*The B.E. Journal of Economic Analysis & Policy*

Topics

*Volume* 12, *Issue* 1 2012 *Article* 8

# Optimal Taxation of Wealth Transfers When Bequests are Motivated by Joy of Giving

Johann K. Brunner* Susanne Pech†

*Johannes Kepler University of Linz, johann.brunner@jku.at
†Johannes Kepler University of Linz, susanne.pech@jku.at

**Recommended Citation**
Johann K. Brunner and Susanne Pech (2012) "Optimal Taxation of Wealth Transfers When Bequests are Motivated by Joy of Giving," *The B.E. Journal of Economic Analysis & Policy*: Vol. 12: Iss. 1 (Topics), Article 8.

DOI: 10.1515/1935-1682.3070

# Optimal Taxation of Wealth Transfers When Bequests are Motivated by Joy of Giving*

Johann K. Brunner and Susanne Pech

## Abstract

Inherited wealth creates a second distinguishing characteristic of individuals, in addition to earning abilities. We incorporate this fact into a model of optimal labor-income taxation, with bequests motivated by joy of giving. We find that taxes on bequests or on inheritances allow further redistribution if, in the parent generation, initial wealth and earning abilities are positively related. However, these taxes distort the bequest decision and thus, the overall effect on social welfare is ambiguous. On the other hand, a tax on all expenditures of a generation (a uniform tax on consumption plus bequests) has the same redistributive effect as an inheritance tax but does not distort the bequest decision.

**KEYWORDS:** optimal taxation, inheritance tax, expenditure tax, intergenerational wealth transfer

*We are grateful to two anonymous referees for many helpful comments.

# 1 Introduction

The tax on estates or inheritances has always been a highly controversial issue, particularly in recent years, when it was abolished in some countries such as Sweden and Austria.[1] In the political discussion, opponents consider it morally inappropriate to use the moment of death as a cause for imposing a tax. They stress its negative economic consequences, in particular, on capital accumulation and on family business. Supporters find these consequences exaggerated and claim that a tax on bequests is desirable for redistributive reasons, contributing to "equality of opportunity".

From the perspective of economic theory, the essential issue is to formulate an appropriate model that allows a discussion of how a shift from labor-income taxation to a tax on intergenerational wealth transfers affects social welfare. In this paper we suggest an extended optimal-taxation model in the tradition of Mirrlees (1971). The starting point of our analysis is the following: inherited wealth creates a second distinguishing characteristic of individuals, in addition to earning abilities. The role of taxes on estates or inheritances cannot be well understood without an explicit consideration of this fact.

Typically, former contributions discussing bequest taxation in an optimal-taxation framework have focused on the specifics of leaving bequests, as compared to other ways of spending the budget, that is, to the consumption of goods.[2] Such an analysis, referring to a standard result in optimal-taxation theory (Atkinson and Stiglitz, 1976), leads to the question of whether preferences are separable between leisure and consumption plus bequests, in which case an income tax alone suffices and spending need not be taxed at all, or whether leaving bequests represents a complement or a substitute to enjoying leisure. We argue in the present paper that this is an inappropriate question, because the Atkinson-Stiglitz result is derived for a model where individuals only differ in earning abilities. What matters is not that bequests represent a particular use of the budget, but the fact that they generate differences in inherited wealth within a generation.

We consider an optimal-taxation model which accounts for these differences, in addition to earning abilities, and analyze the welfare effect of proportional taxes on bequests or inheritances, as well as of a proportional tax on all expenditures (i.e., on consumption plus bequests), which can be im-

[1] However, this tax still exists in most OECD countries.

[2] See Gale and Slemrod (2001, p.33) and Kaplow (2001), as well as Blumkin and Sadka (2003) in the context of a dynastic model.

posed in addition to an optimal nonlinear income tax. While we have studied a similar model (Brunner and Pech, 2012) with bequests motivated by pure altruism, we now assume a joy-of-giving motive for leaving bequests. As is well-known, the altruistic motive leads to the study of dynasties (Barro, 1974) where preferences of parents comprise preferences of all descendant generations. A joy-of-giving motive constitutes a weaker link between generations; individuals just enjoy bequeathing some amount, without an explicit reference to their heirs' welfare position. Empirical studies find quite mixed results (for an overview, see Kopczuk and Lupton, 2007, and Kopczuk, 2010) and thus do not allow for a definitive conclusion with regard to which of the two bequest motives dominates the actual decisions. Instead, it appears likely that a combination of these bequest motives (together with strategic and accidental bequests)[3] is required to explain the observed pattern of bequests. Consequently, the investigation of the effect of taxing bequests, when the latter are motivated by joy of giving, is important.

We find that a tax on bequests of some generation generates a positive welfare effect by allowing increased redistribution (compared to an income tax alone) within this generation, if initial (inherited) wealth is positively related to earning abilities. Indeed, this effect may be the reason why such a tax is frequently considered as increasing the equality of opportunity. It does not occur in a model without differences in initial wealth and has to be contrasted with a negative effect on future generations, due to a distortion of the bequest decision. The positive effect is more likely to dominate, the more double counting of bequests is laundered out of the social welfare function.[4] Moreover, a tax on inheritances received by some generation is found to have essentially the same consequences as a tax on bequests of the previous generation.

The effect of a bequest tax turns out to be similar, irrespective of whether a joy-of-giving or an altruistic bequest motive prevails. On the other hand, however, the bequest motive makes a major difference for the expenditure tax. It is generally true that taxing the expenditures of an individual of some generation $t$ is equivalent to taxing her initial wealth, if combined with an optimal nonlinear labor-income tax. With an altruistic motive, this equivalence extends to the distorting effect on the previous generation $t-1$,

[3] For a typology of bequest motives, see Cremer and Pestieau (2006) or Kopczuk (2010).

[4] Several authors (Hammond, 1987, Milgrom, 1993, and Harsanyi, 1995) argue that external preferences (oriented towards other individuals) should not be included in the social welfare function in order to avoid inappropriate welfare weights (double counting). We follow a flexible approach (see also Michel and Pestieau, 2004) by introducing a parameter whose value (between zero and one) indicates the extent to which double counting of bequests is avoided by the social planner.

whose bequests generate initial wealth of generation $t$. But with a joy-of-giving motive, this equivalence no longer holds; the expenditure tax does not distort the bequest decision of generation $t-1$, because the bequeathing individuals do not take into account a tax imposed on expenditures of generation $t$. As a consequence, the expenditure tax has an unambiguously welfare-increasing effect, if initial wealth is positively related with earning abilities, .

Some earlier studies pay attention to the fact that inheritances create a second distinguishing characteristic in addition to earning abilities. However, to our knowledge this literature does not provide a unified framework for an analysis of the role of bequest taxation within an optimal tax system. Cremer et al. (2001) resume the discussion of indirect taxes (in addition to an optimal nonlinear income tax), given that individuals differ in endowments as well as abilities. They assume, however, that inheritances are unobservable and concentrate on the structure of indirect tax rates. Similarly, Cremer et al. (2003) and Boadway et al. (2000) study the desirability of a tax on capital income as a surrogate for the taxation of inheritances, which are considered unobservable.

In contrast to these contributions, we study a comprehensive tax system where a nonlinear tax on labor income can be combined with taxes on inherited wealth and on expenditures. We take all these variables as being observable (only abilities are unobservable). This is indeed the basis upon which real-world tax systems, including the tax on bequests, operate. In particular, notwithstanding problems of observability, if we want to know whether the inheritance tax should be retained or abolished from a welfare-theoretic point of view, the analysis must be based on the assumption of observable initial wealth.

In Section 2 we introduce the basic model and discuss the consequences of bequest taxation, given that initial wealth is exogenous. In Section 3 we incorporate the distortion of the bequest decision of the previous generation. Section 4 provides some discussion of the results.

## 2 The model

We start with a simple model of two generations that allows us to analyze the consequences of bequest taxation on social welfare, including redistribution. In some period $t$ there exist two individuals $i = L, H$ with abilities $\omega_{Lt} < \omega_{Ht}$. Individuals live for one period only. By supplying labor $l_{it}$ they earn gross income $z_{it} = \omega_{it} l_{it}$ and net income $x_{it}$. They are endowed with initial wealth $e_{it}$, which is used, together with net income, for own consumption $c_{it}$ and for leaving bequests $b_{it}$. For the moment, we take initial wealth as exogenous, resulting from bequests left by the previous generation, whose behavior will be studied in the next section. Preferences are identical for the individuals and can be described by the separable utility function $u(c_{it}, l_{it}) + h(b_{it})$, where both $u$ and $h$ are assumed to be strictly concave and twice differentiable with $\partial u/\partial c_{it} > 0$, $\partial u/\partial l_{it} < 0$, $h' > 0$.

In period $t + 1$ a generation of heirs exists that again consists of two individuals with differing working abilities $\omega_{Lt+1} < \omega_{Ht+1}$. They earn gross income $z_{jt+1} = \omega_{jt+1} l_{jt+1}$, and net income $x_{jt+1}$, which they spend, together with inheritances $e_{jt+1}$, for own consumption $c_{jt+1}$, $j = L, H$. Being the last generation, they do not leave bequests. Their preferences are described by the utility function $u(c_{jt+1}, l_{jt+1})$. Their parents (generation $t$) determine how bequests $b_{it}$ are assigned to inheritances $e_{jt+1}$, that is, the parents may split their estates and leave some share to each individual of the next generation.[5] As our results hold for any such assignment, we need not specify this decision in detail. We just describe it as a linear rule, $e_{jt+1} = \alpha^j_{Lt} b_{Lt} + \alpha^j_{Ht} b_{Ht}$, with $\alpha^j_{it} \geq 0$, $\sum_{j=L,H} \alpha^j_{it} = 1, i = L, H$, and write generally $e_{jt+1}(b_{Lt}, b_{Ht})$ to indicate the dependency.

The government imposes nonlinear taxes on labor income in the periods $t$ and $t+1$. We will analyze the role of three additional tax instruments, namely a proportional tax $\tau_{bt}$ on bequests $b_{it}$, a proportional tax $\tau_{et}$ on initial wealth $e_{it}$, and a proportional tax $\tau_t$ on total expenditures $c_{it} + b_{it}$.

Indirect utility of an individual $i$ of generation $t$ is defined for given values of gross income $z_{it}$, net income $x_{it}$, initial wealth $e_{it}$ and given tax rates $\tau_{bt}$, $\tau_{et}$ and $\tau_t$:

$$\begin{aligned} v^i_t(x_{it}, z_{it}, e_{it}, \tau_{bt}, \tau_{et}, \tau_t) \equiv \max\{&u(c_{it}, z_{it}/\omega_{it}) + \\ &+ h(b_{it}) | (1 + \tau_t)(c_{it} + (1 + \tau_{bt})b_{it}) \leq x_{it} + (1 - \tau_{et})e_{it}\}, \end{aligned} \tag{1}$$

where we assume that either $\tau_{bt}$ or $\tau_t$ may exist, but not both. The RHS of (1) also gives us the demand functions $c_{it}$ and $b_{it}$ with the same arguments as $v^i_t$. Moreover, we use $v^j_{t+1}(x_{jt+1}, z_{jt+1}, e_{jt+1}) \equiv u(x_{jt+1} + e_{jt+1}, z_{jt+1}/\omega_{jt+1})$ to indicate utility of individual $j$ in period $t+1$ for any given $(x_{jt+1}, z_{jt+1})$-bundle and any given inherited wealth $e_{jt+1}$.

In this framework, finding the optimal nonlinear income tax in both periods is equivalent to determining the optimal gross- and net-income bundles $(x_{Lt}, z_{Lt})$, $(x_{Ht}, z_{Ht})$ and $(x_{Lt+1}, z_{Lt+1})$, $(x_{Ht+1}, z_{Ht+1})$ for both generations. We

[5] This is in contrast to the model with altruistic preferences, where transmission necessarily takes place from a parent to the child within a dynasty.

assume that the government's objective is

$$\sum_{i=L,H} f_{it}(u(c_{it}(\cdot), z_{it}/\omega_{it}) + \delta h(b_{it}(\cdot))) + \beta \sum_{j=L,H} f_{jt+1} v^j_{t+1}(\cdot)$$

with weights $f_{it}$, $f_{jt+1}$ and a social discount factor $\beta \leq 1$ of the welfare of the future generation. As already mentioned in Section 1, we introduce the parameter $\delta$, $0 \leq \delta \leq 1$, as indicating the extent to which the government avoids double counting of bequests. With $\delta = 0$ there is full laundering out and with $\delta = 1$ there is no laundering out, that is, there is full double counting. Note that $u(c_{it}(\cdot), z_{it}/\omega_{it}) + \delta h(b_{it}(\cdot)) = v^i_t(\cdot) + (\delta - 1)h(b_{it}(\cdot))$.

Public spending is $g_t$ and $g_{t+1}$ in the two periods, and we assume that the government can transfer any amount $s_t$ of resources over the two periods. Thus, the maximization problem reads as

$$\max_{\substack{x_{it}, z_{it}, x_{jt+1}, \\ z_{jt+1}, i,j=L,H, s_t}} \sum_{i=L,H} f_{it}(v^i_t(\cdot) + (\delta - 1)h(b_{it}(\cdot))) + \beta \sum_{j=L,H} f_{jt+1} v^j_{t+1}(\cdot), \tag{2}$$

$$\text{s.t.} \quad v^H_t(x_{Ht}, z_{Ht}, e_{Ht}, \tau_{bt}, \tau_{et}, \tau_t) \geq v^H_t(x_{Lt}, z_{Lt}, e_{Ht}, \tau_{bt}, \tau_{et}, \tau_t), \tag{3}$$

$$v^H_{t+1}(x_{Ht+1}, z_{Ht+1}, e_{Ht+1}) \geq v^H_{t+1}(x_{Lt+1}, z_{Lt+1}, e_{Ht+1}) \tag{4}$$

$$\sum_{i=L,H} x_{it} \leq \sum_{i=L,H} z_{it} + \tau_{et} \sum_{i=L,H} e_{it} + \tau_{bt} \sum_{i=L,H} b_{it} + \tau_t \sum_{i=L,H} (c_{it} + b_{it}) - g_t - s_t, \tag{5}$$

$$\sum_{j=L,H} x_{jt+1} \leq \sum_{j=L,H} z_{jt} - g_{t+1} + s_t. \tag{6}$$

(5) and (6) are the public resource constraints, and (3) and (4) represent the self-selection constraints in the two periods. We assume that in the optimum only the self-selection constraints for the respective high-able individual is binding, so that the corresponding constraints for the low-able individual can be neglected. Moreover, we assume that the single-crossing condition is fulfilled: $-(\partial v^i_t(\cdot)/\partial z_{it})/(\partial v^i_t(\cdot)/\partial x_{it})$ is larger for type $L$ than for type $H$, at any admissible $(x, z)$-bundle; the same is assumed for period $t + 1$.

We can split the above problem into two subproblems by defining $W_{t+1}(b_{Lt}, b_{Ht}, s_t)$ as the maximal value of $\sum_{j=L,H} f_{jt+1} v^j_{t+1}(x_{jt+1}, z_{jt+1}, e_{jt+1})$ subject to (4) and (6). Note that each $e_{jt+1}$ depends on $b_{Lt}, b_{Ht}$, as discussed above. The two-period problem (2) - (6) is then equivalent to

$$\max_{x_{it}, z_{it}, i=L,H, s_t} \sum_{i=L,H} f_{it}(v^i_t(\cdot) + (\delta - 1)h(b_{it}(\cdot))) + \beta W_{t+1}(b_{Lt}, b_{Ht}, s_t) \tag{7}$$

$$\text{s.t.} \quad (3) \text{ and } (5).$$

As is standard in studies of indirect taxes in a Mirrlees-type model, we assume that the tax authority cannot identify individuals through information obtained when collecting the taxes on bequests, initial wealth or expenditures. Otherwise, given a fixed relation between the individuals' abilities and the levels of these variables, it would be possible for the government to impose an optimal differentiated lump-sum tax as a first-best solution.[6]

Differentiation of the optimal value function $S_t(\tau_{bt}, \tau_{et}, \tau_t)$ of (7), (3), (5) gives us the effect of the three indirect taxes ($\mu_t$ and $\lambda_t$ denote the Lagrangian multipliers associated with (3) and (5), respectively):

**Proposition 1** *Consider a model with two generations and exogenous inherited wealth of generation $t$.*

*a) Let $\tau_t = 0$. The welfare effect of a tax $\tau_{bt}$ on bequests left by generation $t$ reads as*

$$\begin{aligned}\frac{\partial S_t}{\partial \tau_{bt}} &= \mu_t \frac{\partial v_t^H[L]}{\partial x_{Lt}}(b_{Ht}[L] - b_{Lt}) + \beta \sum_{i=L,H} \frac{\partial W_{t+1}}{\partial b_{it}} \frac{\partial b_{it}^{com}}{\partial \tau_{bt}} + \\ &\quad +(\delta - 1) \sum_{i=L,H} f_{it} h'_{it} \frac{\partial b_{it}^{com}}{\partial \tau_{bt}} + \lambda_t \tau_{bt} \sum_{i=L,H} \frac{\partial b_{it}^{com}}{\partial \tau_{bt}}.\end{aligned}$$

*b) Let $\tau_{bt} = 0$. The welfare effect of a tax $\tau_{et}$ on initial wealth of generation $t$ reads as*

$$\frac{\partial S_t}{\partial \tau_{et}} = \mu_t \frac{\partial v_t^H[L]}{\partial x_{Lt}}(e_{Ht} - e_{Lt}).$$

*c) Let $\tau_{bt} = 0$. The welfare effect of a tax $\tau_t$ on total expenditure of generation $t$ reads as*

$$\frac{\partial S_t}{\partial \tau_t} = \mu_t \frac{\partial v_t^H[L]}{\partial x_{Lt}}(e_{Ht} - e_{Lt}) \frac{1 - \tau_{et}}{1 + \tau_t}.$$

**Proof.** See Appendix. ■

In the above formula, $h'_{it}$ abbreviates $dh/db_{it}$. The upper index "*com*" denotes the compensated demand function, and the symbol $[L]$ indicates mimicking, that is, a situation when the high-able individual opts for the $(x_{Lt}, z_{Lt})$-bundle designed for the low-able individual. It turns out that the welfare effects

[6] This assumption is in line with actual behavior of tax authorities. Moreover, one can show that the following results remain essentially unchanged, if a stochastic instead of a fixed relation between abilities and the levels of bequests, initial wealth and expenditures is assumed, so that identification is indeed not possible from these levels (Brunner and Pech, 2008).

of $\tau_t$ and $\tau_{et}$ are unambiguously positive, if within generation $t$ the high-able individual is endowed with more initial wealth than the low-able individual, because the self-selection constraint is binding, then the multiplier is positive, as is the marginal utility of net income. The interesting point is that both taxes have essentially the same welfare consequences, although the tax $\tau_{et}$ on (exogenous) initial wealth is lump-sum while the expenditure tax $\tau_t$ is, in principle, distorting, because it falls on total consumption which depends upon endogenous labor income. However, it is possible to adjust the optimal tax on labor income in such a way that the distorting effect of $\tau_t$ is eliminated. Remarkably, an adjustment of the labor-income tax also allows to compensate for all possibly negative effects on bequests of generation $t$, thus on welfare of generation $t+1$; no such effect occurs in Propositions 1b and 1c.

The positive effect on welfare comes from a relaxation of the self-selection constraint induced by an increase of $\tau_{et}$ (or $\tau_t$). The intuition can be explained as follows: assume, as a first-step, that after an increase of $\tau_{et}$ by $\Delta\tau_{et}$, each individual $i$ of generation $t$ is exactly compensated through an increase of net labor income $x_{it}$ by $\Delta\tau_{et}e_{it}$. If $e_{Ht} > e_{Lt}$, the high-able individual experiences a larger increase of net labor income than the less-able individual, which makes mimicking less attractive and gives slack to the self-selection constraint. Then in a second step, additional redistribution from the high- to the low-able individual becomes possible, which increases social welfare.

To see why a tax $\tau_{et}$ on initial wealth and a tax $\tau_t$ on all expenditures have identical effects (up to a normalizing factor), note that in a model without initial wealth a proportional expenditure tax (uniform for all goods) and a proportional labor-income tax are equivalent (Atkinson and Stiglitz, 1976, p. 64). A straightforward extension is that, if a proportional expenditure tax is introduced, any given *nonlinear* tax on labor income can be modified in such a way that the expenditure tax is outweighed; the overall effect is zero. This is possible if there is no initial wealth; otherwise the modification of the income tax implies that the expenditure tax falls on initial wealth only.[7]

The welfare effect is more complex for the tax $\tau_{bt}$ on bequests of generation $t$. If $e_{Ht} > e_{Lt}$ and if preferences are weakly separable between labor and consumption, then we find a positive welfare effect due to increased redis-

[7] More formally: Let a budget equation $p_1X_1 + p_2X_2 \leq e + \omega l - \sigma(\omega l)$ be given, for some tax function $\sigma$ and with $X_1, X_2$ denoting the consumption of two goods and $p_1, p_2$ their prices. For some $t$ we define the modified tax function $\sigma'$ by the condition $\omega l - \sigma'(\omega l) = (\omega l - \sigma(\omega l))(1+t)$. Then introducing the expenditure tax with rate $t$ and modifiying $\sigma$ to $\sigma'$ gives us the budget constraint $p_1(1+t)X_1 + p_2(1+t)X_2 \leq e + \omega l - \sigma'(\omega l)$, which clearly is equivalent to the original one, if $e$ is replaced by $e/(1+t)$. That is, we arrive at a tax $t' = t/(1+t)$ on initial wealth, where $1 - t' = 1/(1+t)$.

tribution. Note, however, the mechanism behind this effect: taxing bequests is, in fact, a means of taxing initial wealth to some extent. Bequests $b_{Ht}[L]$ of the $H$-type, when mimicking, are larger than bequests $b_{Lt}$ of the $L$-type, if $e_{Ht} > e_{Lt}$, because given an additive utility function, bequests are a normal good while weak separability implies that the lower labor time of the mimicker does not change the spending decision (for given net income). Thus, the intuition is similar to the one described above: compensating the individuals of generation $t$ by an increase in net income of $\Delta\tau_{bt}b_{Ht}$ and $\Delta\tau_{bt}b_{Lt}$, respectively, gives slack to the self-selection constraint.

In addition to this increased redistribution, several other effects arise which are related to the distortion of the bequest decision caused by the tax $\tau_{bt}$. This distortion leads to a negative own compensated price effect on bequests, and thus affects welfare of generation $t+1$ negatively. As a counterpart, this distortion creates a positive welfare effect on generation $t$, if the social planner wants to avoid double counting of bequests, i.e., if $\delta < 1$.[8] Moreover, as the last term in Proposition 1a shows, a deadweight-loss effect for generation $t$ occurs, which is zero at $\tau_{bt} = 0$. Altogether, the welfare effect is ambiguous; it is the more likely to be positive, the smaller is $\delta$ (less double counting) and the smaller is $\beta$ (more discounting of future welfare).

The results of Proposition 1 are analogous to the corresponding ones for a model with an altruistic bequest motive, found in Brunner and Pech (2012). Obviously, if initial wealth is the same for both individuals, then the (positive) redistributive effect disappears for all three taxes. This is the framework of the well-known Atkinson-Stiglitz (1976) result, which tells us that indirect taxes cannot improve social welfare, if they are imposed in addition to an optimal nonlinear labor-income tax and individuals differ merely in earning abilities. Our result shows that the role of indirect taxes can be understood when the framework is extended to include the fact that, as a consequence of wealth transmission in previous generations, differences in initial wealth already exist.

[8] In case of pure altruism when the parents' preferences are dynastic and include welfare of the descendants, these two effects outweigh each other if double counting is avoided completely. With the joy-of-giving motive (also called paternalistic altruism, see Michel and Pestieau, 2004) the relative importance of these two effects is undetermined even if $\delta = 0$ (no double counting).

# 3 The previous generation

Next, we go a step back and consider a previous, third generation $t-1$. That is, we take inheritances $e_{it}$ of generation $t$ no longer as exogenous, but consider explicitly the bequest decisions of generation $t-1$. For simplicity we assume again that the previous generation consists of two individuals only, who differ in earning abilities $\omega_{Lt-1} < \omega_{Ht-1}$ and are endowed with initial wealth $e_{Lt-1}$ and $e_{Ht-1}$. They earn gross income $z_{it-1} = \omega_{it-1} l_{it-1}$ and use net income $x_{it-1}$ together with initial wealth for own consumption $c_{it-1}$ and bequests $b_{it-1}$. Preferences are again described by the utility function $u(c_{it-1}, l_{it-1}) + h(b_{it-1})$. Bequests of generation $t-1$ constitute the inherited wealth of the next generation $t$, where we assume as in Section 2 that each individual of the heir generation $t$ receives some share of the bequests left by each individual of generation $t-1$.

In period $t-1$, the government determines optimal nonlinear income taxes for all three periods $t-1$, $t$ and $t+1$. Moreover, it considers proportional taxes $\tau_{et}$ on inheritances and $\tau_t$ on expenditures of generation $t$, which are assumed to be announced at a time such that generation $t-1$ is able to adapt its behavior. We analyze the welfare consequences of these taxes; in Section 2 these two taxes were shown to be completely equivalent in a model with exogenous inheritances of generation $t$.

How exactly these taxes influence the decision of the generation $t-1$ depends on the bequest motive: in our model, bequests are regarded as some form of consumption; it is the amount left to the heirs, which per se provides utility to the bequeathing individuals. In contrast to the dynastic model, the welfare positions of the descendants are irrelevant for the bequest decision of the parents, and hence taxes affecting only welfare of the descendant generation are also irrelevant.

Still, it has to be discussed how a direct tax $\tau_{et}$, imposed on inherited wealth of generation $t$, affects the previous generation. Taking the bequest-as-consumption model literally, one might again argue that the anticipation of $\tau_{et}$ does not change anything for generation $t-1$, because individuals simply care for what they leave as (gross) bequests to their descendants. On the other hand, however, it seems more plausible to model the bequeathing generation $t-1$ as caring for the amount that actually goes to their heirs, that is, for bequests net of the inheritance tax $\tau_{et}$. Accordingly, we use $b_{it-1}$ as indicating bequests net of an inheritance tax $\tau_{et}$, then gross bequests are $b_{it-1}/(1-\tau_{et})$.[9]

[9] Note that this is consistent with the assumption made in Section 2 that individuals also care for net bequests $b_{it}$ in case of a bequest tax $\tau_{bt}$ (with gross bequests being $b_{it}(1+\tau_{bt})$).

Following these considerations, we define indirect utility of individual $i$ of generation $t-1$ analogous to (1) as:

$$v_{t-1}^{i}(x_{it-1}, z_{it-1}, e_{it-1}, \tau_{et}) \equiv \max\{u(c_{it-1}, z_{it-1}/\omega_{it-1})+ \\ + h(b_{it-1}) | c_{it-1} + b_{it-1}/(1-\tau_{et}) \leq x_{it-1} + e_{it-1}\}. \tag{8}$$

(8) also determines demand functions for consumption $c_{it-1}(\cdot)$ and net bequests $b_{it-1}(\cdot)$. By some way of sharing, the latter result in net inheritances $\varepsilon_{it}(b_{Lt-1}, b_{Ht-1})$ of the next generation $t$, $i = L, H$, such that $\varepsilon_{Lt} + \varepsilon_{Ht} = b_{Lt-1} + b_{Ht-1}$. Indirect utility of generation $t$ is now defined similar to (1) as

$$v_t^i(x_{it}, z_{it}, \varepsilon_{it}, \tau_t) \equiv \max\{u(c_{it}, z_{it}/\omega_{it}) + h(b_{it}) | (1+\tau_t)(c_{it}+b_{it}) \leq x_{it} + \varepsilon_{it}\}. \tag{9}$$

The intertemporal maximization problem of the government reads as (with weights $f_{it-1}$ for individuals of generation $t-1$, public resource requirements $g_{t-1}$, public resource transfer $s_{t-1}$, and $W_{t+1}$ defined as in Section 2):

$$\max_{\substack{x_{it-1}, z_{it-1}, x_{it}, z_{it}, \\ i=L,H, s_{t-1}, s_t}} \sum_{i=L,H} f_{it-1}(v_{t-1}^i(\cdot) + (\delta-1)h(b_{it-1}(\cdot)))+ \tag{10}$$
$$+ \beta \sum_{i=L,H} f_{it}(v_t^i(\cdot) + (\delta-1)h(b_{it}(\cdot)))+$$
$$+ \beta^2 W_{t+1}(b_{Lt}, b_{Ht}, s_t),$$
$$\text{s.t.} \quad v_{t-1}^H(x_{Ht-1}, z_{Ht-1}, e_{Ht-1}, \tau_{et}) \geq \\ v_{t-1}^H(x_{Lt-1}, z_{Lt-1}, e_{Ht-1}, \tau_{et}), \tag{11}$$
$$v_t^H(x_{Ht}, z_{Ht}, \varepsilon_{Ht}, \tau_t) \geq v_t^H(x_{Lt}, z_{Lt}, \varepsilon_{Ht}, \tau_t), \tag{12}$$
$$\sum_{i=L,H} x_{it-1} \leq \sum_{i=L,H} z_{it-1} - g_{t-1} - s_{t-1}, \tag{13}$$
$$\sum_{i=L,H} x_{it} \leq \sum_{i=L,H} z_{it} + T_{et} + \tau_t \sum_{i=L,H} (c_{it} + b_{it}) - \\ - g_t + s_{t-1} - s_t, \tag{14}$$

where $T_{et} \equiv \tau_{et}(\varepsilon_{Lt} + \varepsilon_{Ht})/(1-\tau_{et})$ are the revenues from the inheritance tax and $\delta$ again denotes the extent of laundering out. As in Section 2, we consider a subproblem of (10) - (14) by defining

$$W_t(b_{Lt-1}, b_{Ht-1}, s_{t-1}, T_{et}, \tau_t) \equiv \\ \max_{x_{it}, z_{it}, i=L,H, s_t} \sum_{i=L,H} f_{it}(v_t^i(\cdot) + (\delta-1)h(b_{it}(\cdot))) + \beta W_{t+1}(b_{Lt}, b_{Ht}, s_t), \tag{15}$$
$$\text{s.t.} \qquad (12) \text{ and } (14).$$

Note that the $\varepsilon_{it}, i = L, H$ depend on $b_{Lt-1}, b_{Ht-1}$ and thus on $\tau_{et}$, as does $T_{et}$. We can rewrite the problem (10) - (14) as

$$\max_{\substack{x_{it-1}, z_{it-1}, \\ i=L,H, s_{t-1}}} \quad \sum_{i=L,H} f_{it-1}(v^i_{t-1}(\cdot) + (\delta - 1)h(b_{it-1}(\cdot))) + \beta W_t(\cdot), \tag{16}$$
$$\text{s.t.} \quad (11) \text{ and } (13).$$

In this formulation, the problem is similar to the earlier one given by (7), (3) and (5) for period $t$. We get (let $S_{t-1}$ denote the optimal value function of (16), (11), (13) and $\mu_{t-1}, \mu_t, \lambda_{t-1}$ the Lagrangian multipliers to (11), (12), (13), respectively):

**Proposition 2** *Consider a model with three generations and endogenous inheritances of generation $t$.*

*a) The welfare effect of a tax $\tau_{et}$ on inheritances of generation $t$ reads as*

$$\begin{aligned}\frac{\partial S_{t-1}}{\partial \tau_{et}} = & \frac{\mu_{t-1}}{(1-\tau_{et})^2} \frac{\partial v^H_{t-1}[L]}{\partial x_{Lt-1}} (b_{Ht-1}[L] - b_{Lt-1}) + \beta \sum_{i=L,H} \frac{\partial W_t}{\partial b_{it-1}} \frac{\partial b^{com}_{it-1}}{\partial \tau_{et}} + \\ & + (\delta - 1) \sum_{i=L,H} f_{it-1} h'_{it-1} \frac{\partial b^{com}_{it-1}}{\partial \tau_{et}} + \frac{\lambda_{t-1}\tau_{et}}{1-\tau_{et}} \sum_{i=L,H} \frac{\partial b^{com}_{it-1}}{\partial \tau_{et}}.\end{aligned}$$

*b) The welfare effect of a tax $\tau_t$ on expenditures of generation $t$ reads as*

$$\frac{\partial S_{t-1}}{\partial \tau_t} = \frac{\partial W_t}{\partial \tau_t} = \mu_t \frac{\partial v^H_t[L]}{\partial x_{Lt}} (e_{Ht} - e_{Lt}) \frac{1-\tau_{et}}{1+\tau_t}.$$

**Proof.** See Appendix. ■

The main insight from Proposition 2a is that if the bequeathing individuals care for net bequests, then the tax $\tau_{et}$ on inheritances of generation $t$ is, in fact, a means for increased redistribution within the *previous* generation $t-1$, given that the initial wealth of this generation is positively correlated with abilities. That is, this tax is essentially equivalent to a tax $\tau_{bt-1}$ on bequests of the previous generation, and the same mechanism as described in Section 2 for the effect of the tax $\tau_{bt}$ on bequests of generation $t$ applies, now one period earlier: if $e_{Ht-1} > e_{Lt-1}$ and if preferences are weakly separable between labor and consumption, then indeed $b_{Ht-1}[L] > b_{Lt-1}$ follows and the first term in Proposition 2a is positive. The remaining terms in this formula also show the similarity to the consequences of $\tau_{bt}$: there is a compensated effect due to the distortion of the bequest decision which affects welfare of future generations negatively, but welfare of generation $t-1$ positively, if the social planner wants to avoid double counting of bequests ($\delta < 1$). Finally, a deadweight loss comes in with $\tau_{et} > 0$. Specifically, one observes that a tax rate $\tau_{bt-1} = \tau_{et}/(1-\tau_{et})$ imposes the same burden as $\tau_{et}$; in view of the linear sharing rule and $\varepsilon_{Lt} + \varepsilon_{Ht} = b_{Lt-1} + b_{Ht-1}$, both raise revenues $\tau_{bt-1}(b_{Ht-1} + b_{Lt-1}) = \tau_{et}(e_{Lt} + e_{Ht})$ with gross inheritances $e_{it} = \varepsilon_{it}/(1-\tau_{et})$.

Proposition 2b shows that the consideration of a previous generation, whose bequest decision is influenced by the inheritance tax $\tau_{et}$, destroys the equivalence established in Proposition 1 between this tax and a general expenditure tax $\tau_t$ in a model with fixed inheritances. As already discussed above, it follows from the specific bequest motive that taxes which only affect the well-being of generation $t$, but not the amount of their inheritances, do not influence the behavior of the previous generation. Thus, it turns out that the welfare effect of the tax $\tau_t$ on the expenditures of generation $t$ is unambiguously positive and is indeed the same as found in Proposition 1.

It should also be noted that, for the same reason, a tax $\tau_{bt}$ on bequests left by generation $t$ does not affect the previous generation $t-1$, and the welfare effect found in Proposition 1 also remains valid in the three-generations model. Moreover, given the analogous structure of the maximization problems in this and the foregoing section, a tax $\tau_{et-1}$ on inheritances of generation $t-1$ or a tax $\tau_{t-1}$ on all expenditures of generation $t-1$ have the analogous welfare effects as described in Proposition 1 for generation $t$.

An important aspect of our model is that the government is assumed to be able to transfer resources over generations. As a consequence, it is irrelevant into which generation's budget the revenues from an inheritance or an expenditure tax run; there is always an adequate shifting of resources according to the social discount factor $\beta$. In particular, the mechanism (analogous to that explained for $\tau_{bt}$ below Proposition 1) by which the inheritance tax $\tau_{et}$ allows increased redistribution within generation $t-1$ (given $e_{Ht-1} > e_{Lt-1}$) relies on the ability of the planner to use revenues from a marginal increase of $\tau_{et}$ - in period $t$ - to compensate the individuals in $t-1$ by increasing their net incomes $x_{it-1}$. In other words, the inheritance tax $\tau_{et}$ is completely equivalent to a bequest tax $\tau_{bt-1}$, which would not be the case in a model without public resource shifting over time, where it matters which generation benefits from the tax revenues.[10]

[10] One can, however, show that the *redistributive* effect of $\tau_{et}$ within generation $t-1$ works in the same way if revenues from $\tau_{et}$ go to generation $t$ and there is no possibility of shifting public resources over time. The inheritance tax $\tau_{et}$ discourages mimicking in $t-1$ and gives slack to the self-selection constraint. However, some intergenerational redistribution in favor of generation $t$ clearly arises, compared to a tax $\tau_{bt-1}$ whose revenues go to generation $t-1$.

We close this section by a short discussion on the optimal tax rates. Assuming that the introduction of a tax $\tau_{et}$ has a positive welfare effect (the formula in Proposition 2a is positive when evaluated at $\tau_{et} = 0$), one can find a characterization of the optimal tax rate from the condition that the deadweight loss offsets the welfare effect.[11] For the expenditure tax $\tau_t$ no distortion of the bequest decision occurs, hence there is no balancing effect in our model. In this case, to determine an optimal rate would require an extension of the model by including adverse effects which in reality prevent the government from setting tax rates too high. An immediate possibility is to allow for tax avoidance; such a model was studied in Brunner et al. (2010). Another extension would be to consider a combination of bequest motives: in the present paper we showed that assuming a joy-of-giving bequest motive, in fact, no distortion is associated with an expenditure tax. This result does not, however, hold for a pure altruistic bequest motive (Brunner and Pech, 2012); thus, a balancing effect would arise in a model with a mix of bequest motives.

# 4 Discussion

In the present paper we have studied whether a marginal shift from labor taxation to taxes on the intergenerational transmission of wealth increases social welfare, given that leaving bequests is motivated by joy of giving. For a tax on bequests, imposed on the parent generation, we have found that it indeed allows further redistribution within this generation, on top of what can be attained by a labor-income tax alone, if initial wealth of the more-able individual is larger than that of the less-able individual. The same is true for a tax on inheritances, imposed on the heirs, but anticipated by the parents: it also allows further redistribution within the parent generation. On the other hand, these taxes distort the bequest decision, and, as a consequence, the overall effect on social welfare is ambiguous. This result is similar to what was found in a model with dynastic preferences (pure altruism).

A major difference between the two models arises for the expenditure tax. With an altruistic motive, this tax is completely equivalent to an inheritance tax, both having the two effects just described. With a joy-of-giving motive it is also true - as we have seen - that an expenditure tax on generation $t$ taxes, indirectly, inheritances of generation $t$. It increases social welfare of this generation by allowing further redistribution, given that inherited wealth of the high-able individual is larger than that of the low-able individual. However,

[11] Compare the derivation of the formulas characterizing optimal tax rates in the model with an altruistic bequest motive in Brunner and Pech (2012).

it turns out that no distortive effect on the bequest decision of the previous generation $t-1$ is caused by the expenditure tax in period $t$, in contrast to the inheritance tax. The reason is that with a joy-of-giving motive individuals of generation $t-1$ only care about the net amount left to their descendants. They do not, by definition, care about taxes imposed on the expenditures of generation $t$, even if these expenditures are (partly) financed out of the bequests they leave.

The essential point here is foresight of the individuals. In the altruistic model it is assumed that parents have perfect foresight of all factors influencing the descendants' utility positions, which directly enter the parents' preferences. Thus, taxes imposed on expenditures of the heirs also matter for the parents' decision. No such foresight is assumed with the joy-of-giving motive: only factors which directly affect the bequeathed amount (such as a bequest or inheritance tax), distort the parents' decision.

Ultimately, which model is closer to reality is an empirical question, but it seems to be a quite difficult task to test the extent of foresight over generations. Empirical studies trying to discriminate between the altruistic and the joy-of-giving motive analyze whether parents differentiate bequests according to the recipient child's welfare position, which clearly is a different question than foresight of future taxes.

Our starting point was that in view of the Atkinson-Stiglitz result the role of indirect taxes can only be understood in a model where, as a consequence of wealth transmission over prior generations, differences in initial wealth exist in addition to differences in abilities. Then indirect taxes allow more redistribution than a labor-income tax alone, provided that initial wealth of the more-able individual is larger than that of the less-able individual. Indeed, if differences in initial wealth are neglected, then double counting of the benefits of bequests (which increase parent utility and have a positive external effect on the heirs) leads to the reverse result that a subsidy of bequests is unambiguously desirable (see Kaplow, 2001, Blumkin and Sadka, 2003, and Farhi and Werning, 2010).[12] In our model it is the parameter $\delta$ that captures the degree to which double counting of bequests is laundered out by the social planner. The smaller is $\delta$, the lower is the extent of double counting, and, as Propositions 1a and 2a show, the more probable it is that the positive welfare effect of introducing a bequest or inheritance tax dominates the negative effect on the future generation.

[12] From the perspective of practical economic policy this is a very surprising conclusion. Our appoach puts it into perspective by revealing the redistributive potential of bequest taxation, and provides a theoretical underpinning for the frequently expressed view that bequest taxation enhances "equality of opportunity".

Finally, it should be noted that our analysis remains essentially unchanged if an arbitrary number of generations is considered. On the one hand, it is obvious that with a joy-of-giving motive, taxes imposed on inheritances of some generation $t$ never influence bequest decisions of generations earlier than $t-1$. On the other hand, it is always possible to follow the procedure applied in Section 3, that is, to define a function $W$ that depends on bequests of some generation and captures their welfare consequences for all future generations. Thus, the model underlying Proposition 2 is indeed so general as to allow studying the welfare effects of inheritance taxation for arbitrarily many future generations.

# Appendix

## Proof of Proposition 1

From the Lagrangian to the maximization problem (7), (3) and (5) we derive the first-order conditions with respect to $x_{Lt}$ and $x_{Ht}$,

$$\begin{aligned} &f_{Lt}(\frac{\partial v_t^L}{\partial x_{Lt}} + (\delta-1)h'_{Lt}\frac{\partial b_{Lt}}{\partial x_{Lt}}) + \beta\frac{\partial W_{t+1}}{\partial b_{Lt}}\frac{\partial b_{Lt}}{\partial x_{Lt}} \\ &-\mu_t\frac{\partial v_t^H[L]}{\partial x_{Lt}} + \lambda_t[-1+\tau_{bt}\frac{\partial b_{Lt}}{\partial x_{Lt}} + \tau_t(\frac{\partial c_{Lt}}{\partial x_{Lt}} + \frac{\partial b_{Lt}}{\partial x_{Lt}})] = 0, \end{aligned} \tag{A1}$$

$$\begin{aligned} &f_{Ht}(\frac{\partial v_t^H}{\partial x_{Ht}} + (\delta-1)h'_{Ht}\frac{\partial b_{Ht}}{\partial x_{Ht}}) + \beta\frac{\partial W_{t+1}}{\partial b_{Ht}}\frac{\partial b_{Ht}}{\partial x_{Ht}} \\ &+\mu_t\frac{\partial v_t^H}{\partial x_{Ht}} + \lambda_t[-1+\tau_{bt}\frac{\partial b_{Ht}}{\partial x_{Ht}} + \tau_t(\frac{\partial c_{Ht}}{\partial x_{Ht}} + \frac{\partial b_{Ht}}{\partial x_{Ht}}) = 0. \end{aligned} \tag{A2}$$

a) Let $\tau_t = 0$. The derivative of the optimal value function $S_t$ with respect $\tau_{bt}$ is found by differentiating the Lagrangian:

$$\begin{aligned} \frac{\partial S_t}{\partial \tau_{bt}} = &\sum_{i=L,H}[f_{it}(\frac{\partial v_t^i}{\partial \tau_{bt}} + (\delta-1)h'_{it}\frac{\partial b_{it}}{\partial \tau_{bt}}) + \beta\frac{\partial W_{t+1}}{\partial b_{it}}\frac{\partial b_{it}}{\partial \tau_{bt}}] \\ &+\mu_t(\frac{\partial v_t^H}{\partial \tau_{bt}} - \frac{\partial v_t^H[L]}{\partial \tau_{bt}}) + \lambda_t\sum_{i=L,H}(b_{it} + \tau_{bt}\frac{\partial b_{it}}{\partial \tau_{bt}}). \end{aligned} \tag{A3}$$

We use $\partial v_t^i/\partial\tau_{bt} = -b_{it}\partial v_t^i/\partial x_{it}$, $\partial v_t^H[L]/\partial\tau_{bt} = -b_{Ht}[L]\partial v_t^H[L]/\partial x_{Lt}$ as well as the Slutsky equation $\partial b_{it}/\partial\tau_{bt} = \partial b_{it}^{com}/\partial\tau_{bt} - b_{it}\partial b_{it}/\partial x_{it}$ in order to transform (A3) to

$$\begin{aligned}\frac{\partial S_t}{\partial \tau_{bt}} &= \sum_{i=L,H}[f_{it}(-b_{it}\frac{\partial v_t^i}{\partial x_{it}} + (\delta-1)h'_{it}(\frac{\partial b_{it}^{com}}{\partial \tau_{bt}} - b_{it}\frac{\partial b_{it}}{\partial x_{it}})) \\ &+ \beta\frac{\partial W_{t+1}}{\partial b_{it}}(\frac{\partial b_{it}^{com}}{\partial \tau_{bt}} - b_{it}\frac{\partial b_{it}}{\partial x_{it}})] - \mu_t(b_{Ht}\frac{\partial v_t^H}{\partial x_{Ht}} \\ &- b_{Ht}[L]\frac{\partial v_t^H[L]}{\partial x_{Lt}}) + \lambda_t\sum_{i=L,H}[b_{it} + \tau_{bt}(\frac{\partial b_{it}^{com}}{\partial \tau_{bt}} - b_{it}\frac{\partial b_{it}}{\partial x_{it}})].\end{aligned} \tag{A4}$$

Multiplying (A1) by $b_{Lt}$ and (A2) by $b_{Ht}$ and adding both to (A4) gives us the formula in Proposition 1a.

b) We set $\tau_{bt} = 0$. By use of $\partial v_t^i/\partial \tau_{et} = -e_{it}\partial v_t^i/\partial x_{it}$, $\partial v_t^H[L]/\partial \tau_{et} = -e_{Ht}\partial v_t^H[L]/\partial x_{Lt}$, $\partial c_{it}/\partial \tau_{et} = -e_{it}\partial c_{it}/\partial x_{it}$ and $\partial b_{it}/\partial \tau_{et} = -e_{it}\partial b_{it}/\partial x_{it}$, the welfare effect of $\tau_{et}$ can be written as

$$\begin{aligned}\frac{\partial S_t}{\partial \tau_{et}} &= \sum_{i=L,H}[-f_{it}e_{it}(\frac{\partial v_t^i}{\partial x_{it}} + (\delta-1)h'_{it}\frac{\partial b_{it}}{\partial x_{it}}) - \beta e_{it}\frac{\partial W_{t+1}}{\partial b_{it}}\frac{\partial b_{it}}{\partial x_{it}}] \\ &- \mu_t e_{Ht}(\frac{\partial v_t^H}{\partial x_{Ht}} - \frac{\partial v_t^H[L]}{\partial x_{Lt}}) + \lambda_t\sum_{i=L,H}[e_{it} - \tau_t e_{it}(\frac{\partial c_{it}}{\partial x_{it}} + \frac{\partial b_{it}}{\partial x_{it}})].\end{aligned} \tag{A5}$$

Multiplying (A1) and (A2) by $e_{Lt}$ and $e_{Ht}$, respectively, and adding both to (A5) gives us the formula of Proposition 1b.

c) Let $\tau_{bt} = 0$. We determine the derivative of the optimal value function $S_t$ with respect to $\tau_t$:

$$\begin{aligned}\frac{\partial S_t}{\partial \tau_t} &= \sum_{i=L,H}[f_{it}(\frac{\partial v_t^i}{\partial \tau_t} + (\delta-1)h'_{it}\frac{\partial b_{it}}{\partial \tau_t}) + \beta\frac{\partial W_{t+1}}{\partial b_{it}}\frac{\partial b_{it}}{\partial \tau_t}] + \mu_t(\frac{\partial v_t^H}{\partial \tau_t} \\ &- \frac{\partial v_t^H[L]}{\partial \tau_t}) + \lambda_t\sum_{i=L,H}[(c_{it} + b_{it}) + \tau_t(\frac{\partial c_{it}}{\partial \tau_t} + \frac{\partial b_{it}}{\partial \tau_t})].\end{aligned} \tag{A6}$$

The individual budget equation can be written as $c_{it} + b_{it} = B_{it}$, where $B_{it} \equiv (x_{it} + (1-\tau_{et})e_{it})/(1+\tau_t)$. Thus,

$$\frac{\partial c_{it}}{\partial \tau_t} = \frac{\partial c_{it}}{\partial B_{it}}\frac{\partial B_{it}}{\partial \tau_t} = -(c_{it} + b_{it})\frac{\partial c_{it}}{\partial x_{it}}$$

(use $\partial B_{it}/\partial \tau_t = -(x_{it} + (1-\tau_{et})e_{it})/(1+\tau_t)^2 = -(c_{it} + b_{it})/(1+\tau_t)$ and $\partial c_{it}/\partial x_{it} = (\partial c_{it}/\partial B_{it})/(1+\tau_t)$); equivalently $\partial b_{it}/\partial \tau_t = -(c_{it} +$

$b_{it})\partial b_{it}/\partial x_{it}$. Using these terms, together with $\partial v_t^i/\partial\tau_t = -(c_{it}+b_{it})\partial v_t^i/\partial x_{it}$ and $\partial v_t^H[L]/\partial\tau_t = -(c_{Ht}[L]+b_{Ht}[L])\partial v_t^H[L]/\partial x_{Lt}$ in (A6) yields

$$\begin{aligned}\frac{\partial S_t}{\partial\tau_t} = \sum_{i=L,H}&[-f_{it}(c_{it}+b_{it})(\frac{\partial v_t^i}{\partial x_{it}} + (\delta-1)h'_{it}\frac{\partial b_{it}}{\partial x_{it}}) \\ &-\beta(c_{it}+b_{it})\frac{\partial W_{t+1}}{\partial b_{it}}\frac{\partial b_{it}}{\partial x_{it}}] - \mu_t[(c_{Ht}+b_{Ht})\frac{\partial v_t^H}{\partial x_{Ht}} \\ &-(c_{Ht}[L]+b_{Ht}[L])\frac{\partial v_t^H[L]}{\partial x_{Lt}}] + \lambda_t\sum_{i=L,H}[(c_{it}+b_{it}) \\ &-\tau_t(c_{it}+b_{it})(\frac{\partial c_{it}}{\partial x_{it}} + \frac{\partial b_{it}}{\partial x_{it}})].\end{aligned} \tag{A7}$$

Multiplying (A1) and (A2) by $(c_{Lt}+b_{Lt})$ and $(c_{Ht}+b_{Ht})$, respectively, and adding both to (A7) gives us

$$\frac{\partial S_t}{\partial\tau_t} = \mu_t\frac{\partial v_t^H[L]}{\partial x_{Lt}}(c_{Ht}[L]+b_{Ht}[L]-c_{Lt}-b_{Lt}). \tag{A8}$$

Inserting the budget equations of type $H$, when mimicking, and of type $L$, i. e., $c_{Ht}[L]+b_{Ht}[L] = (x_{Lt}+(1-\tau_{et})e_{Ht})/(1+\tau_t)$ and $c_{Lt}+b_{Lt} = (x_{Lt}+(1-\tau_{et})e_{Lt})/(1+\tau_t)$ into (A8), we obtain the formula of Proposition 1c.

## Proof of Proposition 2

From the Lagrangian of the maximization problem (16), (11) and (13) we derive the first-order conditions with respect to $x_{Lt-1}, x_{Ht-1}$ and $s_{t-1}$ (note that $T_{et}$ can be written as $T_{et} = \tau_{et}\sum_{i=L,H} b_{it-1}/(1-\tau_{et})$):

$$\begin{aligned}f_{Lt-1}(\frac{\partial v_{t-1}^L}{\partial x_{Lt-1}} + (\delta-1)h'_{Lt-1}\frac{\partial b_{Lt-1}}{\partial x_{Lt-1}}) + \beta(\frac{\partial W_t}{\partial b_{Lt-1}}\frac{\partial b_{Lt-1}}{\partial x_{Lt-1}} \\ +\frac{\partial W_t}{\partial T_{et}}\frac{\tau_{et}}{1-\tau_{et}}\frac{\partial b_{Lt-1}}{\partial x_{Lt-1}}) - \mu_{t-1}\frac{\partial v_{t-1}^H[L]}{\partial x_{Lt-1}} - \lambda_{t-1} = 0,\end{aligned} \tag{A9}$$

$$\begin{aligned}f_{Ht-1}\frac{\partial v_{t-1}^H}{\partial x_{Ht-1}} + (\delta-1)h'_{Ht-1}\frac{\partial b_{Ht-1}}{\partial x_{Ht-1}}) + \beta(\frac{\partial W_t}{\partial b_{Ht-1}}\frac{\partial b_{Ht-1}}{\partial x_{Ht-1}} \\ +\frac{\partial W_t}{\partial T_{et}}\frac{\tau_{et}}{1-\tau_{et}}\frac{\partial b_{Ht-1}}{\partial x_{Ht-1}}) + \mu_{t-1}\frac{\partial v_{t-1}^H}{\partial x_{Ht-1}} - \lambda_{t-1} = 0,\end{aligned} \tag{A10}$$

$$\beta\frac{\partial W_t}{\partial s_{t-1}} - \lambda_{t-1} = 0. \tag{A11}$$

a) Differentiation of the Lagrangian with respect to $\tau_{et}$ gives

$$\begin{aligned}\frac{\partial S_{t-1}}{\partial \tau_{et}} = \sum_{i=L,H} f_{it-1}(\frac{\partial v^i_{t-1}}{\partial \tau_{et}} + (\delta-1)h'_{it-1}\frac{\partial b_{it-1}}{\partial \tau_{et}})+ \\ \beta\sum_{i=L,H}[\frac{\partial W_t}{\partial b_{it-1}}\frac{\partial b_{it-1}}{\partial \tau_{et}} + \frac{\partial W_t}{\partial T_{et}}(\frac{b_{it-1}}{(1-\tau_{et})^2} \\ + \frac{\tau_{et}}{1-\tau_{et}}\frac{\partial b_{it-1}}{\partial \tau_{et}})] + \mu_{t-1}(\frac{\partial v^H_{t-1}}{\partial \tau_{et}} - \frac{\partial v^H_{t-1}[L]}{\partial \tau_{et}}).\end{aligned} \tag{A12}$$

We find from definition (8) that $\partial v^i_{t-1}/\partial \tau_{et} = -b_{it-1}/(1-\tau_{et})^2(\partial v^i_{t-1}/\partial x_{it-1})$, $\partial v^H_{t-1}[L]/\partial \tau_{et} = -b_{Ht-1}[L]/(1-\tau_{et})^2(\partial v^H_{t-1}[L]/\partial x_{Lt-1})$. Using these relations, together with the Slutsky equation $\partial b^{com}_{it-1}/\partial \tau_{et} = \partial b_{it-1}/\partial \tau_{et} + b_{it-1}/(1-\tau_{et})^2(\partial b_{it-1}/\partial x_{it-1})$, in (A12), and adding (A9) multiplied by $b_{Lt-1}/(1-\tau_{et})^2$ and (A10) multiplied by $b_{Ht-1}/(1-\tau_{et})^2$ we obtain

$$\begin{aligned}\frac{\partial S_{t-1}}{\partial \tau_{et}} = \sum_{i=L,H} f_{it-1}(\delta-1)h'_{it-1}\frac{\partial b^{com}_{it-1}}{\partial \tau_{et}} + \beta\sum_{i=L,H}[\frac{\partial W_t}{\partial b_{it-1}}\frac{\partial b^{com}_{it-1}}{\partial \tau_{et}} \\ + \frac{\partial W_t}{\partial T_{et}}(\frac{b_{it-1}}{(1-\tau_{et})^2} + \frac{\tau_{et}}{1-\tau_{et}}\frac{\partial b^{com}_{it-1}}{\partial \tau_{et}})] \\ - \lambda_{t-1}\sum_{i=L,H}\frac{b_{it-1}}{(1-\tau_{et})^2} + \mu_{t-1}\frac{\partial v^H_{t-1}}{\partial x_{Lt-1}}\frac{b_{Ht-1}[L]-b_{Lt-1}}{(1-\tau_{et})^2}.\end{aligned} \tag{A13}$$

Finally, observe for the optimal value $W_t$ of the problem (15), (12) and (14) that $\partial W_t/\partial T_{et} = \partial W_t/\partial s_{t-1}$, thus (use (A11)) $\beta\partial W_t/\partial T_{et} = \lambda_{t-1}$. Using this relation in (A13) gives us the formula in Proposition 2a.

b) All $t-1$ variables are independent of $\tau_t$, thus the problem (10) - (14) comes down to problem (15), (12) and (14) for fixed $e_{it}$. Then the proof of Proposition 1c applies.

# References

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

Barro, R. J. (1974), Are Government Bonds Net Wealth? *Journal of Political Economy 82 (6)*, 1095-1117.

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

Boadway, R., Marchand, M. and Pestieau P. (2000), Redistribution with Unobservable Bequests: A Case for Taxing Capital Income, *Scandinavian Journal of Economics 102,* 253-267.

Brunner, J. K. and Pech, S. (2008), Optimum Taxation of Inheritances, *Working paper 0806*, Department of Economics, University of Linz.

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

Brunner, J. K., Eckerstorfer, P. and Pech, S. (2010), Optimal Taxes on Wealth and Consumption in the Presence of Tax Evasion, *Working paper 1004*, Department of Economics, University of Linz.

Cremer, H. and Pestieau, P. (2006), Wealth transfer taxation: a survey of the theoretical literature, in: S. C. Kolm and J. M. Ythier (Eds.), *Handbook on Altruism, Giving and Reciprocity*, Vol. 2, Amsterdam: North Holland, 1108-1134.

Cremer, H., Pestieau, P. and Rochet, J.-Ch. (2001), Direct versus Indirect Taxation: The Design of the Tax Structure Revisited, *International Economic Review 42,* 781-799.

Cremer, H., Pestieau, P. and Rochet, J.-Ch. (2003), Capital Income Taxation When Inherited Wealth is Not Observable, *Journal of Public Economics 87,* 2475-2490.

Farhi, E. and Werning, I. (2010), Progressive Estate Taxation, *Quarterly Journal of Economics 125,* 635-673.

Gale, W. G. and Slemrod, J. (2001), Overview, in: W. G. Gale, J. R. Hines Jr. and J. Slemrod (Eds.), *Rethinking Estate and Gift Taxation*, Washington D. C.: Brookings Institution Press, 1-64.

Hammond, P. J. (1987), Altruism, in: J. Eatwell, M. Milgate and P. Newman (Eds.), *The New Palgrave: A Dictionary of Economics,* First Edition, London: Macmillan Press, 85-87.

Harsanyi, J. C. (1995), A theory of prudential values and a rule utilitarian theory of morality, *Social Choice and Welfare 12*, 319-333.

Kaplow, L. (2001), A Framework for Assessing Estate and Gift Taxation, in: W. G. Gale, J. R. Hines Jr. and J. Slemrod (Eds.), *Rethinking Estate and Gift Taxation*, Washington D. C.: Brookings Institution Press, 164-204.

Kopczuk, W. (2009), Economics of Estate Taxation: A Brief Review of Theory and Evidence, *Tax Law Review 63*, 139-157.

Kopczuk, W. and Lupton, J. P. (2007), To Leave or Not to Leave: The Distribution of Bequest Motives, *Review of Economic Studies 74*, 207-235.

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

Milgrom, P. (1993), Is sympathy an economic value? Philosophy, economics, and the contingent valuation method, in J. A. Hausman (Ed.), *Contingent Valuation: A critical Assessment*, Amsterdam: Elsevier, 417-435.

Mirrlees, J. (1971), An Exploration in the Theory of Optimum Income Taxation, *Review of Economic Studies 38*, 175-208.