between donor and non-donor parents and between recipient and non-recipient children, they also now influence redistribution between wage-types.

To explore this, it is useful to begin with the case where $\alpha = 0$ to suppress the use of p_i to redistribute between recipient and non-recipient children. We can show in this case 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 are the same as in the full-information case (22). Only the income tax rates are distorted. With positive net social surplus from bequests, donor parents will be better off than non-donor parents of a given wage-type and it follows from (22) that $p_1 < p_2$, a contradiction.¹⁵ 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.

Suppose now that $\alpha = 1$ so that the bequest tax must take into account redistribution between recipient and non-recipient children. From the binding incentive constraint on the children, we know that $r^2 > r^1$ and consequently, $R^{21} > R^{11}$ and $R^{22} > R^{12}$.

¹⁵ 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 (27), $W'(V^2) < \lambda$. 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 (25) is satisfied at $p_2 > 1/\delta$ and $p_1 < 1/\delta$, which again leads to $p_1 < p_2$.

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 bequest from a high-wage donor. The relative magnitude of the effective prices on net bequest continues to depend on 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 $p_1 > p_2$ 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 $p_1 < p_2$: this favours 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 p_2 should be higher or lower than p_1 , but relative to the full-information case, p_1 would tend to be reduced relative to p_2 .

Summary of Results with Imperfect Information

When the government cannot observe the wage-types of either parents or children and impose nonlinear income taxes both that are not conditioned on donor or inheritor status along with a less than 100% 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, $p_i = (1 - \tau_i)/(1 - t)$ for each type- i person, so both the absolute size of the inheritance tax and the bequest tax credit are indeterminate.
2. If all parents are donors ($d_1 = d_2 = 1$), bequests will be fully subsidized, provided there is a positive social weight on the utility of bequests and the welfare of children count. If $\alpha = 0$, then a uniform bequest tax that is decreasing in the social weight δ will be optimal and the bequest tax will be zero for sufficiently small social weights.
3. With heterogenous bequest behaviour, bequest taxation also serves to redistribute between parents and children of differing wage-types.
 - (a) If $\alpha = 0$, the optimal value of p_i will be higher for the high-wage donor parents.
 - (b) If the welfare of children count, the relative size of p_1 and p_2 are ambiguous.

6 Extensions

6.1 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 $p_{ik} = (1 - \tau_i)/(1 - t_k)$.¹⁶ 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 that the fourth instrument is also optimal. To see this, note that the effective prices of net bequests are related as follows:

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

so only three of the prices can be chosen independently. This holds regardless of the assumed information of the government. Allowing the government to differentiate the inheritance tax by the wage-type of the receiving child gives it 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 2 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 r^1 relative to r^2 , rewrite the equity condition for the children, (21), as:

$$\lambda = \frac{n^{1d}}{n^1} \bar{W}'(R^1) + \left(1 - \frac{n^{1d}}{n^1}\right) W'(r^1) = \frac{n^{2d}}{n^2} \bar{W}'(R^2) + \left(1 - \frac{n^{2d}}{n^2}\right) W'(r^2) \tag{29}$$

where $\bar{W}'(R^k)$ is the average marginal utility of type- k recipients. We know that $\bar{W}'(R^k) < W'(r^k)$ for $k = 1, 2$, but the relative size of $\bar{W}'(R^1)$ and $\bar{W}'(R^2)$ is ambiguous. However, as we show in the Appendix 3, $\bar{W}'(R^1) < \bar{W}'(R^2)$ if $\pi = 1$ and $\bar{W}'(R^1) = \bar{W}'(R^2)$ if $\pi = 1/2$. So, we expect that $\bar{W}'(R^1) < \bar{W}'(R^2)$ for $\pi > 1/2$. Given that, the following result is apparent from (37) and (38). Given that $\bar{W}'(R^1) \leq \bar{W}'(R^2)$, $W'(r^1) < W'(r^2)$, so $r^1 > r^2$. This implies from above that t_2 should

¹⁶ 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 that such uncertainty is resolved prior to making the bequest decision.

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 .

6.2 Income effects of bequest taxation

Up to now, we have ignored any income effects on the bequest decision. Suppose instead preferences of donor and non-donor parents are given by

$$U(x, b, \ell) = v(x) + f(b) - h(\ell), \quad u(x, \ell) = v(x) - h(\ell), \quad (30)$$

respectively, where $v(x)$ is a strictly concave function.¹⁷ To relate this discussion to our previous analysis, we continue to assume that the government uses both a linear inheritance tax and a bequest tax credit so the relevant policy variable is the effective price of bequests and that the income tax cannot be conditioned on the level of bequests. But for simplicity, assume that the effective price of bequests is the same for all.¹⁸ As before, it is useful to consider the parent's problem in two stages. In the second stage, a parent decides how to allocate a given after-tax income c between private consumption x and bequests b given $x + pb = c$. Non-donor parents consume their after-tax income, so $x = c$. Donor parents choose b such that $f'(b) = pv'(c - pb)$, which yields optimal bequests $b(p, c)$. Bequests will be decreasing in the effective price of net bequests and increasing in after-tax income.¹⁹

In the first stage, parents decide how much income to earn which for a given income tax system yields both income y and after-tax or disposable income c . The marginal rate of substitution between disposable income and earned income for donor and non-donor parents is given by

$$\text{MRS}_{cy} = \frac{h'(y/w_i)/w_i}{v'(c - pb(p, c))}, \quad \text{MRS}_{cy} = \frac{h'(y/w_i)/w_i}{v'(c)}, \quad i = 1, 2, \quad (31)$$

respectively. For a given (c, y) -bundle, donor parents will be better off and will have a lower marginal rate of substitution between disposable income and earned income than non-donor parents of a given wage-type, provided bequests are positive. In other words, the indifference curves of donor and non-donor parents of a given wage-type will exhibit a single-crossing property in (c, y) -space (with positive bequests). These indifference curves coincide when there are no income effects on bequests. We also obtain the standard single-crossing property between high- and low-wage parents of a given donor status.

¹⁷ The additively separable form of $U(x, b, \ell)$ is maintained so as to allow the government to continue to weight unambiguously the utility of bequests to donors by less than unity.

¹⁸ The linear bequest tax could be made more progressive if it were accompanied by a lump-sum transfer to donors. We neglect that for simplicity.

¹⁹ Totally differentiating the donor parent's first-order condition yields $b_p = (v' - pbv'')/(f'' + p^2v'') < 0$ and $b_c = (pv'')/(f'' + p^2v'') > 0$.

Consider the *laissez-faire* when all four types of parents set $MRS_{cy} = 1$ and $p = 1$. As donor parents leave positive bequests when $p = 1$,²⁰ they will choose to work more than non-donor parents of a given wage-type. We also obtain the standard result that for a given donor status, high-wage parents earn more income than low-wage parents. Consequently, high-wage donor parents will leave greater bequests than low-wage donor parents in the *laissez-faire*. In this case, we do not need to assume that there is a greater share of high-wage donor parents to generate the result that the average bequest for high-wage parents is greater than the average bequest for low-wage parents. Assuming that preferences of children are the same as non-donor parents as before, we can also conclude that inheriting children will work less than non-inheriting children of the same wage-type.²¹ These results can be contrasted to the case when preferences are quasilinear in consumption where donor parents of both wage-types leave the same bequests and labour supply is independent of both donor and inheritance status.

Consider now the case for taxing bequests. Having income effects does not change the government's incentive to tax bequests when it discounts the utility of bequests to donors. There is still a social externality in this case since the government puts less weight on bequests than donors do so would like to influence donors' behaviour. At the same time, the government has an incentive to subsidize bequests when the welfare of children is given full weight in the government's objective.

Things become complicated when considering the redistributive role for bequest taxation. Focus on the imperfect-information case. Although we assume donor status is not observable, the government can use the income tax schedule to separate donor from non-donor parents by choosing separate consumption-income bundles based on donor status subject to various incentive constraints. In addition to the incentive constraints between donor parents and non-donor parents of differing wage-types, there will now be incentive constraints between donor parents of one wage-type and non-donor parents of the other wage-type. Which incentive constraints are binding will depend on the underlying individual preferences, wage distribution and government policy. For illustrative purpose, assume that the MRS_{cy} 's of the four types of parents for a given (c, y) -bundle ascend from the lowest to the highest in the following order: (1) high-wage donor parents, (2) low-wage donor parents, (3) high-wage non-donor parents and (4) low-wage non-donor parents. Assume also that in the *laissez-faire* the ranking of parents' social utility descends in the same order: (1) highest, then (2), etc. In this case, when an optimal income tax is imposed, the incentive constraints will be downward binding—(1) will be binding on (2), (2) on (3) and (3) on (4).²² The question then is whether a bequest tax can improve upon the optimal nonlinear income tax schedule.

By the Atkinson–Stiglitz theorem, a bequest tax cannot improve redistributive outcomes between (1) high-wage donor and (2) low-wage donor parents. Nor can it help between (3) high-wage non-donor and (4) low-wage non-donor parents. At best, it

²⁰ We assume that $f'(0) - v'(y) > 0$ where y is determined from $h'(y/w)/w = v'(y)$.

²¹ Indifference curves of children of a given wage-type and differing levels of inheritances will exhibit a single-crossing property in (c, y) -space.

²² If the ranking of the social utilities in the *laissez-faire* of (2) low-wage donor and (3) high-wage non-donor parents is reversed, then the incentive constraint between (2) and (3) would be upward-binding.

can improve redistribution between (2) low-wage donor and (3) high-wage non-donor parents. We can show that a positive bequest tax will be useful for this purpose given the assumed social ranking. To see this, suppose we start from an optimal nonlinear income tax system with no bequest tax ($p = 1$). Consider the introduction of a small proportional bequest tax and an adjustment to the optimal nonlinear income tax schedule to keep utilities of the four types constant. The high-wage non-donor parents do not leave a bequest so they are unaffected by this tax and their nonlinear income tax schedule does not need to be adjusted. The low-wage donor parents leave a bequest whether or not they mimic. If they mimic, they earn the income of the high-wage non-donor parent and receive the same net income. Consequently, they leave a bequest conditional on the net income of the individual they are mimicking. Their nonlinear income tax (assuming they do not mimic) can be adjusted so as to compensate them for the introduction of the bequest tax, but since they pay the bequest tax whether or not they mimic they will now be worse off if they mimic. In this case, the introduction of a small proportional bequest tax could weaken the incentive constraint between the low-wage, donor parent and the high-wage non-donor parent and social welfare can be improved. By a similar argument, if the ranking of the parents' social utility between (2) and (3) was reversed, the incentive constraint would bind in the other direction and a small proportional bequest subsidy could be a useful policy alongside the optimal nonlinear income tax system. The ability of the income tax to differentiate between donors and non-donors of the same wage-type because of income effects reduces the need to use the bequest tax for that purpose. Bequest taxation can, however, still be used to relax the self-selection constraint between parents who differ in both wage and donor status even with weak separability.

Bequest taxation in this environment will also be useful to redistribute between inheriting and non-inheriting children. With a positive correlation between the wage-types of children and parents and since high-wage donors given greater bequests than low-wage donor parents, there will be a positive association between a child's wage-type and the inheritance they received. Consequently, as argued by [Brunner and Pech \(2012a, b\)](#), bequest taxation can be used to reduce inequities within the children's generation. Overall, whether the optimal bequest tax is positive or negative depends on the relative importance of the social externality effect, redistribution between inheriting and non-inheriting children, and improving redistribution between heterogeneous parents.

7 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 non-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 bequests 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 donees 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 useful next step.

Acknowledgments The authors thank two anonymous referees and the guest editor of this special issue, Monika Büttler, for their helpful comments. The authors are also grateful for comments received on earlier versions of this paper from participants 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 by seminar participants at the University of Victoria and at the University of Oslo and by participants at the 70th Annual Congress of the International Institute of Public Finance are also appreciated. Financial support from the Social Sciences and Humanities Research Council of Canada and the Canada Research Chair Programme is gratefully acknowledged.

Appendix 1: First-order conditions on c_i , y_i , \bar{c}_k and \bar{y}_k , and p_i

The first-order conditions on c_i , y_i , \bar{c}_k , \bar{y}_k , and p_i can be written as follows, using (10), (12) and (13):

$$n_1 d_1 W'(V^1) + n_1(1 - d_1)W'(v^1) - (\gamma + \gamma^d) - \lambda n_1 = 0 \quad (c_1)$$

$$\begin{aligned} -n_1 d_1 W'(V^1) \frac{h'(\ell_1)}{w_1} - n_1(1 - d_1)W'(v^1) \frac{h'(\ell_1)}{w_1} \\ + (\gamma + \gamma^d) \frac{h'(\hat{\ell}_2)}{w_2} + \lambda n_1 = 0 \end{aligned} \quad (y_1)$$

$$n_2 d_2 W'(V^2) + n_2(1 - d_2)W'(v^2) + (\gamma + \gamma^d) - \lambda n_2 = 0 \quad (c_2)$$

$$\begin{aligned}
 & -n_2d_2W'(V^2)\frac{h'(\ell_2)}{w_2} - n_2(1 - d_2)W'(v^2)\frac{h'(\ell_2)}{w_2} \\
 & - (\gamma + \gamma^d)\frac{h'(\ell_2)}{w_2} + \lambda n_2 = 0 \tag{y_2}
 \end{aligned}$$

$$\begin{aligned}
 & n_1\pi d_1\alpha W'(R^{11}) + \left(n_1\pi(1 - d_1) + n_2(1 - \pi)(1 - d_2)\right)\alpha W'(r^1) \\
 & + n_2(1 - \pi)d_2\alpha W'(R^{12}) - \phi - \lambda(n_1\pi + n_2(1 - \pi)) = 0 \tag{\bar{c}_1}
 \end{aligned}$$

$$\begin{aligned}
 & -n_1\pi d_1\alpha W'(R^{11})\frac{h'(\bar{\ell}_1)}{w_1} - \left(n_1\pi(1 - d_1) + n_2(1 - \pi)(1 - d_2)\right)\alpha W'(r^1)\frac{h'(\bar{\ell}_1)}{w_1} \\
 & - n_2(1 - \pi)d_2\alpha W'(R^{12})\frac{h'(\bar{\ell}_1)}{w_1} + \phi\frac{h'(\widehat{\ell}_2)}{w_2} + \lambda(n_1\pi + n_2(1 - \pi)) = 0 \tag{\bar{y}_1}
 \end{aligned}$$

$$\begin{aligned}
 & n_2\pi d_2\alpha W'(R^{22}) + \left(n_2\pi(1 - d_2) + n_1(1 - \pi)(1 - d_1)\right)\alpha W'(r^2) \\
 & + n_1(1 - \pi)d_1\alpha W'(R^{21}) + \phi - \lambda(n_2\pi + n_1(1 - \pi)) = 0 \tag{\bar{c}_2}
 \end{aligned}$$

$$\begin{aligned}
 & -n_2\pi d_2\alpha W'(R^{22})\frac{h'(\bar{\ell}_2)}{w_2} - \left(n_2\pi(1 - d_2) + n_1(1 - \pi)(1 - d_1)\right)\alpha W'(r^2)\frac{h'(\bar{\ell}_2)}{w_2} \\
 & - n_1(1 - \pi)d_1\alpha W'(R^{21})\frac{h'(\bar{\ell}_2)}{w_2} - \phi\frac{h'(\bar{\ell}_2)}{w_2} + \lambda(n_2\pi + n_1(1 - \pi)) = 0 \tag{\bar{y}_2}
 \end{aligned}$$

$$\begin{aligned}
 & -n_1d_1W'(V^1)(b(p_1) + (1 - \delta)p_1b'(p_1)) \\
 & + n_1d_1\left[\pi\alpha W'(R^{11}) + (1 - \pi)\alpha W'(R^{21})\right]b'(p_1) \\
 & + \lambda n_1d_1\left((p_1 - 1)b'(p_1) + b(p_1)\right) + \gamma^db(p_1) = 0 \tag{p_1}
 \end{aligned}$$

$$\begin{aligned}
 & -n_2d_2W'(V^2)(b(p_2) + (1 - \delta)p_2b'(p_2)) \\
 & + n_2d_2\left[\pi\alpha W'(R^{22}) + (1 - \pi)\alpha W'(R^{12})\right]b'(p_2) \\
 & + \lambda n_2d_2\left((p_2 - 1)b'(p_2) + b(p_2)\right) - \gamma^db(p_2) = 0 \tag{p_2}
 \end{aligned}$$

where either $\gamma = 0$ or $\gamma^d = 0$, depending on whether in the optimum $p_1 < p_2$ or $p_1 > p_2$, respectively, and ϕ is the multiplier on the incentive constraint for the children.

Appendix 2: 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:

$$\begin{aligned}
 \mathcal{W} = & n_1d_1\left(\pi W(V^1(c_1, y_1, p_{11})) + (1 - \pi)W(V^1(c_1, y_1, p_{12}))\right) \\
 & + n_1(1 - d_1)W(v^1(c_1, y_1)) \\
 & + n_2d_2\left(\pi W(V^2(c_2, y_2, p_{22})) + (1 - \pi)W(V^2(c_2, y_2, p_{21}))\right)
 \end{aligned}$$

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

and the government’s budget constraint (2) can be rewritten as (with corresponding multiplier λ)

$$\begin{aligned}
 &n_1(y_1 - c_1) + (n_1\pi + n_2(1 - \pi))(\bar{y}_1 - \bar{c}_1) + n_2(y_2 - c_2) + (n_2\pi + n_1(1 - \pi)) \\
 &\times (\bar{y}_2 - \bar{c}_2) + n_1\pi d_1(p_{11} - 1)b(p_{11}) + n_1(1 - \pi)d_1(p_{12} - 1)b(p_{12}) \\
 &+ n_2\pi d_2(p_{22} - 1)b(p_{22}) + n_2(1 - \pi)d_2(p_{21} - 1)b(p_{21}) = 0.
 \end{aligned}$$

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

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

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

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

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

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

which reduces to

$$\begin{aligned} \frac{\partial \mathcal{L}^*}{\partial t_2} &= n_1 d_1 (1 - \pi) \pi \left[W'(R^{21}) - W'(R^{11}) \right] b'(p_1) \frac{(1 - \tau_1)}{(1 - t)^2} \\ &\quad + n_2 d_2 \pi (1 - \pi) \left[W'(R^{22}) - W'(R^{12}) \right] b'(p_2) \frac{(1 - \tau_2)}{(1 - t)^2} \end{aligned}$$

Appendix 3: Size of $\bar{W}'(R^1)$ relative to $\bar{W}'(R^2)$

Suppose first that $\pi = 1$ so parent and child wage-types are perfectly correlated. The equity condition (22) can be written:

$$\begin{aligned} \lambda &= d_1 W'(r^1 + b(p_1)) + (1 - d_1) W'(r^1) \\ &= d_2 W'(r^2 + b(p_2)) + (1 - d_2) W'(r^2) \end{aligned}$$

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:

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

$i = 1, 2.$

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 right-hand side of (31') becomes less than the left-hand side (since $W'(R^2) < W'(r^2)$). Given d_1 , the derivative of the left-hand 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 (22) becomes:

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

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

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

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

where we can write

$$E(\bar{W}'(R^1))(n_1 d_1 + n_2 d_2) = n_1 d_1 W'(R^{11}) + n_2 d_2 W'(R^{12})$$

$$E(\bar{W}'(R^2))(n_1 d_1 + n_2 d_2) = n_2 d_2 W'(R^{22}) + n_1 d_1 W'(R^{21})$$

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

References

- Archibald, C., & Donaldson, D. (1976). Non-paternalism and externalities. *Canadian Journal of Economics*, 9, 492–507.
- Atkinson, A. B., & Stiglitz, J. E. (1976). The design of tax structure: Direct vs. indirect taxation. *Journal of Public Economics*, 6, 55–75.
- Brunner, J. K., & Pech, S. (2012a). Optimal taxation of bequests in a model with initial wealth. *Scandinavian Journal of Economics*, 114, 1368–1392.
- Brunner, J. K., & Pech, S. (2012b). Optimal taxation of wealth transfers when bequests are motivated by joy of giving. *The B.E. Journal of Economic Analysis and Policy*. doi:10.1515/1935-1682.3070.
- Cremer, H., & Pestieau, P. (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, pp. 1107–34). Amsterdam, North-Holland: Elsevier.
- Cremer, H., & Pestieau, P. (2011). The tax treatment of intergenerational wealth transfer. *CESifo Economic Studies*, 57, 365–401.
- Farhi, E., & Werning, I. (2010). Progressive estate taxation. *Quarterly Journal of Economics*, 125, 635–673.
- Farhi, E., & Werning, I. (2013). Estate taxation with altruism heterogeneity. *American Economic Review: Papers and Proceedings*, 103, 489–495.
- Frank, R. H. (2008). Should public policy respond to positional externalities? *Journal of Public Economics*, 92, 1777–1786.
- Hammond, P. J. (1987). Altruism. In J. Eatwell, M. Milgate, & P. Newman (Eds.), *The new palgrave: A dictionary of economics* (pp. 85–87). Basingstoke, UK: Palgrave Macmillan.
- Hochman, H. M., & Rodgers, J. D. (1969). Pareto optimal redistribution. *American Economic Review*, 59, 542–557.
- Kaplow, L. (1998). Tax policy and gifts. *American Economic Review Papers and Proceedings*, 88, 283–288.
- Kaplow, L. (2001). A framework for assessing estate and gift taxation. In W. G. Gale, J. R. Hines, & J. Slemrod (Eds.), *Rethinking estate and gift taxation* (pp. 164–204). Washington: Brookings Institution.
- Kaplow, L. (2008). *The theory of taxation and public economics*. Princeton: Princeton University Press.

- Kopczuk, W. (2012). Taxation of intergenerational transfers and wealth. In A. Auerbach, R. Chetty, M. Feldstein, & E. Saez (Eds.), *Handbook of public economics* (Vol. 5, pp. 329–390). Amsterdam: Elsevier.
- Kopczuk, W. (2013). Incentive effects of inheritances and optimal estate taxation. *American Economic Review: Papers and Proceedings*, 103, 472–477.
- Mankiw, G. N. (2006). *The estate tax debate*. Greg Mankiw's Blog: Random Observations for Students of Economics. <http://gregmankiw.blogspot.ca/2006/06/estate-tax-debate.html>. Accessed 6 June 2015.
- McCaffery, E. J. (1994). Uneasy case for wealth transfer taxation. *Yale Law Journal*, 104, 283–365.
- Meade Report (1978). *The structure and reform of direct taxation*. Report of a committee chaired by Professor J. E. Meade. London: George Allen and Unwin.
- Mirrlees, J.A. (2007). Taxation of gifts and bequests, slides for a talk at the centenary of James Meade Conference.
- Mirrlees, J.A. (2011). Optimal taxation of saving and inheritance, slides for a talk at the University of St. Gallen.
- Mirrlees, J. A., Adam, S., Besley, T., Blundell, R., Bond, S., Chote, R., et al. (2011). *Tax by design: The Mirrlees review*. Oxford: Oxford University Press.
- Oswald, A. (1983). Altruism, jealousy and the theory of optimal non-linear taxation. *Journal of Public Economics*, 20, 77–87.
- Pauly, M. V. (1973). Income distribution as a local public good. *Journal of Public Economics*, 2, 35–58.
- Piketty, T., & Saez, E. (2013). A theory of optimal inheritance taxation. *Econometrica*, 81, 1851–1886.
- Thurow, L. S. (1971). The income distribution as a pure public good. *Quarterly Journal of Economics*, 85, 327–336.