

Pareto-improving bequest taxation

Volker Grossmann · Panu Poutvaara

Published online: 24 May 2008
© Springer Science+Business Media, LLC 2008

Abstract Altruistic parents may transfer resources to their offspring by providing education, and by leaving bequests. We show that in the presence of wage taxation, a small bequest tax may improve efficiency in an overlapping-generations framework with only intended bequests, by enhancing incentives of parents to invest in their children's education. We also calculate an optimal mix of wage and bequest taxes with alternative parameter combinations. In all cases, the optimal wage tax rate is clearly higher than the optimal bequest tax rate, but the latter is generally positive when the required government revenue in the economy is sufficiently high. If educational investment is partly unobservable for the government, these results qualitatively hold also when allowing for education subsidies.

Keywords Bequest taxation · Bequests · Education · Pareto improvement

JEL Classification H21 · H31 · D64 · I21

Electronic supplementary material The online version of this article (<http://dx.doi.org/10.1007/s10797-008-9080-1>) contains supplementary material, which is available to authorized users.

V. Grossmann (✉)

Department of Economics, University of Fribourg, Bd. de Pérolles 90, 1700 Fribourg, Switzerland
e-mail: volker.grossmann@unifr.ch

V. Grossmann · P. Poutvaara
CESifo, Munich, Germany

V. Grossmann · P. Poutvaara
Institute for the Study of Labor (IZA), Bonn, Germany

P. Poutvaara
Department of Economics, University of Helsinki, P.O. Box 17, Arkadiankatu 7, Helsinki 00014, Finland
e-mail: panu.poutvaara@helsinki.fi

1 Introduction

Normative analyses of estate taxation suggest that the case for taxing bequests is rather weak.¹ For instance, a strong case against bequest taxation comes from infinite-horizon, Ramsey-type models. As it is well known, this kind of framework can be interpreted as a model of individuals with a Barro-type form of altruism (Barro 1974) who live one period, so that bequest taxation coincides with capital income taxation. Chamley (1986) and Judd (1985) show that in an infinite-horizon framework, the disincentives to accumulate capital and the implied effects on the consumption stream are so strong that the optimal capital income tax converges to zero, despite potential benefits from redistribution across heterogeneous agents.

The Chamley–Judd result of zero capital income taxation in the limit has been qualified by extending the neoclassical growth model to imperfect goods market competition (Judd 2002), unemployment as a result of search frictions in the labor market (Domeij 2005) and human capital formation (Jones et al. 1993, 1997).² A nonzero bequest tax is potentially desirable in finite horizon models as well. For instance, it may derive from the possibility of accidental bequests (Blumkin and Sadka 2003),³ redistribution effects in heterogeneous agent models (e.g., Cremer and Pestieau 2001) or, as pointed out by Kopczuk (2001), from negative externalities arising from wealth inequality.

What the previous literature has in common is its focus on financial bequests as single source of intergenerational transfers. In this paper, altruistic parents face a trade-off between investing in their children's education and leaving bequests. Starting from a second-best world in which wage taxation distorts human capital investment, we show that taxation of intended bequests can be justified for pure efficiency reasons. Even if the wage tax rate is held constant, introducing a bequest tax can be Pareto-improving by enhancing incentives of parents to invest in their children's education. In our model, this holds when the positive effect of bequest taxation on human capital formation is sufficiently high to outweigh the negative effects from reduced wealth accumulation. We also provide numerical results on the optimal tax structure. These demonstrate that with a given revenue requirement and endogenously chosen proportional tax rates on wage income and bequests, the tax rate on bequests depends positively on the extent of the distortion a wage tax causes on educational investments. The results also suggest that the wage tax rate should be considerably higher than the bequest tax rate. The latter is positive when the required government revenue

¹For an excellent survey of the existing literature on optimal bequest taxation under various motives to leave financial bequests, see Cremer and Pestieau (2003).

²Judd (2002) suggests that the capital income tax should be negative if there is imperfect competition, whereas Domeij (2005) shows that whether it should be positive or negative depends on the tightness of the labor market. Jones et al. (1993) show that the optimal long-run tax on capital income is positive in an endogenous growth framework where government spending is productive. Jones et al. (1997) argue that the Chamley–Judd result also fails to hold when there are pure rents, or different types of labor which need to be taxed at the same rate.

³Blumkin and Sadka (2003) provide an important modification of the result that accidental bequests should fully be taxed because such a tax seemingly has lump-sum character. They show that the optimal tax on accidental bequests is typically below 100% when labor supply is endogenous and there is wage taxation.

in the economy is sufficiently high. If educational investment is partly unobservable for the government, these results qualitatively hold also when allowing for education subsidies, although these generally reduce the potentially beneficial role of positive bequest taxation.

Our paper is probably most closely related to the recent contributions of Michel and Pestieau (2004) and Jacobs and Bovenberg (2005). Like Michel and Pestieau (2004), we analyze an optimal mix between wage taxation and bequest taxation in a model with non-Barrobian dynasties. Whereas Michel and Pestieau (2004) assume a “joy of giving” bequest motive and follow the existing literature by focusing on bequests as the only form of intergenerational transfers, we assume that parents receive utility from their offsprings’ disposable income. Hence, parental utility depends on both their financial bequests and educational investment. Focusing on a steady state, Michel and Pestieau (2004) show that bequest taxes should typically be negative when the social planner takes into account the parental bequest motive. In contrast, we derive a plausible condition under which the optimal tax rate on bequests may well be positive. Introducing a positive tax on bequests may even improve the utility of all currently living and future generations, instead of just maximizing the objective function of a social planner attaching certain weights on current and future generations, without requiring a Pareto-improvement.

Jacobs and Bovenberg (2005) analyze optimal linear taxes on capital and labor income with human capital investment and financial savings. They find that the positive tax on capital income serves to alleviate distortions arising from labor income taxation. Our paper differs from their contribution in two crucial respects. First, we analyze an infinitely lasting OLG economy while Jacobs and Bovenberg (2005) assume that the economy lasts only for three periods. The positive capital income taxes that Jacobs and Bovenberg (2005) derive are in line with Jones et al. (1993) who show that even if optimal capital income taxes would converge to zero also in the presence of human capital formation, they are typically positive within a finite time. We identify conditions under which bequest taxes are positive also in the steady-state. Second, Jacobs and Bovenberg (2005) do not consider intergenerational transfers or altruism, which is the focus of this paper. Our contribution to the existing literature thus is to examine the welfare effects of bequest taxation with finite lives when parents can invest in their children’s education.⁴

In the coming section, we present the basic structure of the model. In Sect. 3, we analyze the equilibrium, particularly focusing on the question under which conditions bequest taxation leads to a Pareto-improvement. Section 4 provides numerical illustrations on the optimal mix of (linear) wage and bequest taxation. Section 5 provides an extension to education subsidies and discusses the role of intergenerational externalities due to altruism of parents for the optimal tax structure. The last section concludes. All proofs are relegated to the [Appendix](#).

⁴We are by far not the first ones, however, to analyze the interplay between bequests and investment in education by parents. Blinder (1976) studies intergenerational transfers and life cycle consumption and remarks that differential tax treatment of intergenerational transfers of human capital and bequests should have consequences on the mix of the two. However, he does not provide a formal analysis. Ishikawa (1975) analyzes household decisions concerning education and bequests in the absence of taxation.

2 The model

2.1 Production of final output

In every period, a single homogeneous consumption good is produced according to a neoclassical, constant-returns-to-scale production technology. Output at time t , Y_t , is

$$Y_t = F(K_t, H_t) \equiv H_t f(k_t), \quad k_t \equiv K_t/H_t, \quad (1)$$

where K_t and H_t are the amounts of physical capital and human capital employed in period t , respectively, the latter being measured in efficiency units. $f(\cdot)$ is a strictly monotonically increasing and strictly concave function which fulfills $\lim_{k \rightarrow \infty} f'(k) = 0$ and $\lim_{k \rightarrow 0+} f'(k) = \infty$.⁵

Output is sold to a perfectly competitive world market, with output price normalized to unity. The rate of return to capital, r_t , is internationally given and time-invariant, i.e., $r_t = \bar{r}$. That is, we analyze a small open economy framework with perfectly mobile capital.

Profit maximization of the representative firm in any period t implies that $\bar{r} = f'(k_t)$. Thus, $k_t = (f')^{-1}(\bar{r}) \equiv \bar{k}$. The wage rate per efficiency unit of human capital, w_t , reads $w_t = f(\bar{k}) - \bar{k}f'(\bar{k}) \equiv \bar{w}$ and output is given by $Y_t = H_t f(\bar{k})$.

2.2 Individuals and education technology

In each period t , a unit mass of identical individuals (generation t) is born. An individual lives three periods. In the first period (childhood), individuals live by their parents and acquire education. In the second period (working age), individuals supply their human capital to the labor market, give birth to one child, invest in their children's human capital,⁶ and save for old age. In their final period of life (retirement age), they allocate their wealth between consumption and transfers to their offspring, from now on labeled "bequests." For simplicity, suppose that the financial market is perfect and there is no human capital risk.

An individual born in period t (a member of generation t) with parental investment e_t (in units of the consumption good) in education acquires

$$h_{t+1} = h(e_t), \quad (2)$$

units of human capital in $t + 1$, where $h(\cdot)$ is a strictly monotonically increasing and strictly concave function which fulfills $h(0) \geq 0$, $\lim_{e \rightarrow \infty} h'(e) = 0$ and $\lim_{e \rightarrow 0+} h'(e) = \infty$.⁷ As individuals are identical and of unit mass, the aggregate human capital stock is given by $H_{t+1} = h_{t+1}$. Let s_{t+1} denote the amount of savings

⁵The capital-skill complementarity underlying production function (1) is empirically well supported; see, e.g., Goldin and Katz (1998).

⁶Human capital investments can be thought of as both nonschooling forms of training and private schooling.

⁷For a similar specification and a discussion of diminishing returns to human capital investment, see, e.g., Galor and Moav (2004), among others.

of a member of generation t for retirement. Initially, at $t = 1$, both savings of the currently old generation (born in $t = -1$), s_0 , and the education level of the current middle-aged generation (born at $t = 0$), e_0 , are given. (Hence, the initial stock of human capital, $H_1 = h(e_0)$ is given.)

Utility U_t of a member of generation t is defined over consumption levels $c_{2,t+1}$ and $c_{3,t+2}$ in the working and retirement age, respectively, and *disposable* income of the offspring (born in $t + 1$) in its working age, I_{t+2} .⁸ Assuming additively separable utility, we have

$$U_t = u_2(c_{2,t+1}) + \beta V(c_{3,t+2}, I_{t+2}), \quad (3)$$

$$V(c_{3,t+2}, I_{t+2}) = u_3(c_{3,t+2}) + v(I_{t+2}), \quad (4)$$

where $u_2(\cdot)$, $u_3(\cdot)$, and $v(\cdot)$ are strictly monotonic increasing and strictly concave functions, and $\beta \in (0, 1)$ is a discount factor. The altruism motive reflects the notion that parents care about the economic situation of their offspring. It may be called “joy-of-children-receiving-income,” in contrast to the often assumed “joy-of-giving” motive. In the latter, the bequeathed amount of resources enters utility of parents and parents do not care about other sources of children’s consumption (see Andreoni 1989, for an important early contribution on giving with impure altruism). However, in the present context, in which parents also finance the human capital investment of children, joy of giving would imply that parents value education per se, rather than as a means to earn income.⁹

As will become apparent in Sect. 4, our “joy-of-children-receiving-income” motivation gives rise to externalities of intergenerational transfers which renders nonzero taxes optimal even if no public spending has to be financed. The reason is similar as under a “joy-of-giving” motive. Since parents do not care about children’s utility per se, intergenerational transfers are suboptimal from a social planner’s point of view.

Externalities from intergenerational transfers do not arise under a “dynastic” altruism motive as suggested by Barro (1974), in which parents care about the well being of their offspring. In our context, this would imply a utility function of the form $U_t = u_2(c_{2,t+1}) + \beta u_3(c_{3,t+2}) + \gamma U_{t+1}$, $0 < \gamma < 1$. We do not adopt such a utility function for two reasons. First, the Barrovian bequest motive has been criticized, inter alia, because it means that individuals act as they would be infinitively-living, as implied by the recursive definition of utility. Second, one can easily show that with an internationally given interest rate, such a utility function rules out an interior steady-state solution to the individual optimization problem in the proposed overlapping-generations structure except for a knife-edge parameter constellation. An analysis of a small open economy under perfect capital mobility is, however,

⁸At the cost of some notational complexity, we could introduce either an exogenous consumption for children, or assume that the utility function of the middle-aged parents would have the family consumption as its argument, this being optimally allocated between the parent and the child.

⁹Our bequest motive is linked to Gradstein and Justman (1997), who assume that parents care about the earnings capacity of children. However, in their model, gross rather than net income of children enters parents’ utility and parents do not leave financial bequests. Moreover, our bequest motive is related to Blinder (1976), who assumes that the after-tax bequest enters parents’ utility function.

becoming a more attractive reference point than a closed economy for several countries. For governments facing capital mobility, it has become increasingly difficult to tax capital. A question which arises in this context is if taxing bequests could play a role to avoid shifting too much of the tax burden to labor when financing the public sector.

2.3 Public sector

The government has to finance an exogenous expenditure $G \geq 0$ in each period. In the Chamley–Judd framework, the problem of the government is to choose an optimal intertemporal profile of wage taxes and bequest taxes to finance its expenditures over time. While acknowledging the importance of this traditional approach, we adopt a more challenging criterion of intertemporal Pareto-optimality: We require that each generation has to be made better off; moreover, we assume that the government budget has to be balanced each period, for the following reasons.

From normative perspective, we view the Chamley–Judd framework as fully appropriate for their analysis of infinitely-lived households, but more problematic in an overlapping generations environment. Judd (1985, 2002), Chamley (1986) and Jones et al. (1993, 1997) conclude that it is generally optimal for the government with an intertemporal budget constraint to levy taxes in the initial periods to establish a fund that can be used to pay steady-state expenditures, allowing often tax rates to converge to zero in the long run. In an overlapping generations framework, this would imply sacrificing the utility of a potentially large number of current and future generations to benefit the subsequent generations far away. To avoid the potentially contentious issue of comparing welfare between different generations, we adopt the stricter test of intergenerational Pareto-improvement.

From the positive perspective, we view the idea that a government could tax several generations to collect a fund to benefit subsequent generations rather demanding.¹⁰ Indeed, in most countries governments have accumulated net debt, rather than even started creating large funds that would allow them to pay future expenditures without levying taxes. As a compromise between the normative prediction by the Chamley–Judd framework and the stylized fact that most governments do not collect such funds, we assume that the government budget has to be balanced in each period. Naturally, lifting such restriction would widen the scope for an intertemporal Pareto improvement.

For the equilibrium analysis of the coming section, we follow the tradition by Judd (1985, 2002), Chamley (1986), and Jones et al. (1993, 1997) by assuming that for financing G the government has to use linear taxes on wages and bequests. There are

¹⁰If the results that Chamley, Judd, and Jones et al. (1993, 1997) derive in an infinitely-lived agent framework would be extrapolated to a world of overlapping generations, their findings would suggest as an optimal tax policy to levy potentially high taxes during several generations to accumulate funds that would finally generate enough interest to allow future governments to pay for expenditures. However, such funds could tempt generations alive in any given period in future to spend at least part of assets, rather than just the interest that a social planner alive several generations ago intended them to receive. Furthermore, it is not evident that current generations would be willing to sacrifice their utility to accumulate assets that would be used to improve the standards of living after several generations.

no other taxes. We thereby focus on interactions between wage and bequest taxation. We consider these interactions to be the most interesting ones in our framework for the following reasons. First, labor income taxation is the main source of government revenue in all advanced countries. Second, as intuitive and as will become apparent, it directly distorts human capital investment. Since the novel feature of our analysis is to study bequest taxation in a model in which altruism of parents is reflected by both financial bequests *and* educational investment, it seems natural to examine the desirability of a positive bequest tax *conditional* on the extent of the distortion caused by wage taxation. However, in Sect. 5 we discuss the additional role of education subsidies in our framework. (For tractability reasons, these discussions are based on numerical analyses only.) Finally, we assume that positive lump-sum taxes are non-feasible.

3 Equilibrium analysis

This section analyzes the equilibrium for given tax rates. First, individual decisions are studied. Second, we examine the evolution of the level of human capital investment and the level of bequests. Third, and most important, we analyze the impact of bequest taxation on individual utility. In particular, we ask: Can bequest taxation raise welfare of all generations from the time when a bequest tax is introduced onwards?

3.1 Individual decisions

The pretax bequest received by a member of generation t in her working age (i.e., in $t + 1$) is denoted by b_{t+1} . τ_w and τ_b denote the tax rates on wage income and bequest, respectively, where $\tau_w, \tau_b < 1$. Thus, disposable income of a member of generation t at date $t + 1$ is given by

$$I_{t+1} = (1 - \tau_w)\bar{w}h(e_t) + (1 - \tau_b)b_{t+1} + T_{t+1}, \quad (5)$$

where $T_{t+1} \geq 0$ denotes a potential lump-sum transfer. The possibility of lump-sum transfers is introduced for a conceptual reason and will play a minor role in what follows. It specifies that any tax revenue which may exceed G is redistributed in a lump-sum fashion. Such tax revenue could in principle accrue when introducing a bequest tax while holding the tax on wage income constant. As we are interested in the question whether introducing a bequest tax may be efficiency-enhancing, we will examine whether the impact of a marginal increase in τ_b on utility of each generation around $\tau_b = 0$ is positive, which effectively means that we have $T_{t+1} = 0$ for all t .

The government budget constraint in period $t + 1$ is

$$\tau_w \bar{w}h(e_t) + \tau_b b_{t+1} = G + T_{t+1}. \quad (6)$$

Individual budget constraints at date $t + 1$ and $t + 2$ are given by

$$c_{2,t+1} + s_{t+1} + e_{t+1} = I_{t+1}, \quad (7)$$

$$c_{3,t+2} + b_{t+2} = (1 + \bar{r})s_{t+1}, \quad (8)$$

where s_{t+1} denotes working-life savings for retirement. Throughout the paper, we focus on interior solutions of the utility maximization problem in each period. Using (3)–(8), it is straightforward to show that a member of generation t in $t + 1$ (with income I_{t+1}) chooses savings for her old age (s_{t+1}), educational investment for her child (e_{t+1}) in her working age and bequests in retirement age (b_{t+2}) according to first-order conditions

$$\frac{u'_2(c_{2,t+1})}{\beta u'_3(c_{3,t+2})} = 1 + \bar{r}, \quad (9)$$

$$\frac{u'_2(c_{2,t+1})}{\beta v'(I_{t+2})} = (1 - \tau_w) \bar{w} h'(e_{t+1}), \quad (10)$$

and

$$\frac{u'_3(c_{3,t+2})}{v'(I_{t+2})} = 1 - \tau_b, \quad (11)$$

respectively. Optimality condition (9) is standard: the marginal rate of substitution between present and future consumption is equal to the interest rate factor. According to (10), the marginal rate of substitution between present consumption and children's income equals the marginal (net) return of children to human capital investment, whereas (11) says that the marginal rate of substitution between future consumption and (future) bequests equals the net receiving of children per unit of bequests, $1 - \tau_b$.

For later use, note that parental decisions imply that a member of generation t receives income

$$I_{t+1} = \bar{w} h(e_t) + b_{t+1} - G \quad (12)$$

in $t + 1$, according to (5) and (6).¹¹

3.2 Educational investments

We first look at educational investments. By combining (9)–(11) and observing the properties of education technology $h(e)$, it is easy to see that the following results hold.

Proposition 1 (Education) *For any $t \geq 1$, human capital investment, $e_t \equiv e^*(\tau_b, \tau_w)$, is time-invariant, unique, and implicitly given by*

$$(1 - \tau_w) \bar{w} h'(e^*) = (1 - \tau_b)(1 + \bar{r}). \quad (13)$$

Corollary 1 *Educational investment e^* and thus, for all $t \geq 1$ equilibrium output, $Y_{t+1} = h(e^*) f(\bar{k}) \equiv Y^*$, are increasing in τ_b and decreasing in τ_w .*

¹¹Note that combining (8), (11), and (12) implies $u'_3((1 + \bar{r})s_0 - b_1) = (1 - \tau_b)v'(\bar{w}h(e_0) + b_1 - \bar{G})$, i.e., bequest b_1 left by members of the initially old generation is determined by initial conditions: investment e_0 in their offspring's education and savings s_0 in their working age.

According to Proposition 1, the optimal educational investment, e^* , is reached when the marginal after-tax return to education equals the after-tax return on one unit of bequest when invested in the financial market. An important implication of this is that e^* , and thus the gross domestic product, Y^* , is increasing in the degree of bequest taxation (Corollary 1). This is because an increase in τ_b induces parents, who care about net income of their offspring, to substitute away from financial transfers (in retirement age) and invest more in children's education (in working age). This result is novel in the literature on bequest taxation. The other result—that higher earnings taxation (i.e., an increase in τ_w) reduces incentives to invest in education—is standard and straightforward.

3.3 Bequest taxation and efficiency

We now turn to the question whether bequest taxation can lead to a Pareto-improvement. In the remainder of this section, we consider the impact on utility of introducing a small tax on bequests levied from period 2 onwards and announced in period 1. As already indicated, the wage tax rate τ_w is kept constant throughout this analysis. Note that this is a rather demanding test for the desirability of a bequest tax as we could alternatively assume that at the same time the wage tax could be lowered when marginally increasing τ_b . We find (as proven, like all subsequent formal results, in the Appendix)

Lemma 1 *By levying a small bequest tax from period 2 onwards, (i) the currently middle-aged generation unambiguously gains (is unaffected) if $\tau_w > (=) 0$, and (ii) a Pareto-improvement occurs if and only if*

$$\frac{1 + \bar{r} + \tau_w}{1 - \tau_w} \frac{\partial e^*}{\partial \tau_b} \Big|_{\tau_b=0} + \frac{\partial b_{t+1}}{\partial \tau_b} \Big|_{\tau_b=0} \geq 0 \quad (14)$$

for $t \geq 1$.¹²

For the initially middle-aged generation, income (I_1) is not affected by the bequest tax from period 2 onwards. (Consequently, also utility of the initially old generation is unaffected.) Given that human capital investment is distorted ($\tau_w > 0$) utility of members of the initially middle-aged generation increases after introducing a small bequest tax τ_b . This is because human capital investment rises (Corollary 1), which positively affects their offspring's income. Regarding the generations born *after* the initially middle-aged, two potentially counteracting effects are relevant. The first one is again the unambiguously positive impact of τ_b on $e^*(\tau_b, \tau_w)$, according to Corollary 1. However, the effect on welfare also depends on how the bequests received from parents are affected. Thus, if the amount of intergenerational transfers declines, utility may decline after introducing bequest taxation despite the positive effect from an increase in human capital investments. Hence, a priori, it is not clear whether

¹²Note that evaluating at $\tau_b = 0$ means that no revenue is generated from bequest taxation.

bequest taxes can raise welfare of all generations. The positive impact of bequest taxation on human capital formation has to be weighted against the potential reduction in bequests.

When the optimal bequest tax is positive, its intuition can be summarized as follows. In absence of a bequest tax, a positive tax on labor distorts the composition of intergenerational transfers in favor of bequests. Thus, parents will invest too little in their children's education. To reduce this distortion in educational investment, the government may levy a bequest tax.¹³ Starting from a zero tax rate on financial bequests, introducing a bequest tax—although generating a distortion in the level of bequests—also alleviates the distortion in the composition of intergenerational transfers. At least a small positive tax on bequests would be optimal as the new distortion it generates is of second-order relative to the initial distortion it alleviates.

As general conclusions are difficult to obtain, we attempt to gain insight into this issue from an example which allows explicit analytical solutions. From now on, we consider utility specifications

$$u_2(c) = u_3(c) = \ln c \quad \text{and} \quad v(I) = \ln(I - \chi), \quad (15)$$

where $\chi > 0$ may be interpreted as “subsistence income” of children from the perspective of parents. It is a measure of the strength of the bequest motive. To simplify further, let us also employ the standard specification

$$\beta(1 + \bar{r}) = 1. \quad (16)$$

Moreover, let us define

$$\begin{aligned} \Gamma^*(\tau_b, \tau_w) &\equiv (1 + \beta)\chi - (\beta + \tau_b)(\bar{w}h(e^*(\tau_b, \tau_w)) - G) \\ &\quad - (1 - \tau_b)e^*(\tau_b, \tau_w), \end{aligned} \quad (17)$$

$$\begin{aligned} \Gamma_0(\tau_b, \tau_w) &\equiv (1 + \beta)\chi + (\beta + \tau_b)G - (1 + \beta)\bar{w}h(e^*(\tau_b, \tau_w)) \\ &\quad + (1 - \tau_b)[\bar{w}h(e_0) - e^*(\tau_b, \tau_w)]. \end{aligned} \quad (18)$$

Note that both expressions are positive if χ is sufficiently large, which is assumed for the next result.

Lemma 2 *Under specifications (15) and (16), if $\Gamma^* > 0$ and $\Gamma_0 > 0$, then the evolution of bequests is characterized by*

$$b_2 = \frac{\Gamma_0(\tau_b, \tau_w)}{1 + \beta + \beta(1 - \tau_b)} + c(\tau_b)b_1 \equiv B_0(b_1; \tau_b, \tau_w) \quad (19)$$

and for $t \geq 1$,

$$b_{t+2} = \frac{\Gamma^*(\tau_b, \tau_w)}{1 + \beta + \beta(1 - \tau_b)} + c(\tau_b)b_{t+1} \equiv B^*(b_{t+1}; \tau_b, \tau_w), \quad (20)$$

¹³Note that we do not allow for positive externalities of human capital formation (which could generate endogenous growth). Rather, the only distortion of educational investments comes from wage taxation. Assuming instead that positive externalities from education exist would make a positive tax on bequests even more desirable.

where

$$c(\tau_b) \equiv \frac{1 - \tau_b}{1 + \beta + \beta(1 - \tau_b)} < 1. \quad (21)$$

Thus, intergenerational transfers converge to steady state level

$$b^*(\tau_b, \tau_w) \equiv \frac{\Gamma^*(\tau_b, \tau_w)}{2\beta + \tau_b(1 - \beta)} > 0. \quad (22)$$

The assumptions in Lemma 2 thus imply that a unique and stable steady state with a positive amount of bequest exists. In order to examine the dynamic process and the welfare implications of introducing a bequest tax, we suppose that the economy is initially in a steady state with no bequest taxation ($\tau_b = 0$). That is, defining revenue from wage income taxation as $R_w(\tau_b, \tau_w) \equiv \tau_w \bar{w} h(e^*(\tau_b, \tau_w))$, we set the wage tax rate at $\tau_w = \tau_w^0$ as given by $R_w(0, \tau_w^0) = G$; hence, we have initial conditions $e_0 = e^*(0, \tau_w^0)$ and $b_1 = b^*(0, \tau_w^0)$. The next result implies that to establish a Pareto-improvement we only need to check whether the introduction of a bequest tax in $t = 1$ benefits the initially young generation (i.e., raises U_1) and the steady state generation (i.e., raises U_t as $t \rightarrow \infty$).

Lemma 3 Assume $e_0 = e^*(0, \tau_w^0)$ and $b_1 = b^*(0, \tau_w^0)$. Under the assumptions of Lemma 2, announcing in period $t = 1$ that a small tax is levied on bequests from period 2 onwards generates an intertemporal Pareto-improvement if and only if condition (14) holds for both $t = 1$ and $t \rightarrow \infty$.

Recall from Lemma 1 that a Pareto-improvement is obtained when the amount of bequest is not reduced too much in response to the introduction of the bequest tax from period 2 onward. Figure 1 shows the evolution of bequests *after* introduction of the bequest tax. Let \hat{b} be the level of bequest such that, when starting at \hat{b} in period 1, bequests immediately jump to the steady state level b^* in period 2. If $b_1 < \hat{b}$, the amount of bequests increases over time from period 2 onward. Thus, if the generation which is middle-aged when the bequest tax is introduced does not reduce bequests b_2 too much, so that generation 1 is made better off, all generations are made better off. That is, if condition (14) holds for $t = 1$, it holds for all $t > 1$ as well. In contrast, if $b_1 > \hat{b}$, bequests decrease over time from period 2 onward, eventually reaching steady state value b^* (point A in Fig. 1). Thus, if b^* is not reduced too much by the bequest tax, also bequests during the transition to the steady state will decline sufficiently little so to leave every generation better off.

To obtain explicit characterizations in what follows, we further specify

$$h(e) = e^{1/2}. \quad (23)$$

Then (13) and (16) imply that

$$\hat{e}^* = \left[\frac{\beta(1 - \tau_w)}{2(1 - \tau_b)} \right]^2, \quad (24)$$

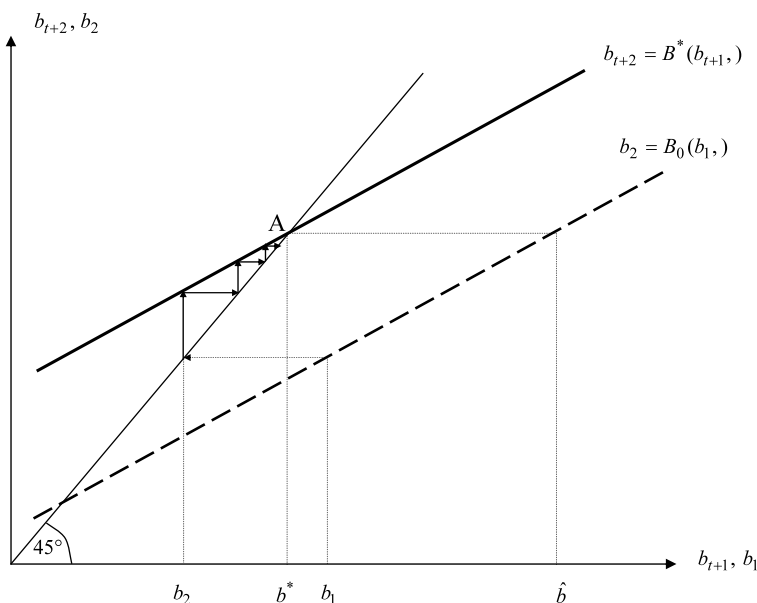


Fig. 1 The evolution of bequests, illustrated for the case $b^* < b_1 < \hat{b}$

where we use the notation $\hat{x} \equiv x/\bar{w}^2$ when a variable (or parameter) x is adjusted by \bar{w}^2 . Using (23), the tax revenue from labor income taxation is given by $R_w = \tau_w \bar{w}^2 (\hat{e}^*)^{1/2}$. Thus, the wage tax rate which finances $\hat{G} = G/\bar{w}^2$ in absence of bequest taxation is given by¹⁴

$$\tau_w^0 = 0.5 \left(1 - \sqrt{1 - \frac{8\hat{G}}{\beta}} \right) \equiv \tilde{\tau}_w^0(\hat{G}). \quad (25)$$

Lemma 4 Under specifications (15), (16), and (23):

- (i) $\Gamma^*(0, \tau_w^0) > 0$ if and only if $\hat{\chi} > 0.75\beta^2(1 - \tilde{\tau}_w^0(\hat{G}))^2/(1 + \beta) \equiv f(\hat{G})$.
- (ii) For both $t = 1$ and $t \rightarrow \infty$, $\partial b_{t+1}/\partial \tau_b|_{\tau_b=0} < 0$.

For Lemma 2, we assumed that $\Gamma^*(\tau_b, \tau_w) > 0$ to obtain a positive steady state level of financial bequests, $b^*(\tau_b, \tau_w)$. Part (i) of Lemma 4 shows that this condition indeed holds for our specifications, given that there is no bequest taxation, if the bequest motive, measured by “adjusted” subsistence income, $\hat{\chi} = \chi/\bar{w}^2$, is sufficiently strong. The relevant threshold, $f(\hat{G})$, is decreasing in τ_w^0 . Thus, if human capital investment is not always more attractive than financial bequests, which may be the case if τ_w^0 is low, then there is a steady state where parents use both ways to bequeath, investing in education and transferring financial wealth. Part (ii) implies that

¹⁴It is easy to see that the economy is on the downward-sloping part of the Laffer curve with respect to revenue from labor income taxation (i.e., $\partial R_w/\partial \tau_w < 0$) if and only if $\tau_w > 0.5$. There are two solutions to $R_w(0, \tau_w^0) = G$. Equation (25) shows the smaller root, as for the larger root $\tau_w^0 > 0.5$ holds.

intergenerational transfers decline in all periods after introduction of a small bequest tax.

We are now ready to study under which circumstances the introduction of a bequest tax, despite its negative effect on the level of bequests leads to a Pareto-improvement.

Proposition 2 Assume $\hat{\chi} > f(\hat{G})$ and initially $e_0 = e^*(0, \tau_w^0)$, $b_1 = b^*(0, \tau_w^0)$. Under specifications (15), (16), and (23), levying a small bequest tax improves welfare of each generation if $\hat{G} \geq \underline{\hat{G}}$, where $\underline{\hat{G}}$ is implicitly given by

$$\frac{2\hat{G}(1+5\beta+8\beta^2)}{\beta(1+\beta)} - 0.5\left(1 + \sqrt{1 - \frac{8\hat{G}}{\beta}}\right) - \frac{4(1-\beta)\hat{\chi}}{\beta^2} = 0. \quad (26)$$

$\underline{\hat{G}}$ is increasing in $\hat{\chi}$.

Proposition 2 suggests that a positive bequest tax may be efficiency-enhancing even if not used to lower the wage tax. This desirability of a distorting new tax arises in an initial steady state with labor income taxation only,¹⁵ if the public expenditure level and, therefore, the initial wage tax rate are sufficiently high ($\tau_w^0 > \tilde{\tau}_w(\hat{G})$). The reason is that for $\tau_w^0 > \tilde{\tau}_w(\hat{G})$ the human capital investment decision is severely distorted by labor income taxation and, therefore, the incentive to raise educational investment may dominate the effect from a reduction in the amount of bequests on utility. For instance, suppose $\beta = 0.9$ and $\hat{\chi} = 0.5$. In this case, $\tilde{\tau}_w(\hat{G}) \cong 20.8\%$. If $\beta = 0.9$ and $\hat{\chi} = 1$, then $\tilde{\tau}_w(\hat{G}) \cong 26.5\%$. This shows that bequest taxation may be optimal at rather moderate wage taxation, even when leaving the wage tax unchanged. Note that for given $\hat{\chi}$, the adjusted threshold for public expenditure, $\underline{\hat{G}}$, is independent of \bar{w}^2 and, therefore, $\tilde{\tau}_w(\hat{G})$ is independent of the wage level. A higher $\hat{\chi}$ raises threshold expenditure $\underline{\hat{G}}$ since it positively affects the equilibrium level of financial bequest (which is distorted by bequest taxation) for any period, while leaving educational investment unchanged.

4 Optimal tax structure

In the previous section, we proved that introducing a small bequest tax may raise welfare of all generations, even if the wage tax rate is kept constant. In this section, we analyze what would be an optimally chosen combination of wage and bequest taxation (in absence of lump-sum transfers), with a given government revenue requirement. To abstract from transition issues, we focus on maximizing the utility of steady-state generations,¹⁶ assuming that the government budget is balanced in each period. That is, $\tau_w \bar{w}h(e^*) + \tau_b b^* = G$ holds.

¹⁵Assumption $\hat{\chi} > f(\hat{G})$ implies $b_1 > 0$, according to Lemma 2 and part (i) of Lemma 4.

¹⁶As shown in the proof of Proposition 2, introducing a small bequest tax leads to a Pareto improvement if it benefits the steady state generation. This suggests that all generations are made better off under the optimal tax mix for steady state generations, compared to a situation where there is only wage taxation.

According to (3), (4), (12), (7), and (8), the social planner's objective function is then given by

$$U^* \equiv u_2(\bar{w}h(e^*) + b^* - G - s^* - e^*) + \beta u_3((1 + \bar{r})s^* - b^*) + \beta v(\bar{w}h(e^*) + b^* - G). \quad (27)$$

Under specifications (15), (16), and (23), up to an additive constant this is equivalent to

$$\hat{U}^* \equiv \ln(\sqrt{\hat{e}^*} + \hat{b}^* - \hat{G} - \hat{s}^* - \hat{e}^*) + \beta \ln\left(\frac{\hat{s}^*}{\beta} - \hat{b}^*\right) + \beta \ln(\sqrt{\hat{e}^*} + \hat{b}^* - \hat{G}). \quad (28)$$

From (22),

$$\hat{b}^* = \frac{(1 + \beta)\hat{\chi} - (\beta + \tau_b)(\sqrt{\hat{e}^*} - \hat{G}) - (1 - \tau_b)\hat{e}^*}{2\beta + \tau_b(1 - \beta)}. \quad (29)$$

Moreover, using first-order condition (9) together with (7), (8), and (12), it is easy to see that

$$\hat{s}^* = \frac{\beta}{1 + \beta}(\sqrt{\hat{e}^*} + 2\hat{b}^* - \hat{G} - \hat{e}^*). \quad (30)$$

Together with (24), (28)–(30) show that the optimal solution to the social planner's problem is independent of wage rate \bar{w} , provided that χ is proportional to \bar{w}^2 .

Table 1 shows numerical results for the optimal tax rates $(\tau_w^{\text{opt}}, \tau_b^{\text{opt}}) \equiv \arg \max_{(\tau_w, \tau_b)} \hat{U}^*$ s.t. $\tau_w(\hat{e}^*)^{1/2} + \tau_b\hat{b}^* = \hat{G}$, for $\hat{\chi} \in \{0.5, 1\}$ and $\beta = 0.9$, conditional on the (adjusted) government expenditure level, \hat{G} . We also report the ratio of financial bequests to wage income, $\hat{b}^*/\sqrt{\hat{e}^*}$, under both regimes without bequest taxation (i.e., $(\tau_w, \tau_b) = (\tau_w^0, 0)$) and under the optimal tax mix, $(\tau_w^{\text{opt}}, \tau_b^{\text{opt}})$; we use notation $(\hat{b}^*/\sqrt{\hat{e}^*})^0$ and $(\hat{b}^*/\sqrt{\hat{e}^*})^{\text{opt}}$, respectively. Moreover, we report the implied ratio of government expenditure as fraction of the gross national product (GNP), initially and under the optimal tax mix, denoted by g^0 and g , respectively.¹⁷

Our numerical results again suggest that the optimal bequest tax rate is generally positive when the (adjusted) government revenue requirement, \hat{G} , is sufficiently high. This is consistent with the intuition of Proposition 2: Using bequest taxes can raise efficiency when an excessive use of a wage tax would be too distorting. With a low revenue requirement, however, it is optimal to moderately tax wages and use tax revenue to subsidize bequests. Moreover, also when \hat{G} is high, the optimal bequest tax rate is significantly lower than the wage tax rate. The intuition for these results is the following. Investment in human capital exhibits decreasing returns to scale, while financial markets provide constant returns to scale. At the same time as taxing wages reduces investment in human capital, it also increases the rate of return to marginal investment. This partly counteracts the distortion created by the tax wedge. When the government chooses tax rates to balance marginal distortions

¹⁷GNP equals the sum of wage income and interest income from savings, $\text{GNP} = \bar{w}h(e^*) + \bar{r}s^*$. Our specifications imply that $G/\text{GNP} = \hat{G}/(\sqrt{\hat{e}^*} + \frac{1-\beta}{\beta}\hat{s}^*)$.

Table 1 Social optimum in the steady state—numerical results (in percent for τ_w , τ_b , and g); $\beta = 0.9$

\hat{G}	τ_w^0	τ_w^{opt}	τ_b^{opt}	$(\frac{\hat{b}^*}{\sqrt{e^*}})^0$	$(\frac{\hat{b}^*}{\sqrt{e^*}})^{\text{opt}}$	g^0	g
0	0	5.6	−8.4	0.42	0.66	0	0
0.02	4.7	8.1	−4.6	0.52	0.66	4.3	4.6
0.04	9.9	10.5	−0.8	0.63	0.65	9.0	9.1
0.06	15.8	13.0	2.9	0.76	0.65	14.3	13.6
0.08	23.1	15.4	6.5	0.95	0.65	20.6	18.0
0.10	33.3	17.8	10.0	1.26	0.65	28.9	22.4

(a) $\hat{\chi} = 0.5$

\hat{G}	τ_w^0	τ_w^{opt}	τ_b^{opt}	$(\frac{\hat{b}^*}{\sqrt{e^*}})^0$	$(\frac{\hat{b}^*}{\sqrt{e^*}})^{\text{opt}}$	g^0	g
0	0	6.1	−3.3	1.60	1.87	0	0
0.02	4.7	7.4	−1.3	1.75	1.87	3.8	4.0
0.04	9.9	8.6	0.6	1.93	1.87	8.0	7.9
0.06	15.8	9.8	2.5	2.16	1.87	12.7	11.8
0.08	23.1	11.0	4.3	2.74	1.87	18.0	14.5
0.10	33.3	12.2	6.1	3.02	1.88	24.9	17.9

(b) $\hat{\chi} = 1$

from collecting any given revenue, it is optimal to distort human capital investment relatively more. For the same reason, when G is low, taxing the return to education and subsidizing bequests may improve the welfare of the steady-state generations by encouraging parents to transfer in aggregate more resources to their children. Also note that optimal tax rates are nonzero even in the case where $\hat{G} = 0$. Why an optimal tax on bequests could be negative (and, therefore, the optimal wage tax positive) even when there is no public sector? The answer relies on intergenerational externalities that intergenerational transfers generate. Each generation chooses the level of transfers to the subsequent generation taking into account only its own joy-of-children-receiving. Subsidizing financial bequests encourages more giving while taxing wages introduces a negative distortion. A priori, there is no reason why the social planner should abstain doing the former in order to avoid the latter, given that returns to education are diminishing.

We can also see from Table 1 how the size of financial bequests relative to wage income depends on public expenditure. In the absence of bequest taxation, increasing the tax rate on labor income results in parents transferring relatively more resources through bequests. In the examples we report, in the absence of bequest taxes, the size of bequests varies between 42 and 126% of the lifetime wage income with $\hat{\chi} = 0.5$, and between 160 and 302% with $\hat{\chi} = 1$. When the bequest tax rate is set optimally, the range is 65 to 66% with $\hat{\chi} = 0.5$ and 187 to 188% with $\hat{\chi} = 1$. This suggests that optimal taxation stabilizes the composition of intergenerational transfers when the general level of public expenditures changes.¹⁸

¹⁸The same holds with respect to educational investment and, therefore, with respect to GDP, $Y^* = h(e^*)f(\bar{k})$. These insights are also valid for the extensions in the coming section.

According to the last two columns, for sufficiently high public spending levels, GNP is clearly higher under the optimal tax mix than in the case where bequest taxation is not available, which implies $g < g^0$. Moreover, comparing panels (a) and (b) of Table 1 suggests that a higher desired income level for children, $\hat{\chi}$, reduces the ratio of public expenditure to GNP as well as the level of g^0 above which the optimal bequest tax becomes positive. The results suggest that bequest tax rates should be positive already at moderate levels of the ratio of government expenditure to GNP.

5 Discussion and extensions

5.1 Optimal tax mix with education subsidies

So far, we have abstracted from the instrument of education subsidies for stimulating educational investment. Partly, this may be justified because human capital investments are often unobservable to tax authorities, in a similar manner as the optimal tax literature typically posits that work effort is not observable.¹⁹ We will now extend our analysis to allow for subsidies on partly observable education expenditure. For this purpose, suppose each unit of investment in education, e , is subsidized by a constant rate τ_e . However, the effective subsidy rate is restricted to $\tau_e \leq \bar{\tau}_e < 1$. For instance, if the government can levy a subsidy rate of up to 100% for observed education expenditure and only 50% of total expenditure e is observable, then $\bar{\tau}_e = 0.5$.

Individual budget constraints at date $t + 1$ and $t + 2$ are again given by (7) and (8) where now net income is $I_{t+1} = (1 - \tau_w)\bar{w}h(e_t) + (1 - \tau_b)b_{t+1} + \tau_e e_{t+1}$. Whereas first-order conditions associated with the individual optimization problem (9) and (11) remain the same, (10) changes to

$$\frac{u'_2(c_{2,t+1})}{\beta v'(I_{t+2})} = \frac{1 - \tau_w}{1 - \tau_e} \bar{w}h'(e_{t+1}). \quad (31)$$

Thus, (13) becomes

$$\frac{1 - \tau_w}{1 - \tau_e} \bar{w}h'(e^*) = (1 - \tau_b)(1 + \bar{r}), \quad (32)$$

i.e., a higher subsidy rate τ_e raises educational investment. The government budget constraint now includes payments for the education subsidy; it is given by $\tau_w \bar{w}h(e_t) + \tau_b b_{t+1} = G + \tau_e e_{t+1}$ (like in the previous section, we focus on $T_{t+1} = 0$). Net income is thus still given by (12).

Before we analyze the optimal tax mix cum education subsidy, suppose first that there is no bequest taxation. If the social planner is constrained to $\tau_b = 0$, then the optimal tax policy is to set $\tau_w = \tau_e = \tau$ such that G is financed, irrespective of our specification of functional forms. In this case, there is no distortion of educational

¹⁹Trostel (1993) estimates that a substantial fraction of the costs of education are nonverifiable, even when abstracting from any effort costs. In their paper on human capital investment and capital income taxation, Jacobs and Bovenberg (2005) find that taxing capital income is optimal with subsidies to human capital investment when at least a share of these investments is nonverifiable.

Table 2 Social optimum in the steady state with education subsidy—numerical results; $\beta = 0.9$, $\hat{\chi} = 1$

\hat{G}	$\bar{\tau}_e$	τ_w^{opt}	τ_b^{opt}	g
0	0.25	25.9	-7.9	0
0.05	0.25	27.6	-2.8	9.7
0.1	0.25	30.3	2.0	19.3
0.15	0.25	33.0	6.7	28.8
0	0.5	45.9	-13.7	0
0.05	0.5	48.0	-8.3	9.5
0.1	0.5	50.3	-3.1	19.1
0.15	0.5	52.4	2.0	28.5

Note: $\bar{\tau}_e$ is also the optimal educational subsidy for social planner's problem (34)

investment. Hence, if $\tau_w = \tau_e = \tau$ under $\tau_b = 0$, then taxation has the same impact as lump-sum taxation. We will now show that in steady state $\tau_w = \tau_e = \tau$ may not be optimal, however, if taxing (or subsidizing) bequests is feasible.

Using $\beta(1 + \bar{r}) = 1$ and $h(e) = e^{1/2}$, we obtain

$$\hat{e}^* = \left[\frac{\beta(1 - \tau_w)}{2(1 - \tau_b)(1 - \tau_e)} \right]^2. \quad (33)$$

Again focusing on the steady state and using specification (15), the objective function of the social planner remains \hat{U}^* , as defined in (28), where \hat{b}^* and \hat{s}^* are still given by (29) and (30), respectively.²⁰ Under partial unobservability of educational investments, the social planner's problem now is

$$\begin{aligned} \max_{\tau_b, \tau_w, \tau_e} \quad & \hat{U}^* \\ \text{s.t.} \quad & \tau_w(\hat{e}^*)^{1/2} + \tau_b \hat{b}^* = \hat{G} + \tau_e \hat{e}^*, \quad \tau_e \leq \bar{\tau}_e. \end{aligned} \quad (34)$$

As a first robust numerical result, we find that it is optimal to fully subsidize education, i.e., under partial unobservability of educational investment, to choose the maximal effective education subsidy, $\bar{\tau}_e$. In Table 2, we report the optimal tax mix (τ_w, τ_b) as well as the implied ratio of public spending to GNP, g , conditional on \hat{G} , for $\bar{\tau}_e \in \{0.25, 0.5\}$, $\beta = 0.9$ and $\hat{\chi} = 1$.

The results suggest that the bequest tax rate is generally nonzero, being negative if \hat{G} is low and positive if \hat{G} is high. The threshold level of public spending which calls for positive taxation of bequests clearly depends on the fraction of educational investment which is observable, as generally observed education spending should be fully subsidized. In fact, education subsidies should be paid even when the government revenue requirement is zero, implying a rather high wage tax to finance both education subsidies and bequest subsidies.

Thus, it is clearly the case that the availability of education subsidies significantly weakens the beneficial role of positive bequest taxation and even calls for higher bequest subsidies in case of low revenue requirements. However, there may be reasons

²⁰Recall that first-order conditions (9) and (11) hold with and without education subsidy and that net income is still given by (12).

why wage taxes cannot be raised to a level which finances both an optimal education subsidy and a subsidy on bequests, like international labor mobility. Suppose that for exogenous reasons there is an upper bound to labor taxation, $\bar{\tau}_w$, but no limit on education subsidies. One can show that in this case, there is an interior solution $\tau_e \in (0, 1)$, and again we have $\tau_b^{\text{opt}} < 0$ if \hat{G} is low and $\tau_b^{\text{opt}} > 0$ if \hat{G} is high (results available upon request). This suggests that a government which operates under constraints for either wage taxation or subsidization of education should well consider taxing bequests if the revenue requirement is high.

5.2 Social optimum without the altruistic component

In the optimal tax literature, the altruistic component is sometimes laundered out in the social planner's objective function. How does the optimal tax mix change as a consequence?²¹ Replacing \hat{U}^* by

$$\ln(\sqrt{\hat{e}^*} + \hat{b}^* - \hat{G} - \hat{s}^* - \hat{e}^*) + \beta \ln\left(\frac{\hat{s}^*}{\beta} - \hat{b}^*\right)$$

and leaving out education subsidies, we obtain numerical results for the steady state as summarized in Table 3.

As already discussed in Sect. 4, nonzero tax rates in the case where there is no public revenue requirement are called for only if the social planner takes intergenerational externalities due to altruism of parents into account. A striking contrast to Sect. 4 is that now the optimal bequest tax rate, τ_b^{opt} , is always positive and not substantially lower than the wage tax τ_w^{opt} .

It is important to note, however, that allowing the social planner to choose an education subsidy, τ_e , like in the previous subsection, leaves the basic insights of our discussion above unchanged. That is, education should be subsidized at rate $\bar{\tau}_e$ and an efficiency-enhancing role of positive bequest taxes requires that either a substantial fraction of educational investment is unobservable (implying that $\bar{\tau}_e$ is low) or that there is a binding constraint for the wage tax ($\bar{\tau}_w$). (Results are available upon request.)

Table 3 Social optimum in the steady state without altruistic component in objective function—numerical results; $\beta = 0.9$, $\hat{\chi} = 1$, $\tau_e = 0$

\hat{G}	τ_w^{opt}	τ_b^{opt}	g
0	0	0	0
0.02	2.6	1.2	3.6
0.04	4.2	2.9	7.5
0.06	5.8	4.6	11.3
0.08	7.3	6.3	15.0
0.1	8.9	7.9	18.6
0.12	10.4	9.6	22.5
0.14	11.9	11.2	26.2

²¹We are grateful to an anonymous referee for suggesting to address this question.

6 Conclusion

Altruistic parents may transfer resources to their offspring by providing education and by leaving bequests. Parental altruism is often seen as an argument against bequest taxation, the reason being that bequest taxation would distort the accumulation of capital intergenerationally in the same way as capital income taxation would distort consumption profile and savings over the individual life cycle. In this paper, we show that this intuition needs no longer hold true in the presence of education and wage taxation. Wage taxes reduce the rate of return that children receive on parental investments in education. This induces parents, who value the after-tax resources that their children receive, to reduce investment in education, and leave bequests instead. We show that a small bequest tax may improve efficiency in an overlapping-generations framework with only intended bequests, even when the wage tax remains unchanged. This is because the bequest tax may mitigate the distortion of educational investment caused by wage taxation.

In addition to deriving a general criterion for the desirability of a small bequest tax when the wage tax rate is left unchanged, we also analyze what would be an optimal mix of wage taxes and bequest taxes with given government revenue requirement. Certain clear patterns emerge. First of all, the optimal bequest tax is generally positive when the government revenue requirement is sufficiently high, although always lower than the wage tax rate. This may hold true even if governments are also able to levy education subsidies, provided that these are limited due to the partial unobservability of education expenditure. In any case, our results suggest that bequest taxation may be efficiency-enhancing if education subsidies are low. Moreover, the case for taxing bequests is strengthened if the social planner does not take into account the altruistic motive of parents.

Our results have certain surprising implications for the US debate on estate taxation, which centers around the conventional wisdom that taxation of intended bequests gives rise to a typical equity-efficiency trade-off (see Gale and Slemrod 2001, for a review of the debate). Currently, descendants of only 2% of Americans who die pay estate taxes. Even proponents of the estate tax are willing to raise the exempted amount further. We find that this policy, while popular, need not be optimal even from an efficiency point of view. It might well be optimal to tax also smaller bequests, possibly at a relatively low rate, and use the tax revenue to lower wage taxes. Such policy would boost the incentives of altruistic parents among the currently exempted 98% of population to transfer resources to their children more through education.

In this paper, we deliberately abstracted from distributional issues, in order to highlight the efficiency argument in favor of bequest taxation. An important topic for future research would be to introduce heterogeneity with respect to learning abilities and/or initial wealth of households as well as stochastic components in the transmission of abilities from parent to child. For instance, one could study the evolution of the distribution of income and wealth and how it is affected by the tax instruments. Such analyses would call for a computational model with a suitably calibrated distribution of various shocks.

Acknowledgements We are grateful to Josef Falkinger, Vesa Kanninen, Katarina Keller, and seminar participants at the CESifo Area Conference on Public Economics in Munich in April 2005, at the

Copenhagen Business School in May 2005, and the Annual Meeting of German Economists (Verein für Socialpolitik) in Bonn in September 2005 for valuable comments and suggestions. Moreover, we are indebted to Stephanie Bade for excellent research assistance.

Appendix

Proof of Lemma 1 Part (i) is proven first. Note that the currently middle-aged generation is born in $t = 0$. Also note from (12) that their income, I_1 , is initially given, as e_0 and b_1 (the latter depending on both e_0 and s_0) are given. Observing $e_1 = e^*$, we have

$$U_0 = u_2(I_1 - s_1 - e^*) + \beta u_3((1 + \bar{r})s_1 - b_2) + \beta v(\bar{w}h(e^*) + b_2 - G), \quad (\text{A.1})$$

according to (3), (4), (12), (7), and (8). Differentiating with respect to τ_b , using (by applying the envelope theorem) both $u'_2(c_{2,1}) = (1 + \bar{r})\beta u'_3(c_{3,2})$ and $v'(I_2) = u'_3(c_{3,2})/(1 - \tau_b)$, according to (9) and (11), and finally, using $\bar{w}h'(e^*)/(1 - \tau_b) = (1 + \bar{r})/(1 - \tau_w)$, according to (13), leads to

$$\frac{\partial U_0}{\partial \tau_b} = \beta u'_3(c_{3,2}) \left[(1 + \bar{r}) \frac{\tau_w}{1 - \tau_w} \frac{\partial e^*}{\partial \tau_b} + \frac{\tau_b}{1 - \tau_b} \frac{\partial b_2}{\partial \tau_b} \right]. \quad (\text{A.2})$$

Thus, $\partial U_0 / \partial \tau_b|_{\tau_b=0} > (=) 0$ if $\tau_w > (=) 0$, according to Corollary 1. This confirms part (i).

We now turn to part (ii). Utility of generation $t \geq 1$ is

$$U_t = u_2(\bar{w}h(e_t) + b_{t+1} - G - s_{t+1} - e_{t+1}) + \beta u_3((1 + \bar{r})s_{t+1} - b_{t+2}) + \beta v(\bar{w}h(e_{t+1}) + b_{t+2} - G). \quad (\text{A.3})$$

Taking into account that $e_{t+1} = e^*$ for all $t \geq 0$ stays the same, differentiating and using first-order condition (10) w.r.t. s_{t+1} gives

$$\frac{\partial U_t}{\partial \tau_b} = u'_2 \bar{w}h' \frac{\partial e^*}{\partial \tau_b} + u'_2 \frac{\partial b_{t+1}}{\partial \tau_b} - u'_2 \frac{\partial e^*}{\partial \tau_b} - \beta u'_3 \frac{\partial b_{t+2}}{\partial \tau_b} + \beta v' \bar{w}h' \frac{\partial e^*}{\partial \tau_b} + \beta v' \frac{\partial b_{t+2}}{\partial \tau_b}. \quad (\text{A.4})$$

Using again the first-order conditions associated with the individual optimization problem, this simplifies as

$$\begin{aligned} \frac{\partial U_t}{\partial \tau_b} &= (1 + \bar{r})\beta u'_3 \bar{w}h' \frac{\partial e^*}{\partial \tau_b} + (1 + \bar{r})\beta u'_3 \frac{\partial b_{t+1}}{\partial \tau_b} - (1 + \bar{r})\beta u'_3 \frac{\partial e^*}{\partial \tau_b} \\ &\quad - \beta u'_3 \frac{\partial b_{t+2}}{\partial \tau_b} + \beta \frac{u'_3}{1 - \tau_b} \bar{w}h' \frac{\partial e^*}{\partial \tau_b} + \beta \frac{u'_3}{1 - \tau_b} \frac{\partial b_{t+2}}{\partial \tau_b}. \end{aligned} \quad (\text{A.5})$$

We obtain condition (14) by using (13), factoring out $\beta(1 + \bar{r})u'_3$ and evaluating at $\tau_b = 0$. \square

Proof of Lemma 2 Substituting $c_{2,t+1} = I_{t+1} - s_{t+1} - e_{t+1}$ and $c_{3,t+2} = (1 + \bar{r})s_{t+1} - b_{t+2}$ from (7) and (8), respectively, into (9), and using $u_2(c) = u_3(c) = \ln c$, leads to

$$s_{t+1} = \frac{\beta(1 + \bar{r})(I_{t+1} - e_{t+1}) + b_{t+2}}{(1 + \bar{r})(1 + \beta)} \quad (\text{A.6})$$

for all $t \geq 0$. Moreover, substituting $c_{3,t+2} = (1 + \bar{r})s_{t+1} - b_{t+2}$ from (8) into (11), and using $u_3(c) = \ln c$ and $v(I) = \ln(I - \chi)$ yields $I_{t+2} - \chi = (1 - \tau_b)[(1 + \bar{r})s_{t+1} - b_{t+2}]$. Substituting (12) and (A.6) into this expression and using both $e_{t+1} = e^*$ for $t \geq 0$ and $\beta(1 + \bar{r}) = 1$ from specification (16) implies that bequests evolve over time according to (19) and (20). As $c(\tau_b) < 1$, the dynamic process governing the evolution of bequests is stable. Finally, setting $b_{t+1} = b_{t+2} \equiv b^*$ in (20), observing (21) and solving for b^* gives us (22). This concludes the proof. \square

Proof of Lemma 3 If $\tau_b > 0$, then $e_0 < e^*(\tau_b, \tau_w^0)$, according to Corollary 1. Consequently, we have $\Gamma_0(\tau_b, \tau_w) < \Gamma^*(\tau_b, \tau_w)$, according to (17) and (18), and thus, $B_0(b; \cdot) < B^*(b; \cdot)$, according to (19) and (20). Figure 1 depicts $b_2 = B_0(b_1; \cdot)$ as dashed line and $b_{t+2} = B^*(b_{t+1}; \cdot)$ as solid line for $\tau_b > 0$. The steady state level of bequest with $\tau_b > 0$, b^* , is given by point A. Let \hat{b} be given by $B_0(\hat{b}; \cdot) = b^*$. Now if $b_1 < \hat{b}$ as in Fig. 1, then $b_2 < b^*$ and, for all $t \geq 1$, b_{t+2} increases over time to b^* . In this case, if condition (14) holds for $t = 1$, it also holds for all $t > 1$. If $b_1 = \hat{b}$, then $b_2 = b_{t+2} = b^*$ for all $t \geq 1$. Finally, if $b_1 > \hat{b}$, then $b_2 > b^*$ and, for all $t \geq 1$, b_{t+2} decreases over time to b^* . In this case, if condition (14) holds for $t \rightarrow \infty$ (i.e., for $b_{t+1} = b^*$), it also holds for all $t \geq 1$. This concludes the proof. \square

Proof of Lemma 4 First, note that

$$b^*(0, \tau_w^0) = \frac{(1 + \beta)\chi - \beta(1 - \tau_w^0)\bar{w}h(e^*(0, \tau_w^0)) - e^*(0, \tau_w^0)}{2\beta}, \quad (\text{A.7})$$

according to (17), (22), and (by definition of τ_w^0) $\bar{w}h(e^*(0, \tau_w^0)) - G = (1 - \tau_w^0)\bar{w}h(e^*(0, \tau_w^0))$. Using $h(e) = e^{1/2}$ and substituting $e^*(0, \tau_w) = [\beta(1 - \tau_w)\bar{w}]^2/4$ from (24) into (A.7) leads to

$$b^*(0, \tau_w^0) = \frac{(1 + \beta)\chi - 3e^*(0, \tau_w^0)}{2\beta}. \quad (\text{A.8})$$

Part (i) follows by using (24). To confirm part (ii), take partial derivatives of (22) and (19) with respect to τ_b , by using (17) and (18), respectively. By evaluating the resulting expressions at $(\tau_b, \tau_w) = (0, \tau_w^0)$ and noting that

$$\left. \frac{\partial e^*(\tau_b, \tau_w)}{\partial \tau_b} \right|_{\tau_b=0} = 2e^*(0, \tau_w), \quad (\text{A.9})$$

according to (24), we obtain

$$\begin{aligned} & \left. \frac{\partial b^*(\tau_b, \tau_w^0)}{\partial \tau_b} \right|_{\tau_b=0} \\ &= - \frac{(1 - \tau_w^0) \bar{w} h(e^*(0, \tau_w^0)) + (2 \frac{2-\tau_w^0}{1-\tau_w^0} - 1) e^*(0, \tau_w^0) + (1 - \beta) b^*(0, \tau_w^0)}{2\beta} \end{aligned} \quad (\text{A.10})$$

and

$$\begin{aligned} & \left. \frac{\partial B_0(b_1; \tau_b, \tau_w^0)}{\partial \tau_b} \right|_{\tau_b=0} \\ &= - \frac{(1 - \tau_w^0) \bar{w} h(e^*(0, \tau_w^0)) + (2 \frac{1+\beta(2-\tau_w^0)}{\beta(1-\tau_w^0)} - 1) e^*(0, \tau_w^0) + (1 - \beta) b_1}{1 + 2\beta}. \end{aligned} \quad (\text{A.11})$$

Both derivatives are negative. This concludes the proof. \square

Proof of Proposition 2²² According to Lemma 3 and the assumptions of Proposition 2, a Pareto-improvement is reached if

$$\Omega^* \equiv \frac{\frac{1}{\beta} + \tau_w^0}{1 - \tau_w^0} \left. \frac{\partial e^*(\tau_b, \tau_w^0)}{\partial \tau_b} \right|_{\tau_b=0} + \left. \frac{\partial b^*(\tau_b, \tau_w^0)}{\partial \tau_b} \right|_{\tau_b=0} \geq 0 \quad (\text{A.12})$$

and

$$\Omega_0 \equiv \frac{\frac{1}{\beta} + \tau_w^0}{1 - \tau_w^0} \left. \frac{\partial e^*(\tau_b, \tau_w^0)}{\partial \tau_b} \right|_{\tau_b=0} + \left. \frac{\partial B_0(b^*(0, \tau_w^0); \tau_b, \tau_w^0)}{\partial \tau_b} \right|_{\tau_b=0} \geq 0 \quad (\text{A.13})$$

simultaneously hold. It is tedious but straightforward to show that substituting (A.9) and (A.10) into (A.12) and using $b_1 = b^*(0, \tau_w^0)$ as given in (A.8) implies

$$\Omega^* = \frac{1 + \beta}{4\beta^2} \left(\frac{e_0}{1 - \tau_w^0} [q\tau_w^0 - 1] - (1 - \beta)\chi \right), \quad (\text{A.14})$$

where $q \equiv (1 + 5\beta + 8\beta^2)/(1 + \beta)$. Similarly, substituting (A.9) and (A.11) into (A.13) and using (A.8) implies

$$\Omega_0 = \frac{1 + \beta}{2\beta(1 + 2\beta)} \left(\frac{e_0}{1 - \tau_w^0} [(1 + 8\beta)\tau_w^0 - 1] - (1 - \beta)\chi \right). \quad (\text{A.15})$$

Note that $q < 1 + 8\beta$. Thus, if $\Omega^* \geq 0$, then $\Omega_0 > 0$. Substituting $e_0 = [\beta(1 - \tau_w)\bar{w}/2]^2$ into (A.14), we find that $\Omega^* \geq 0$ if and only if $\beta^2(1 - \tau_w^0) \times$

²²A more detailed proof is presented in a technical appendix, available from the authors upon request.

$(q\tau_w^0 - 1) \geq 4(1 - \beta)\hat{\chi}$. Note that the left-hand side of the condition is increasing in τ_w^0 whenever $\tau_w^0 \leq 0.5$. It is, therefore, increasing in G . Furthermore, substituting τ_w^0 from (25) reveals that the condition holds with equality if $G = \underline{G}$. This concludes the proof. \square

References

- Andreoni, J. (1989). Giving with impure altruism: Applications to charity and ricardian equivalence. *Journal of Political Economy*, 97, 1447–1458.
- Barro, R. J. (1974). Are government bonds net wealth? *Journal of Political Economy*, 82, 1095–1117.
- Blinder, A. S. (1976). Intergenerational transfers and life cycle consumption. *American Economic Review Papers and Proceedings*, 66, 87–89.
- Blumkin, T., & Sadka, E. (2003). Estate taxation with intended and accidental bequests. *Journal of Public Economics*, 88, 1–21.
- Chamley, C. P. (1986). Optimal taxation of capital income in general equilibrium with infinite lives. *Econometrica*, 54, 607–622.
- Cremer, H., & Pestieau, P. (2001). Non-linear taxation of bequests, equal sharing rules and the trade-off between intra- and inter-family inequalities. *Journal of Public Economics*, 79, 35–53.
- Cremer, H., & Pestieau, P. (2003). *Wealth transfer taxation: A survey* (CESifo Working Paper No. 1061).
- Domeij, D. (2005). Optimal capital taxation and labor market search. *Review of Economic Dynamics*, 8, 623–650.
- Gale, W. G., & Slemrod, J. (2001). Rethinking estate and gift taxation: Overview. In W. Gale, J. Hines & J. Slemrod (Eds.), *Rethinking estate and gift taxation* (pp. 1–64). Washington: Brookings Institution Press.
- Galor, O., & Moav, O. (2004). From physical to human capital accumulation: Inequality in the process of development. *Review of Economic Studies*, 71, 1001–1026.
- Goldin, C., & Katz, L. F. (1998). The origins of technology-skill complementarity. *Quarterly Journal of Economics*, 113, 693–732.
- Gradstein, M., & Justman, M. (1997). Democratic choice of an education system: Implications for growth and income distribution. *Journal of Economic Growth*, 2, 169–183.
- Ishikawa, T. (1975). Family structures and family values in the theory of income distribution. *Journal of Political Economy*, 83, 987–1008.
- Jacobs, B., & Bovenberg, A. L. (2005). *Human capital and optimal positive taxation of capital income* (CEPR Discussion Paper No. 5047).
- Jones, L. E., Manuelli, R. E., & Rossi, P. E. (1993). Optimal taxation in models of endogenous growth. *Journal of Political Economy*, 101, 485–517.
- Jones, L. E., Manuelli, R. E., & Rossi, P. E. (1997). On the optimal taxation of capital income. *Journal of Economic Theory*, 73, 93–117.
- Judd, K. (1985). Redistributive taxation in a perfect foresight model. *Journal of Public Economics*, 28, 59–84.
- Judd, K. (2002). Capital income taxation with imperfect competition. *American Economic Review Papers and Proceedings*, 92, 417–421.
- Kopczuk, W. (2001). *Optimal estate taxation in the steady state*. University of Columbia (mimeo).
- Michel, P., & Pestieau, P. (2004). Fiscal policy in an overlapping generations model with bequest-as-consumption. *Journal of Public Economic Theory*, 6, 397–407.
- Trostel, P. A. (1993). The effects of taxation on human capital. *Journal of Political Economy*, 101, 327–350.