DISCUSSION PAPER SERIES

IZA

IZA DP No. 2277

Pareto-Improving Bequest Taxation

Volker Grossmann
Panu Poutvaara

August 2006

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Pareto-Improving Bequest Taxation

Volker Grossmann
*University of Fribourg, CESifo
and IZA Bonn*

Panu Poutvaara
*University of Helsinki, CESifo
and IZA Bonn*

Discussion Paper No. 2277
August 2006

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
Email: iza@iza.org

IZA Discussion Paper No. 2277
August 2006

ABSTRACT

Pareto-Improving Bequest Taxation*

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. This result holds even if the wage tax rate is held constant when introducing bequest taxation. 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.

JEL Classification: H21, H31, D64, I21

Keywords: bequest taxation, bequests, education, Pareto improvement

Corresponding author:

Panu Poutvaara
Department of Economics
University of Helsinki
P.O. Box 17 (Arkadiankatu 7)
FIN-00014 Helsinki
Finland
E-mail: panu.poutvaara@helsinki.fi

* We are grateful to Josef Falkinger, Vesa Kanniainen, Katarina Keller and seminar participants at the CESifo Area Conference on Public Economics inMunich 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.

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 non-zero 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

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

[2] Judd (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.

[3] Blumkin 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 percent when labor supply is endogenous and there is wage taxation.

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 in the economy is sufficiently high.

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-Barrovian 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 section 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 (linear) tax structure. The last section concludes. All proofs are relegated to an appendix.

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

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

[4] We 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.

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 $\lim_{k\to\infty} f'(k) = 0$ and $\lim_{k\to 0^+} f'(k) = \infty$.[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

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

units of human capital in $t + 1$, where $h(\cdot)$ is a strictly monotonically increasing and

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

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

strictly concave function which fulfills $\lim_{e \to \infty} h'(e) = 0$ and $\lim_{e \to 0^+} h'(e) = \infty$.[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 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]

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

[8]At 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.

[9]Our 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.

As will become apparent in section 4, our "joy-of-children-receiving-income" motivation gives rise to externalities of intergenerational transfers which renders non-zero 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,[10] intergenerational transfers are suboptimal from a social planner's point of view.

2.3 Public Sector

The government has to finance an exogenous expenditure $\bar{G} \geq 0$ 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. There are two reasons for this.

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.[11]

[10]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. Such a bequest motive has been criticized, however, inter alia because it means that individuals act as they would be infinitively-living.

[11]If the results that Chamley, Judd and Jones et al. (1993, 1997) derive in an infinitely-lived agent

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.

Finally, for the equilibrium analysis of the coming section, we assume that for financing $\bar{G}$ the government has to use linear taxes on wages and bequests. There are no other taxes. By this, we follow the tradition by Judd (1985, 2002), Chamley (1986) and Jones et al. (1993, 1997). 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. See, however, section 4 for a discussion of the additional role of education subsidies in our framework. (For tractability reasons, the discussion there is based on a numerical analysis only.)

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

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.

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 pre-tax 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. Thus, disposable income of a member of generation t at date $t+1$ is given by

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

where T_{t+1} denotes a potential lump-sum transfer. The government budget constraint in period $t+1$ is

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

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

$$c_{2,t+1} + s_{t+1} + e_{t+1} = I_{t+1}, \tag{7}$$

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

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

$$\frac{u_2'(c_{2,t+1})}{\beta u_3'(c_{3,t+2})} = 1+\bar{r}, \tag{9}$$

$$\frac{u_2'(c_{2,t+1})}{\beta v'(I_{t+2})} = (1-\tau_w)\bar{w}h'(e_{t+1}), \tag{10}$$

and

$$\frac{u_3'(c_{3,t+2})}{v'(I_{t+2})} = 1 - \tau_b, \tag{11}$$

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

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

in $t + 1$, according to (5) and (6).[12]

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*

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

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.*

According to Proposition 1, the optimal educational investment, e^*, is reached when

[12] Note 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.

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. 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*

$$\frac{1+\bar{r}+\tau_w}{1-\tau_w}\left.\frac{\partial e^*}{\partial \tau_b}\right|_{\tau_b=0}+\left.\frac{\partial b_{t+1}}{\partial \tau_b}\right|_{\tau_b=0}\geq 0 \tag{14}$$

for $t \geq 1$.[13]

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

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

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 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.[14] 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

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

where $\chi > 0$ may be interpreted as "subsistence income" of children from the perspec-

[14]Note 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.

tive of parents. It is a measure of the strength of the bequest motive. To simplify further, let us also employ the standard specification

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

Moreover, let us define

$$\Gamma^*(\tau_b, \tau_w) \equiv (1+\beta)\chi - (\beta+\tau_b)\left(\bar{w}h(e^*(\tau_b,\tau_w)) - \bar{G}\right) - (1-\tau_b)e^*(\tau_b,\tau_w), \tag{17}$$

$$\Gamma_0(\tau_b, \tau_w) \equiv (1+\beta)\chi + (\beta+\tau_b)\bar{G} - (1+\beta)\bar{w}h(e^*(\tau_b,\tau_w)) + (1-\tau_b)\left[\bar{w}h(e_0) - e^*(\tau_b,\tau_w)\right]. \tag{18}$$

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*

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

and, for $t \geq 1$*,*

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

where

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

Thus, intergenerational transfers converge to steady state level

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

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 $R_w(0, \tau_w^0) = \bar{G}$; 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 $t \to \infty$).

Lemma 3. *Suppose* $e_0 = e^*(0, \tau_w^0)$ *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* $t \to \infty$.

<Figure 1>

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 onwards. Fig. 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 onwards. 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 onwards, 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

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

Then (13) and (16) imply that

$$e^*(\tau_b, \tau_w) = \frac{1}{4}\left[\frac{\beta(1-\tau_w)\bar{w}}{1-\tau_b}\right]^2. \tag{24}$$

Lemma 4. *Under specifications (15), (16) and (23):*
(i) $\partial R_w/\partial\tau_w > (=,<)0$ *if and only if* $\tau_w < (=,>)0.5$*;*
(ii) $b^*(0,\tau_w^0) > 0$ *if and only if* $e^*(0,\tau_w^0) < (1+\beta)\chi/3 \equiv \bar{e}(\beta,\chi)$*;*
(iii) for both $t = 1$ *and* $t \to \infty$, $\partial b_{t+1}/\partial\tau_b|_{\tau_b=0} < 0$.

Part (i) of Lemma 4 shows that a Laffer effect with respect to labor income taxation does not occur if tax rate τ_w is sufficiently small. Part (ii) of Lemma 4 implies that steady state bequests in absence of bequest taxation, $b^*(0,\tau_w^0)$ are positive if the bequest motive, measured by χ, is sufficiently strong. Finally, part (iii) implies that 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. *Suppose* $e_0 = e^*(0,\tau_w^0) < \bar{e}(\beta,\chi)$ *and* $b_1 = b^*(0,\tau_w^0)$. *Under specifications (15), (16) and (23), levying a small bequest tax improves welfare of each generation if* $\tau_w^0 > \bar{\tau}_w(\beta)$ *and* $e_0 \geq \underline{e}(\tau_w^0,\beta,\chi)$*, where*

$$\bar{\tau}_w(\beta) \equiv \frac{2-\beta}{2+\beta(4\beta+1)}, \tag{25}$$

with $\bar{\tau}_w(\beta) \in (0,1)$*, and*

$$\underline{e}(\tau_w^0,\beta,\chi) \equiv \frac{(1-\tau_w^0)(1-\beta)(1+\beta)\chi}{\tau_w^0\left(1+5\beta+8\beta^2\right)-1-\beta}, \tag{26}$$

with $\underline{e}(\tau_w^0, \beta, \chi) \in (0, \bar{e}(\beta, \chi))$.

According to Proposition 2, if the initial wage tax rate is sufficiently high ($\tau_w^0 > \bar{\tau}_w(\beta)$), i.e., the human capital investment decision is severely distorted by labor income taxation, a bequest tax may be efficiency-enhancing even if not used to lower the wage tax. (For instance, if $\beta = 0.9$, as used in the numerical analysis of the optimal tax structure in the next section, we have $\bar{\tau}_w(\beta) \approx 0.18$.) In this case, the incentive to raise educational investment may dominate the effect from a reduction in the amount of bequests on utility. Under the specifications of functional forms considered in Proposition 2, efficiency and welfare are indeed raised if, in addition to $\tau_w^0 > \bar{\tau}_w(\beta)$, incentives to invest in education (and thus $e_0 = e^*(0, \tau_w^0)$) are sufficiently high[15] (but low enough to induce positive bequests in the initial steady state; see Lemma 4 (ii)).

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, 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.

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

$$\max_{\tau_b, \tau_w} \left\{ u_2(\bar{w}h(e^*) + b^* - \bar{G} - s^* - e^*) + \beta u_3((1+\bar{r})s^* - b^*) + \beta v(\bar{w}h(e^*) + b^* - \bar{G}) \right\} \tag{27}$$

$$\text{s.t. } \tau_w \bar{w} h(e^*) + \tau_b b^* = \bar{G}. \tag{28}$$

[15] This is ensured if the wage rate $\bar{w}$ is sufficiently high, i.e., the economy is technologically advanced. To see this, recall $e^*(0, \tau_w) = [\beta(1-\tau_w)\bar{w}]^2/4$ and note that $\underline{e}$ as given in (26) is independent of $\bar{w}$.

[16] As 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.

Tab. 1 shows numerical results for the optimal tax rates, denoted τ_w^{opt}, τ_b^{opt}, for different government expenditures with an assumption that $\bar{w} = 1$ and $\beta = 0.9$, for varying levels of χ and $\bar{G}$.

χ	$\bar{G}$	τ_w^0	τ_w^{opt}	τ_b^{opt}	$\frac{b^*(0,\tau_w^0)}{\bar{w}h(e^*(0,\tau_w^0))}$	$\frac{b^*(\tau_b^{opt},\tau_w^{opt})}{\bar{w}h(e^*(\tau_b^{opt},\tau_w^{opt}))}$
0.4	0	0	0.048	-0.112	0.188	0.423
0.4	0.02	0.047	0.079	-0.065	0.269	0.416
0.4	0.04	0.099	0.109	-0.019	0.365	0.412
0.4	0.06	0.158	0.140	0.027	0.484	0.409
0.4	0.08	0.231	0.170	0.071	0.644	0.408
0.4	0.10	0.333	0.199	0.115	0.907	0.409
0.5	0	0	0.056	-0.084	0.423	0.660
0.5	0.02	0.047	0.081	-0.046	0.515	0.656
0.5	0.04	0.099	0.106	-0.008	0.625	0.653
0.5	0.06	0.158	0.130	0.029	0.762	0.652
0.5	0.08	0.231	0.154	0.065	0.949	0.652
0.5	0.10	0.333	0.178	0.100	1.259	0.654

Table 1. Optimal tax rates.

Our numerical results suggest certain general patterns. First of all, the optimal bequest tax rate is generally positive when government revenue requirement, $\bar{G}$, 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 $\bar{G}$ 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 from collecting any given revenue, it is optimal to distort human capital investment relatively more. For the same reason, when $\bar{G}$ 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 non-zero even in the case where $\bar{G} = 0$. 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.

Second, an increase in public expenditures $\bar{G}$ results in an increase in both tax rates τ_b^{opt} and τ_w^{opt} as well as in the ratio between the bequest tax rate and the wage tax rate, $\tau_b^{opt}/\tau_w^{opt}$ (that is, optimal bequest tax rate increases faster than the optimal wage tax rate). With a zero revenue requirement, this ratio is negative, then increasing and approaching unity as $\bar{G}$ increases.

In the last two columns of Tab. 1, we also report the size of bequests relative to the wage income that children receive over their working period, both in the initial situation (without bequest tax) and under the optimal tax mix. The relative size of bequests is increasing in the strength of parents' motive to transfer resources to their children, measured by parameter χ. (Recall that b^* is increasing in χ, whereas e^* is independent of χ.) In the absence of bequest taxation, increasing the wage tax rate 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 19 and 91 percent of the lifetime wage income with $\chi = 0.4$, and between 42 and 126 percent with $\chi = 0.5$. When the bequest tax rate is set optimally, the range is 41 to 42 percent with $\chi = 0.4$ and 65 to 66 percent with $\chi = 0.5$. This suggests that optimal taxation stabilizes the composition of intergenerational transfers when the general level of public expenditures changes.

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.[17] Nevertheless, one may ask if the potentially beneficial role of using bequest taxes suggested by our preceding analysis still holds when education subsidies are feasible. For this purpose, suppose each unit of investment in education, e, is subsidized by a constant rate τ_e. A numerical analysis of this extended model with optimally chosen education subsidies, focusing again on the steady state, suggests that education should indeed be subsidized, at a rate of similar magnitude as the optimal wage tax rate (results not shown).[18] Importantly, however, the main insight from Tab. 1, that bequests should be subsidized with a low government requirement $\bar{G}$ and taxed for a high level of $\bar{G}$, is unaffected. Thus, the qualitative results on the optimality of taxing bequests with a large public sector hold even when education subsidies are available.

5 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

[17]Trostel (1993) estimates that about a quarter of the costs of education are non-verifiable, 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 non-verifiable.

[18]Numerical results are provided in supplementary material which is available from the authors upon request.

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. Moreover, our analysis suggests that when the government revenue requirement increases the ratio between the bequest tax and the wage tax should increase.

Our results have certain surprising implications for the U.S. 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 percent 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 percent of population to transfer resources to their children more through education.

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

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

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

$$\frac{\partial U_0}{\partial \tau_b} = \beta u_3'(c_{3,2}) \left[(1+\bar{r}) \frac{\tau_w}{1-\tau_w} \frac{\partial e^*}{\partial \tau_b} + \frac{\tau_b}{1-\tau_b} \frac{\partial b_2}{\partial \tau_b} \right]. \tag{A.2}$$

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

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

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

$$\frac{\partial U_t}{\partial \tau_b} = u_2' \bar{w}h' \frac{\partial e^*}{\partial \tau_b} + u_2' \frac{\partial b_{t+1}}{\partial \tau_b} - u_2' \frac{\partial e^*}{\partial \tau_b} - \beta u_3' \frac{\partial b_{t+2}}{\partial \tau_b} + \beta v' \bar{w}h' \frac{\partial e^*}{\partial \tau_b} + \beta v' \frac{\partial b_{t+2}}{\partial \tau_b}. \tag{A.4}$$

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

$$\begin{aligned} \frac{\partial U_t}{\partial \tau_b} &= (1+\bar{r})\beta u_3' \bar{w}h' \frac{\partial e^*}{\partial \tau_b} + (1+\bar{r})\beta u_3' \frac{\partial b_{t+1}}{\partial \tau_b} - (1+\bar{r})\beta u_3' \frac{\partial e^*}{\partial \tau_b} \\ &\quad -\beta u_3' \frac{\partial b_{t+2}}{\partial \tau_b} + \beta \frac{u_3'}{1-\tau_b} \bar{w}h' \frac{\partial e^*}{\partial \tau_b} + \beta \frac{u_3'}{1-\tau_b} \frac{\partial b_{t+2}}{\partial \tau_b}. \end{aligned} \tag{A.5}$$

We obtain condition (14) by using (13), factoring out $\beta(1+\bar{r})u_3'$ and evaluating at $\tau_b = 0$. ■

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

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

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 $I_{t+2} - \chi = (1 - \tau_b)\left[(1+\bar{r})s_{t+1} - b_{t+2}\right]$. 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. ■

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). Fig. 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 $t \to \infty$ (i.e., for $b_{t+1} = b^*$), it also holds for all $t \geq 1$. This concludes the proof. ■

Proof of Lemma 4. Part (i) is confirmed by substituting (24) into $R_w = \tau_w \bar{w} \cdot (e^*)^{1/2}$. To prove part (ii), note that

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

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

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

which confirms part (ii). Regarding part (iii), 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

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

according to (24), we obtain

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

and

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

Both derivatives are negative. This concludes the proof. ■

Proof of Proposition 2.[19] First, note that $e_0 < \bar{e}(\beta, \chi)$ implies $b_1 > 0$, according to part (ii) of Lemma 4. According to Lemma 3 and the assumptions of Proposition 2, a Pareto-improvement is reached if

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

and

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

simultaneously hold.

We begin to check (A.12). It is tedious but straightforward to show that substituting

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

(A.9) and (A.10) into (A.12) and using (A.8) implies

$$\Omega^* = \frac{1+\beta}{4\beta^2}\left(\frac{e_0}{1-\tau_w^0}\left[\frac{(1+5\beta+8\beta^2)\tau_w^0}{1+\beta} - 1\right] - (1-\beta)\chi\right). \tag{A.14}$$

Thus, $\Omega^* \geq 0$ if and only if

$$\tau_w^0 > \frac{1+\beta}{1+5\beta+8\beta^2} \equiv q^*(\beta). \tag{A.15}$$

and $e_0 \geq \underline{e}(\tau_w^0, \beta, \chi)$ simultaneously hold, using the definition of $\underline{e}$ in (26). One can show that $\underline{e}(\tau_w^0, \beta, \chi) < \bar{e}(\beta, \chi)$ if and only if $\tau_w^0 > \bar{\tau}_w(\beta)$. Moreover, $\bar{\tau}_w(\beta) > q^*(\beta)$. Thus, $\tau_w^0 > \bar{\tau}_w(\beta)$ implies $\tau_w^0 > q^*(\beta)$. From (25), it is also easy to see that $\bar{\tau}_w(\beta) < 1$.

Now we turn to derive an expression for Ω_0. It is again tedious but straightforward to show that substituting (A.9) and (A.11) into (A.13) and using $b_1 = b^*(0, \tau_w^0)$ as given in (A.8) implies

$$\Omega_0 = \frac{1+\beta}{2\beta(1+2\beta)}\left(\frac{e_0}{1-\tau_w^0}\left[(1+8\beta)\tau_w^0 - 1\right] - (1-\beta)\chi\right), \tag{A.16}$$

Thus, (A.13) is fulfilled if and only if

$$\tau_w^0 > \frac{1}{1+8\beta} \equiv q_0(\beta) \tag{A.17}$$

and

$$e_0 \geq \frac{(1-\beta)\chi(1-\tau_w^0)}{(1+8\beta)\tau_w^0 - 1} \equiv \underline{\underline{e}}(\tau_w^0, \beta, \chi) \tag{A.18}$$

simultaneously hold. One can show that $\tau_w^0 > \bar{\tau}_w(\beta)$ implies $\tau_w^0 > q_0(\beta)$. Moreover, it is straightforward to check that $\underline{e}(\tau_w^0, \beta, \chi) > \underline{\underline{e}}(\tau_w^0, \beta, \chi)$, according to (26) and (A.18). Thus, if $\Omega^* \geq 0$, then $\Omega_0 > 0$. This concludes the proof. ■

References

Andreoni, James (1989). Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence, *Journal of Political Economy* 97, 1447-1458.

Barro, Robert J. (1974). Are Government Bonds Net Wealth?, *Journal of Political Economy* 82, 1095-1117.

Blinder, Alan S. (1976). Intergenerational Transfers and Life Cycle Consumption, *American Economic Review Papers and Proceedings* 66, 87-89.

Blumkin, Tomer and Efraim Sadka (2003). Estate Taxation with Intended and Accidental Bequests, *Journal of Public Economics* 88, 1-21.

Chamley, Christopher P. (1986). Optimal Taxation of Capital Income in General Equilibrium With Infinite Lives, *Econometrica* 54, 607-622.

Cremer, Helmuth and Pierre Pestieau (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, Helmuth and Pierre Pestieau (2003). Wealth Transfer Taxation: A Survey, CESifo Working Paper No. 1061.

Domeij, David (2005). Optimal Capital Taxation and Labor Market Search, *Review of Economic Dynamics* 8, 623-650.

Gale, William G. and Joel Slemrod (2001). Rethinking Estate and Gift Taxation: Overview, in W. Gale, J. Hines and J. Slemrod (eds.), Rethinking Estate and Gift Taxation, Brookingd Institution Press, 2001, 1-64.

Galor, Oded and Omer Moav (2004). From Physical to Human Capital Accumulation: Inequality in the Process of Development, *Review of Economic Studies* 71, 1001-1026.

Goldin, Claudia and Lawrence F. Katz (1998). The Origins of Technology-Skill Complementarity, *Quarterly Journal of Economics* 113, 693-732.

Gradstein, Mark and Moshe Justman (1997). Democratic Choice of an Education System: Implications for Growth and Income Distribution, *Journal of Economic Growth* 2, 169-183.

Ishikawa, Tsuneo (1975). Family Structures and Family Values in the Theory of Income Distribution, *Journal of Political Economy* 83, 987-1008.

Jacobs, Bas and A. Lans Bovenberg (2005). Human Capital and Optimal Positive Taxation of Capital Income. CEPR Discussion Paper No. 5047.

Jones, Larry E., Rodolfo E. Manuelli and Peter E. Rossi (1993). Optimal Taxation in Models of Endogenous Growth. *Journal of Political Economy* 101, 485-517.

Jones, Larry E., Rodolfo E. Manuelli and Peter E. Rossi (1997). On the Optimal Taxation of Capital Income. *Journal of Economic Theory* 73, 93-117.

Judd, Kenneth (1985). Redistributive Taxation in a Perfect Foresight Model, *Journal of Public Economics* 28, 59-84.

Judd, Kenneth (2002). Capital Income Taxation with Imperfect Competition, *American Economic Review Papers and Proceedings* 92, 417-421.

Kopczuk, Wojciech (2001). Optimal Estate Taxation in the Steady State, University of Columbia (mimeo).

Michel, Philippe and Pierre Pestieau (2004). Fiscal Policy in an Overlapping Generations Model with Bequest-as-Consumption, *Journal of Public Economic Theory* 6, 397-407.

Trostel, Philip A. (1993). The Effects of Taxation on Human Capital, *Journal of Political Economy* 101, 327-350.

Supplement (Not Intended for Publication)

Detailed Proof of Proposition 2

First, note that by using $b_1 = b^*(0, \tau_w^0)$, $e_0 = e^*(0, \tau_w^0) = \left[(1-\tau_w^0)\bar{w}\beta\right]^2/4$ and thus $(1-\tau_w^0)\bar{w}h(e_0) = 2e_0/\beta$, we can rewrite (A.10) and (A.11) as

$$\left.\frac{\partial b^*(\tau_b, \tau_w^0)}{\partial \tau_b}\right|_{\tau_b=0} = -\frac{2e_0/\beta + \left(\frac{2(2-\tau_w^0)}{1-\tau_w^0} - 1\right)e_0 + (1-\beta)b_1}{2\beta}, \tag{B.1}$$

and

$$\left.\frac{\partial B_0(b_1; \tau_b, \tau_w^0)}{\partial \tau_b}\right|_{\tau_b=0} = -\frac{2e_0/\beta + \left(\frac{2\left[1+\beta(2-\tau_w^0)\right]}{\beta(1-\tau_w^0)} - 1\right)e_0 + (1-\beta)b_1}{1+2\beta}, \tag{B.2}$$

according to (A.10) and (A.11), respectively. Moreover, using both $\partial e^*(\tau_b, \tau_w^0)/\partial \tau_b|_{\tau_b=0} = 2e_0$ from (A.9) and $b_1 = \left[(1+\beta)\chi - 3e_0\right]/(2\beta)$ from (A.8), and substituting (B.1) and (B.2) into (A.12) and (A.13), respectively, we obtain

$$\Omega^* = \frac{2\left(\frac{1}{\beta} + \tau_w^0\right)e_0}{1-\tau_w^0} - \frac{\frac{2e_0}{\beta} + \left(\frac{2(2-\tau_w^0)}{1-\tau_w^0} - 1\right)e_0 + \frac{(1-\beta)[(1+\beta)\chi - 3e_0]}{2\beta}}{2\beta} \tag{B.3}$$

and

$$\Omega_0 = \frac{2\left(\frac{1}{\beta} + \tau_w^0\right)e_0}{1-\tau_w^0} - \frac{\frac{2e_0}{\beta} + \left(\frac{2\left[1+\beta(2-\tau_w^0)\right]}{\beta(1-\tau_w^0)} - 1\right)e_0 + \frac{(1-\beta)[(1+\beta)\chi - 3e_0]}{2\beta}}{1+2\beta}. \tag{B.4}$$

(B.3) can be rewritten as

$$
\begin{aligned}
\Omega^* &= \frac{e_0}{2\beta\left(1-\tau_w^0\right)}\left[4(1+\beta\tau_w^0)-\frac{2\left(1-\tau_w^0\right)}{\beta}-\left(3-\tau_w^0\right)+\frac{3(1-\beta)\left(1-\tau_w^0\right)}{2\beta}\right]- \\
&\quad \frac{(1+\beta)(1-\beta)\chi}{4\beta^2} \\
&= \frac{1}{4\beta^2}\frac{e_0}{\left(1-\tau_w^0\right)}\left[8\beta(1+\beta\tau_w^0)-4\left(1-\tau_w^0\right)-2\beta\left(3-\tau_w^0\right)+3(1-\beta)\left(1-\tau_w^0\right)\right]- \\
&\quad \frac{(1+\beta)(1-\beta)\chi}{4\beta^2} \\
&= \frac{1}{4\beta^2}\left(\frac{e_0}{\left(1-\tau_w^0\right)}\left[\left(8\beta^2+5\beta+1\right)\tau_w^0-(1+\beta)\right]-(1+\beta)(1-\beta)\chi\right) \\
&= \frac{1+\beta}{4\beta^2}\left(\frac{e_0}{1-\tau_w^0}\left[\frac{\left(8\beta^2+5\beta+1\right)\tau_w^0}{1+\beta}-1\right]-(1-\beta)\chi\right), \qquad \text{(B.5)}
\end{aligned}
$$

which confirms (A.14). We turn next to examine when

$$\underline{e}(\tau_w^0,\beta,\chi)=\frac{(1-\tau_w^0)(1-\beta)(1+\beta)\chi}{\left(8\beta^2+5\beta+1\right)\tau_w^0-1-\beta}<\frac{(1+\beta)\chi}{3}=\bar{e}(\beta,\chi). \tag{B.6}$$

Rewriting inequality (B.6) leads to

$$3(1-\tau_w^0)(1-\beta)<\left(8\beta^2+5\beta+1\right)\tau_w^0-1-\beta, \tag{B.7}$$

which holds if and only if

$$\tau_w^0>\frac{2-\beta}{2+\beta(4\beta+1)}=\bar{\tau}_w(\beta). \tag{B.8}$$

This confirms that $\underline{e}<\bar{e}$ under assumption $\tau_w^0>\bar{\tau}_w(\beta)$ made in Proposition 2. (Moreover, (B.8) immediately implies $\tau_w^0\in(0,1)$.) We can then show that

$$\bar{\tau}_w(\beta)=\frac{2-\beta}{2+\beta(4\beta+1)}>\frac{1+\beta}{8\beta^2+5\beta+1}=q^*(\beta). \tag{B.9}$$

To see this, rewrite (B.9) such that

$$(2-\beta)\left(8\beta^2+5\beta+1\right)>(1+\beta)\left(2+4\beta^2+\beta\right). \tag{B.10}$$

(B.10) reduces to $\beta(3\beta-2)+3>0$, which holds as $\beta \in (0,1)$. Thus, $\tau_w^0 > \bar{\tau}_w(\beta)$ implies $\tau_w^0 > q^*(\beta)$. In sum, this confirms $\Omega^* \geq 0$ if $\tau_w^0 > \bar{\tau}_w(\beta)$ and $e_0 \geq \underline{e}(\tau_w^0, \beta, \chi)$, according to (B.5) and (B.6).

Now we turn to rewrite Ω_0. According to (B.4),

$$\begin{aligned}
\Omega_0 &= \left[(1+2\beta)\left(\frac{2}{\beta}+2\tau_w^0\right) - \frac{2\left(1-\tau_w^0\right)}{\beta} - \frac{2\left(1+\beta(2-\tau_w^0)\right)}{\beta} + 1 - \tau_w^0 + \frac{3(1-\beta)}{2\beta}\right] \times \\
&\quad \frac{e_0}{(1+2\beta)\left(1-\tau_w^0\right)} - \frac{(1+\beta)(1-\beta)\chi}{2\beta(1+2\beta)} \\
&= \left[(1+2\beta)\left(4+4\beta\tau_w^0\right) - 4\left(1-\tau_w^0\right) - 4\left(1+\beta) - 4\beta(1-\tau_w^0)\right) + 2\beta\left(1-\tau_w^0\right) + \right. \\
&\quad \left. 3(1-\beta)\left(1-\tau_w^0\right)\right] \times \frac{e_0}{2\beta(1+2\beta)\left(1-\tau_w^0\right)} - \frac{(1+\beta)(1-\beta)\chi}{2\beta(1+2\beta)} \\
&= \frac{1}{2\beta(1+2\beta)}\left(\frac{e_0}{(1-\tau_w^0)}\left[4\beta+4\beta\tau_w^0+8\beta^2\tau_w^0 - \left(1-\tau_w^0\right)(1+5\beta)\right] - (1+\beta)(1-\beta)\chi\right) \\
&= \frac{1+\beta}{2\beta(1+2\beta)}\left(\frac{e_0}{(1-\tau_w^0)}\left[\frac{\left(8\beta^2+5\beta+1\right)\tau_w^0}{1+\beta} - 1\right] - (1-\beta)\chi\right) \\
&= \frac{1+\beta}{2\beta(1+2\beta)}\left(\frac{e_0}{(1-\tau_w^0)}\left[(8\beta+1)\tau_w^0 - 1\right] - (1-\beta)\chi\right), \qquad \text{(B.11)}
\end{aligned}$$

which confirms (A.16). We turn next to show that

$$\underline{\underline{e}}(\tau_w^0, \beta, \chi) = \frac{(1-\tau_w^0)(1-\beta)\chi}{(8\beta+1)\tau_w^0 - 1} < \frac{(1-\tau_w^0)(1-\beta)(1+\beta)\chi}{\left(8\beta^2+5\beta+1\right)\tau_w^0 - 1 - \beta} = \underline{e}(\tau_w^0, \beta, \chi). \quad \text{(B.12)}$$

Rewriting inequality (B.12) leads to

$$\left(8\beta^2+5\beta+1\right)\tau_w^0 - 1 - \beta < (1+\beta)\left[(8\beta+1)\tau_w^0 - 1\right], \quad \text{(B.13)}$$

which is easily seen to hold if $\beta > 0$, as assumed. Moreover,

$$\bar{\tau}_w(\beta) = \frac{2-\beta}{2+\beta(4\beta+1)} > \frac{1}{8\beta+1} = q_0(\beta) \quad \text{(B.14)}$$

is equivalent to condition $7 > 6\beta$, which is fulfilled. Thus, under the assumptions of Proposition 2, which imply $\Omega^* \geq 0$, we find that $\Omega_0 > 0$ holds. This concludes the proof. ■

Optimal Tax Mix with Education Subsidies

In this appendix we allow for a proportional subsidy on education investment e, at rate τ_e. First, note that under a balanced budget, $T_{t+1} = 0$. Thus, the government budget constraint in period $t+1$ modifies to

$$\tau_w \bar{w} h(e_t) + \tau_b b_{t+1} = \bar{G} + \tau_e e_{t+1}. \tag{C.1}$$

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

$$c_{2,t+1} + s_{t+1} + (1-\tau_e)e_{t+1} = I_{t+1}, \tag{C.2}$$

$$c_{3,t+2} + b_{t+2} = (1+\bar{r})s_{t+1}. \tag{C.3}$$

First-order conditions associated with the individual optimization problem now read

$$\frac{u_2'(c_{2,t+1})}{\beta u_3'(c_{3,t+2})} = 1 + \bar{r}, \tag{C.4}$$

$$\frac{u_2'(c_{2,t+1})}{\beta v'(I_{t+2})} = \frac{1-\tau_w}{1-\tau_e} \bar{w} h'(e_{t+1}), \tag{C.5}$$

$$\frac{u_3'(c_{3,t+2})}{v'(I_{t+2})} = 1 - \tau_b. \tag{C.6}$$

Observing (C.1), income $I_{t+1} = (1-\tau_w)\bar{w}h(e_t) + (1-\tau_b)b_{t+1}$ can be rewritten as

$$I_{t+1} = \bar{w}h(e_t) + b_{t+1} - \tau_e e_{t+1} - \bar{G}. \tag{C.7}$$

It is straightforward to show that, under specifications (15), (16) and (23), (C.2)-(C.7) imply steady state levels

$$e^* = \frac{1}{4}\left[\frac{\beta(1-\tau_w)\bar{w}}{(1-\tau_b)(1-\tau_e)}\right]^2, \tag{C.8}$$

$$b^* = \frac{(1+\beta)\chi - (\beta+\tau_b)\left(\bar{w}(e^*)^{1/2} - \tau_e e^* - \bar{G}\right) - (1-\tau_b)(1-\tau_e)e^*}{2\beta + \tau_b(1-\beta)}, \tag{C.9}$$

$$s^* = \frac{\beta}{1+\beta}\left(\bar{w}(e^*)^{1/2} + 2b^* - e^* - \bar{G}\right). \tag{C.10}$$

Now, observing (C.7), under specifications (15), (16) and (23), the social planner's optimization problem modifies to

$$\max_{\tau_b,\tau_w,\tau_e} \left\{\ln(\bar{w}(e^*)^{1/2} + b^* - \bar{G} - s^* - e^*) + \beta \ln\left(\frac{s^*}{\beta} - b^*\right) + \beta \ln(\bar{w}(e^*)^{1/2} + b^* - \tau_e e^* - \bar{G} - \chi)\right\} \tag{C.11}$$

$$\text{s.t. } \tau_w \bar{w}(e^*)^{1/2} + \tau_b b^* = \bar{G} + \tau_e e^*, \tag{C.12}$$

given the expressions in (C.8)-(C.10).

The following table shows numerical results for the optimal wage income tax rate, the tax rate on bequests and the subsidy on education investment, denoted τ_w^{opt}, τ_b^{opt} and τ_e^{opt}, respectively, for different government expenditures with an assumption that $\bar{w} = 0.7$, $\beta = 0.9$ and $\chi = 0.4$.

$\bar{G}$	τ_w^{opt}	τ_b^{opt}	τ_e^{opt}	$\frac{b^*}{\bar{w}h(e^*)}$
0	0.72	-0.23	0.79	1.06
0.02	0.72	-0.17	0.79	1
0.04	0.72	-0.12	0.76	1.1
0.06	0.72	-0.06	0.77	1
0.08	0.78	-0.03	0.81	1.05
0.10	0.75	0.03	0.77	1.05
0.12	0.69	0.09	0.67	1.15
0.14	0.75	0.12	0.75	1.13
0.16	0.78	0.15	0.75	1.25

Notes:

1. For computational reasons, feasible values for τ_w and τ_b are restricted to the set $\Xi = \{-0.6, -0.57, -0.54, ..., 0.72, 0.75, 0.78\}$, i.e., in steps of 0.03 between values -0.6 and 0.78. For a given parameter constellation, the maximization procedure is as follows. First, it computes for each combination $\tau_w, \tau_b \in \Xi$ the value of τ_e which fulfills the government budget constraint, and then gives for all so-constructed triples

$\{\tau_w, \tau_b, \tau_e\}$ the one which yields the highest utility.

2. The last column reports the implied value of bequests to wage income under the optimal chosen triple $\{\tau_w, \tau_b, \tau_e\}$. ($\tau_e > 0$ means that education is subsidized.)

Figure

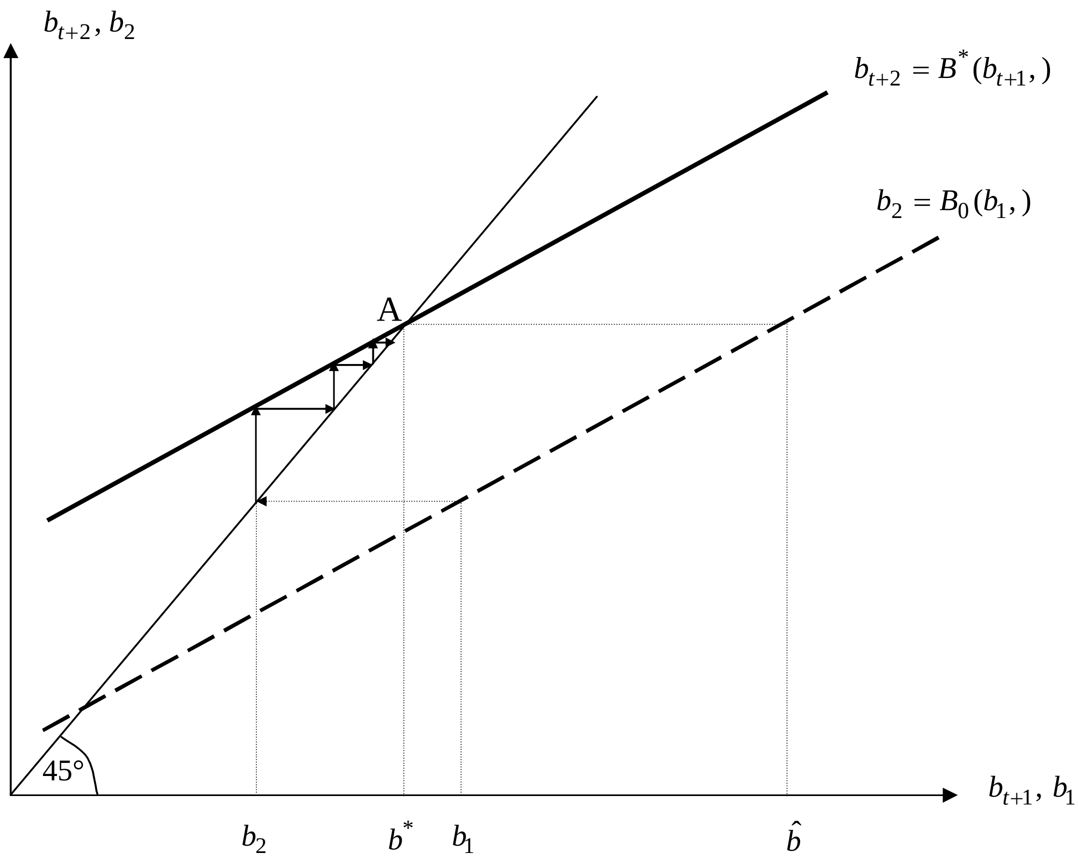

Figure 1: The evolution of bequests, illustrated for the case $b^* < b_1 < \hat{b}$.