DEPARTMENT OF ECONOMICS
JOHANNES KEPLER UNIVERSITY OF LINZ

Optimum taxation of bequests in a model with initial wealth

by

Johann K. Brunner and Susanne Pech

Working Paper No. 1001
January 2010

Johannes Kepler University of Linz
Department of Economics
Altenberger Strasse 69
A-4040 Linz - Auhof, Austria
www.econ.jku.at

johann.brunner@jku.at
phone +43 (0)732-2468-8248
susanne.pech@jku.at
phone +43 (0)732-2468-8593

Optimum taxation of bequests in a model with initial wealth

Johann K. Brunner* and Susanne Pech†

University of Linz‡

January 25, 2010

Abstract

We formulate an optimum-taxation model, where parents leave bequests to their descendants for altruistic reasons. In contrast to the standard model, individuals differ not only in earning abilities, but also in initial (inherited) wealth. In this model a redistributive motive for an inheritance tax - which is equivalent to a uniform tax on all expenditures - arises, given that initial wealth increases with earning abilities. Its introduction increases intertemporal social welfare or has an ambiguous effect, depending on whether the bequeathing generation can adjust their behaviour and whether the external effect related to altruism is accounted for in the social objective.

* *Email:* johann.brunner@jku.at.

† *Email:* susanne.pech@jku.at.

‡ *both:* Department of Economics, Altenberger Straße 69, 4040 Linz, Austria.

1 Introduction

Taxation of estates or inheritances remains to be a heavily discussed issue in tax policy. There are strong movements in many countries to abolish the tax on bequests (it was in fact repealed in Sweden and Austria recently), because it is considered immoral (named a "death tax") and adverse to savings.[1] On the other hand, proponents mainly stress its redistributive effect, they see the tax as an instrument for increasing *equality of opportunity.* The existence of such controversial views may be the consequence of deep-going ideological differences, but it may also be attributed to the missing evidence offered by economists as to the effects of a bequest tax.

In the present paper we want to provide new evidence on this tax by introducing an important aspect into the theoretical analysis, which has been neglected by earlier contributions: as a consequence of having rich or poor parents, individuals are endowed with differing inherited wealth. That is, inheritances create a distinguishing characteristic, which is responsible for inequality within a generation. Indeed, the view that inheritance taxation increases equality of opportunity seems to be based on this observation.

Nevertheless, differences in initial wealth are left out in the usual welfare-theoretic analysis of estate taxation, which is based on the optimum-taxation model in the tradition of Mirrlees (1971) and concentrates on differences in earning abilities only. In such a restricted framework, redistribution is best performed via an income tax alone, there is no need for any indirect tax (Atkinson and Stiglitz 1976). As a consequence, there is no role for a tax on bequests either, because leaving bequests can be seen just as a specific way of spending income, like on consumption of goods[2]. In this framework, even a subsidy on bequests may be considered desirable, if the view that bequests (or, more generally, gifts) create twofold utility, for the donor as well as for the donee, is taken into account - in

[1] But such a tax still exists in most European countries. In the USA, President Obama plans to make the estate tax permanent at a rate of 45% for estates above \$3,5 mill. (Wall Street Journal, Sep. 19, 2009). The rate was 55% in the Clinton era.

[2] To be precise, this result follows, if preferences are weakly separable between consumption and leisure. Otherwise, complementarity or substitutability of some consumption good with leisure plays a role (Corlett and Hague 1953). Saez (2003) considers heterogeneity in tastes and argues, in particular, that more educated individuals have a higher savings rate, which makes taxation of savings desirable. In this paper, we introduce heterogeneity in initial wealth and analyse its consequences.

other words, if a positive externality is attributed to leaving bequests, which calls for a Pigouvian subsidy.[3]

However, the situation is fundamentally different, if the fact that the individuals of some generation are already endowed with (differing) initial wealth, as a result of bequests left by their parents, is introduced into the model: then individuals differ in two characteristics: earning abilities and initial wealth. The aim of this paper is to analyse the role of inheritance taxation, together with optimum taxation of labour income, in an appropriately extended framework. We show that the existence of differing initial wealth matters indeed for determining the welfare effect of inheritance taxation.

To our knowledge, prior contributions did not attempt to provide such an analysis. There are some papers that discuss the consequences of (differing) initial wealth on the structure of indirect taxes and on the desirability of capital income taxation (Boadway et al. 2000, Cremer et al., 2001, 2003). However, these authors assume that bequests are unobservable, they analyse to which extent other taxes can be designed as surrogates. In contrast, we model bequests (that is, initial wealth of the descendants) as being observable (as is labour income), because this is the assumption on which actual tax systems rely.[4]

We assume generally that bequests are motivated by pure altruism of the parents, which means that consumption of their descendants is an argument in their utility function. As is well known, this formulation leads to a model of dynasties (Blumkin and Sadka 2003, among others) and implies, in particular, a precise rule of how estates are allocated to the members of the subsequent generation: at death, parents leave all wealth to their own children.

To keep the analysis as simple as possible, we consider an economy which consists of two individuals, and we begin with a model of two generations. The individuals in the parent generation have differing earning abilities and they use their labour income for consumption and for leaving bequests to their immediate descendants, who live on inheritances only. The planner determines an optimum nonlinear tax on labour income of the

[3] See e. g., Blumkin and Sadka 2003 or Farhi and Werning 2008.

[4] This is not to deny that there are problems of observability. However, in this paper we concentrate on the discussion of whether such a tax is welfare-enhancing, assuming that it can be sufficiently enforced.

parents and considers, in addition, the introduction of a proportional tax on bequests or on inheritances. This is the standard framework for the analysis of bequest taxation[5] - where differences in initial wealth of the parents are completely left out - and we formulate the above-mentioned result that a subsidy on bequests increases intertemporal social welfare, if the external effect is observed, that is, if there is "double-counting" of the welfare effect of bequests.[6] Otherwise it is optimal to have neither a tax nor a subsidy.

Next, we introduce into the model the fact that parents are endowed with given initial wealth. We show that then a (specific) tax on bequests left by the parent generation has an ambiguous effect on intertemporal social welfare. The positive externality associated with bequests calls for a subsidy, as above, but there is also an argument for a tax for redistributive reasons: taxing *bequests of the parents* means indirectly taxing *inheritances (initial wealth) received by the parent generation.* Imposing the tax and redistributing its revenues to the individuals through an appropriate adaptation of the income tax is welfare-increasing, provided that high-able individuals have larger initial wealth than low-able.[7] If the external effect is ignored by the planner, only the latter effect occurs.

Moreover, we show that, for obvious reasons, a direct tax on the given inheritances of the parent generation is definitely desirable, because it allows more redistribution than the optimum labour income tax alone. It has no adverse effects on welfare of later generations, if its revenues are used to adapt the income tax appropriately. What is more surprising, however, is that completely the same result arises for a general tax on all expenditures of the parent generation (that is, on their consumption as well as on the bequests they leave to their descendants). Both taxes are equivalent, though the tax on initial wealth is a lump-sum tax, while the expenditure tax is not.

In a next step we account explicitly for the fact that the parent generation inherited

[5] For our main point, namely the consequences of unequal initial wealth, it is inessential whether one works with a proportional tax (as do Blumkin and Sadka 2003) or a nonlinear (Farhi and Werning 2008).

[6] Double-counting refers to the case that welfare of both generations of the dynasty is summed up in the social objective. As welfare of the first generation already includes welfare of the second generation, the latter is counted twice. For a classification of bequest motives see Cremer and Pestieau (2006).

[7] A positive correlation between initial wealth and abilities appears quite plausible, because empirical evidence shows that individuals with higher income also own more wealth (e.g., Diaz-Gemének et al. 2002) and that a substantial part of wealth results from inheritances (Gale and Scholz 1994). We will assume the existence of such a positive correlation in the following.

their initial wealth from the previous generation. That is, we introduce an earlier third generation (of grandparents) into the model, who also have differing earning abilities, identical to those of their respective descendants, and differing initial wealth (positively correlated with abilities). They also have altruistic preferences, caring for consumption of the following two generations of their dynasty; clearly, bequests left by the grandparent generation constitute initial wealth of the subsequent parent generation. The social planner determines optimum nonlinear labour income taxes for these two generations, knowing that abilities remain the same within a dynasty[8] and being able to credibly commit not to change taxes in the following periods.

In this framework, we first consider the case that grandparents cannot change their behaviour any more when the tax on their bequests (i.e., on inheritances of their descendants) is introduced. We find that this tax still increases social welfare unambiguously, even if welfare of the grandparents (who might be expected to be negatively affected) is included in the social welfare function. The intuitive reason is that altruistic grandparents, caring for consumption of their descendants, recognise the additional redistribution associated with this tax.

However, the situation changes, if we consider a model where the grandparent generation can adapt their decisions and this has to be observed by the planner. As grandparents react in a way which ignores the positive externality of bequests, a negative impact occurs in addition to the positive redistributive effect. Hence, the overall effect on intertemporal social welfare of a tax on inheritances of the parent generation (i.e., on bequests left by the grandparent generation) is ambiguous, if the externality is observed by the planner; otherwise the effect is positive. This result is clearly similar to that found in the two-generation model discussed above (for a tax imposed on the bequests of the parent generation).

On the other hand, if instead a tax on (exogenously given) initial wealth of the grandparent generation or - equivalently - a uniform tax on all their expenditures for consumption and bequests is imposed, and their optimum income tax is adapted appropriately,

[8] Thus, we do not consider the intertemporal wedge related to the "inverse Euler equation", which characterises the optimum allocation, if there is uncertainty over future abilities (see, e. g., Golosov et. al. 2007). In contrast, we concentrate on the pure welfare consequence of (taxation of) inheritances.

this again has a definite positive effect on social welfare, just as described above for such taxes in the two-generation model.

Finally, we show that essentially the same results can be derived in a more complex model with an arbitrary number of individuals and a stochastic relationship between initial wealth and abilities: the already familiar redistributive effect occurs, if expected inheritances increase with abilities.

Altogether, our general conclusion is the following: If one considers a model where the first generation of dynasties is characterised by unequal initial wealth, a redistributive motive for inheritance taxation arises, given that inherited wealth is positively correlated with earning abilities. A negative welfare effect on future generations, because the bequeathing generations react to the tax, counteracts this positive (redistributive) effect, but only if one assumes double-counting of bequests in the social welfare function. Otherwise, there is a definite positive welfare effect.

The plan of the paper is as follows: As a starting point (Section 2) we consider bequest taxation in the standard model with two generations and two types of individuals differing only in earning abilities. In Section 3 the existence of unequal initial wealth is introduced. In Section 4 the consequences of various taxes are analysed in an extended model with three generations. Section 5 contains a generalisation to more types of individuals and a stochastic relation between initial wealth and abilities. Section 6 provides concluding remarks.

2 The standard model

We start with the simplest version of a model of dynasties, similar to that in Blumkin and Sadka (2003) or in the first section of Fahri and Werning (2008). There are only two dynasties (L, H) and each comprises two generations (a parent and a child). We assume that each generation lives for one period and consists of two individuals only. The children do not work at all, they live on inheritances. The parents (generation t), however, do work; they differ in earning abilities $\omega_L < \omega_H$. By working l_{it} units of time they earn gross income $z_{it} = \omega_i l_{it}$ and net income x_{it}, $i = L, H$, which they spend for own consumption c_{it} and bequests b_{it}. Each parent has a single child, to whom she leaves all her bequests. Thus, b_{it} is equal to child consumption c_{it+1}. The government imposes a nonlinear labour income tax in period t and a proportional tax on bequests left by generation t, i.e., on inheritances of generation $t+1$.

Identical preferences of the parents are characterised by pure altruism and can be described by the concave utility function $u(c_{it}, c_{it+1}, l_{it})$, strictly increasing in c_{it} and c_{it+1}, strictly decreasing in l_{it}. Child consumption is assumed to be a normal good, it enters the utility function like own consumption. Utility $U\left(c_{it+1}\right)$ of the child depends only own consumption c_{it+1}, with $U : \mathbb{R} \to \mathbb{R}$ strictly concave and increasing. In later sections, when we introduce a third generation, we will assume additive separability with respect to generations and write the parent's utility as $\widetilde{U}(c_{it}, l_{it}) + \delta U\left(c_{it+1}\right)$, where $0 < \delta \leq 1$ is a discount factor, usually interpreted as representing the "degree of altruism".[9]

Let, for given net income x_{it}, gross income z_{it} and a tax rate τ_{bt} on bequests (= child consumption), indirect utility of a parent for a general utility function u be defined by

$$v_t^i\left(x_{it}, z_{it}, \tau_{bt}\right) \equiv \max\left\{u\left(c_{it}, c_{it+1}, z_{it}/\omega_i\right) \mid c_{it} + \left(1+\tau_{bt}\right)c_{it+1} \leq x_{it}\right\} \tag{1}$$

We first take the tax on bequests as fixed at $\tau_{bt} = 0$ and consider a benevolent government which can impose an optimum nonlinear income tax in order to maximise the welfare of the two generations. This is equivalent to determining two bundles $(x_{Lt,} z_{Lt})$, $(x_{Ht,} z_{Ht})$, subject to a self-selection constraint and the resource constraint. With a social discount factor $\beta \geq 0$ and required government resources g_t the problem reads

$$\max_{(x_{it}, z_{it}), i=L,H} \quad \sum_{i=L,H} f_i v_t^i(\cdot) + \beta \sum_{i=L,H} f_i U\left(c_{it+1}\right), \tag{2}$$

$$\text{s.t.} \quad v_t^H(x_{Ht}, z_{Ht}, \tau_{bt}) \geq v_t^H(x_{Lt}, z_{Lt}, \tau_{bt}), \tag{3}$$

$$x_{Lt} + x_{Ht} \leq z_{Lt} + z_{Ht} + \tau_{bt}(c_{Lt+1} + c_{Ht+1}) - g_t. \tag{4}$$

[9] In a still more specific version, additivity is assumed also between consumption and labour with utility out of consumption being the same for the parent and the child: $u(c_{it}, c_{it+1}, l_{it}) = U\left(c_{it}\right) + \delta U\left(c_{it+1}\right) - h\left(l_{it}\right)$, where $h : \mathbb{R} \to \mathbb{R}$, strictly convex and increasing, describes disutility of labour (see, e.g., Fahri and Werning 2008).

Here we assume that the government puts sufficient weight $f_L > f_H$ on the low-wage individual, such that in the optimum further downward redistribution is desired. Therefore we can neglect the self-selection constraint for the low-wage individual, while the self-selection constraint for the high-wage individual is binding.[10]

Note that in case of $\beta = 0$ the social objective (2) is equal to welfare of the parent generation (which includes welfare of the descendants). In case of $\beta > 0$, the descendants' welfare is included separately as well, which means double-counting, as mentioned in the Introduction.

As a next step we ask how the introduction of a tax on bequests affects social welfare. Let $S_1(\tau_{bt})$ be the optimum value of the foregoing problem and μ the Lagrange multiplier of (3). We find

Proposition 1 *The welfare effect of an introduction of a tax on bequests is*

$$\left.\frac{\partial S_1}{\partial \tau_{bt}}\right|_{\tau_{bt}=0} = \beta \sum_{i=L,H} f_i U'_{it+1} \frac{\partial c^{com}_{it+1}}{\partial \tau_{bt}} + \mu \frac{\partial v^H_t[L]}{\partial x_{Lt}} (c_{Ht+1}[L] - c_{Lt+1}).$$

This effect is negative, given weak separability of preferences between consumption and labour time and $\beta > 0$. It is zero, if $\beta = 0$.

Proof. See Appendix. ■

In this formula, $U'_{it+1} \equiv dU/dc_{it+1}$ for $i = L, H$. The upper index "com" denotes compensated demand and $[L]$ refers to "mimicking", that is, a situation where the high-wage individual opts for the bundle designed for the low-wage individual.

The essential point of this result is that, with the mild assumption of weak separability of preferences between consumption and labour, a subsidy on bequests is welfare increasing: Weak separability implies that the second term in the formula of Proposition 1, i.e. the effect on the self-selection constraint, is zero, because the difference in bequests left by type H, when mimicking type L, and the bequests of type L, is zero: $c_{Ht+1}[L] = c_{Lt+1}$.[11]

[10] We further assume that *agent monotonicity* (Seade 1982) is fulfilled for general preferences $u(c_{it}, c_{it+1}, l_{it})$, i. e., $-(\partial v^L_t/\partial z_{Lt})/(\partial v^L_t/\partial x_{Lt}) > -(\partial v^H_t/\partial z_{Ht})/(\partial v^H_t/\partial x_{Ht})$ at any admissible (x, z).

[11] Mimicking by type H means that she chooses the same bundle of net and gross income as type L. Then the only difference between the two types is in labour supply: type H can earn the same gross income

On the other hand, the first term, the direct welfare effect on the child generation $t+1$ is negative if $\beta > 0$, because the effect of an increase of τ_{bt} on compensated demand for $c_{it+1} (= b_{it})$ is always negative.

This finding is clearly related to the theorem of Atkinson and Stiglitz (1976), which tells us that in case of weak separability an optimum nonlinear income tax is a sufficient instrument for redistribution within a generation, there is no role for a tax on a specific good. The particular issue in the present model is that the good "bequests" (= consumption of the descendant) enters the social welfare function, via both the parent's and the child's utility. As a consequence of this double-counting of c_{it+1}, a subsidy to internalise a positive external effect is desirable.[12] Indeed, if β is zero (no double-counting), we have the Atkinson-Stiglitz outcome.

3 A model with initial wealth

We now introduce the fact that parents already have (differing) initial wealth. Let e_{Lt} and e_{Ht} be initial wealth of the two types, which is inherited from the previous generation, but is taken as exogenous for the moment. We want to clarify the role of three different ways of taxing bequests, namely of a proportional tax τ_{bt} on bequests as before, a proportional tax τ_{et} on initial (inherited) wealth of the parents, and a proportional tax τ_t on all expenditures of the parents, i.e., a uniform tax rate on their own consumption c_{it} and on bequests $b_{it} = c_{it+1}$. The definition of the indirect utility function (1) is modified to

$$v_t^i \left(x_{it}, z_{it}, e_{it}, \tau_{bt}, \tau_{et}, \tau_t\right) \equiv \max \left\{u\left(c_{it}, c_{it+1}, z_{it}/\omega_i\right) \mid (1+\tau_t)(c_{it} + (1+\tau_{bt})\, c_{it+1}) \leq x_{it} + (1-\tau_{et})\, e_{it}\right\}, \tag{5}$$

where we assume that not both τ_{bt} and τ_t exist. As in the model of Section 2, the government determines the labour income tax for given $\tau_{bt} = \tau_{et} = \tau_t = 0$ by maximising

with less working time. But due to weak separability this does not influence the decision of how to spend net income.

[12] Farhi and Werning (2008) show in a model with an optimal nonlinear tax on bequests, that this tax is progressive, that is, the marginal subsidy is lower for high-able individuals.

(2) with respect to the income bundles x_{it}, z_{it}, $i = L, H$, subject to

$$v_t^H(x_{Ht}, z_{Ht}, e_{Ht}, \tau_{bt}, \tau_{et}, \tau_t) \geq v_t^H(x_{Lt}, z_{Lt}, e_{Ht}, \tau_{bt}, \tau_{et}, \tau_t), \tag{6}$$

$$\begin{aligned} x_{Lt} + x_{Ht} \leq z_{Lt} + z_{Ht} + \tau_{et} \sum_{i=L,H} e_{it} + \tau_{bt} \sum_{i=L,H} c_{it+1} + \\ + \tau_t \sum_{i=L,H} (c_{it} + c_{it+1}) - g_t, \end{aligned} \tag{7}$$

Let $S_2(\tau_{bt}, \tau_{et}, \tau_t)$ denote the optimum value of the above maximisation problem with μ as the Lagrange multiplier of the self-selection constraint (6).

Proposition 2 *a) The welfare effect of introducing a tax τ_{bt} on bequests left by the parent generation t reads*

$$\left.\frac{\partial S_2}{\partial \tau_{bt}}\right|_{\tau_{bt}=\tau_{et}=\tau_t=0} = \beta \sum_{i=L,H} f_i U'_{it+1} \frac{\partial c^{com}_{it+1}}{\partial \tau_{bt}} + \mu \frac{\partial v_t^H[L]}{\partial x_{Lt}} (c_{Ht+1}[L] - c_{Lt+1}).$$

In general, the sign of this effect is ambiguous. The first term is negative while the second term is positive, given weak separability of preferences between consumption and labour and if the high-able individual is endowed with more inherited wealth than the low-able. The effect is positive, if $\beta = 0$.

b) The welfare effect of introducing a tax τ_{et} on inherited wealth or a tax τ_t on expenditures of the parent generation t reads

$$\left.\frac{\partial S_2}{\partial \tau_{et}}\right|_{\tau_{bt}=\tau_{et}=\tau_t=0} = \left.\frac{\partial S_2}{\partial \tau_t}\right|_{\tau_{bt}=\tau_{et}=\tau_t=0} = \mu \frac{\partial v_t^H[L]}{\partial x_{Lt}} (e_{Ht} - e_{Lt}).$$

This effect is positive, if the high-able individual is endowed with more inherited wealth than the low-able.

Proof. See Appendix. ■

Thus, the formula describing the effect of τ_{bt} is the same as we found in Section 2. Note, however that now even in case of weak separability we have $c_{Ht+1}[L] > c_{Lt+1}$, if $e_{Ht} > e_{Lt}$. That is, as bequests are assumed to be a normal good, the high-able individual - even when mimicking - will leave more bequests to her descendant, if she is endowed with more initial wealth. In that case, if a tax on bequests is introduced, mimicking becomes less attractive for her, which gives slack to the self-selection constraint and more redistribution via the income tax becomes possible. This positive welfare effect counteracts the consequence of double-counting, which calls for a subsidy (as above the own compensated price effect is negative); the overall welfare effect is ambiguous and depends on the magnitude of both effects.

On the other hand, we find that the consequences of imposing a tax τ_{et} on initial wealth or a tax τ_t on the expenditures of the parent generation t are clear-cut: both increase welfare, if $e_{Ht} > e_{Lt}$. All potentially negative welfare consequences of these two taxes, in particular those on the descendant generation, can be offset by an appropriate adaptation of the nonlinear income tax. It is interesting to observe that the effects of τ_{et} and τ_t are identical, though the first clearly is a lump-sum tax while the second is distorting, because expenditures are endogenous.[13] The positive welfare effect of either tax comes from a relaxation of the self-selection constraint. To get some intuition for this result, consider the effect of a small $\Delta\tau_{et}$, which increases tax revenues by $\Delta\tau_{et}(e_{Lt} + e_{Ht})$. Compensating the individuals by an increase in net income $\Delta x_{it} = \Delta\tau_{et}e_{it}$ would leave welfare of both individuals unchanged. However, as $\Delta x_{Lt} < \Delta x_{Ht}$, if $e_{Lt} < e_{Ht}$, this procedure makes mimicking less attractive and allows, thus, further redistribution of net income by increasing Δx_{Lt} and decreasing Δx_{Ht}. This raises social welfare. Note, moreover, that in the formula in b) the social discount factor β does not appear. That is, the positive effect occurs whether or not there is double-counting.

Finally it should be noted that in this and the following Section 4 we assume that the government does not use information on inherited wealth (which is assumed to be observable) to identify individuals according to their earning ability. Namely, if it is publicly known that the higher-able individual has more inherited wealth, the government could infer the types from the reported amount e_{it} of inheritances and then apply a differentiated lump-sum tax as a first-best instrument. Our assumption that the government does not follow this strategy is in accordance with actual behaviour of tax authorities and is

[13] For a further discussion of this issue see Brunner and Pech (2008).

probably based on the fact that in reality the relation between the two characteristics is stochastic and allows no such identification. Therefore, in Section 5 we drop the simplifying assumption of a fixed relation between observable inherited wealth and unobservable abilities and we show that essentially the same results can be derived in a model with a stochastic relation between initial wealth and abilities.

4 Considering the previous generation

After having demonstrated that (differing) initial wealth of some generation t provides a rationale for bequest taxation, we now introduce into our model the fact that this initial wealth occurs due to bequests left by the previous generation $t-1$. In other words, we take initial wealth e_{Lt}, e_{Ht} no longer as exogenous, but incorporate the decisions of generation $t-1$ and analyse how this affects the welfare consequences of bequest taxation.

We assume again that pure altruism motivates the bequest decision of generation $t-1$. That is, this generation care for own activities as well as for the activities of the following generations. As already mentioned in Section 2, we assume from now on that the utility function is additively separable with respect to generations (similar to Blumkin and Sadka 2003); hence the utility function of an individual i of generation $t-1$ reads $\widetilde{U}(c_{it-1}, l_{it-1}) + \delta\widetilde{U}(c_{it}, l_{it}) + \delta^2 U(c_{it+1})$, with a specific (additive) version being $U(c_{it-1}) - h(l_{it-1}) + \delta U(c_{it}) - \delta h(l_{it}) + \delta^2 U(c_{it+1})$. Now a dynasty comprises three generations, and we assume that also generation $t-1$ consists of two types of individuals, with abilities ω_i, $i = L, H$, resp. Moreover, each individual knows that her descendant, to whom she leaves all her bequests, has the same earning ability. Members of generation $t-1$ have initial wealth e_{Lt-1}, e_{Ht-1}.

In a first step, we are interested in the effects of a tax τ_{et}, imposed on inheritances received by generation t, and of a tax τ_t, imposed on all expenditures of generation t, where we assume that generation $t-1$ are already aware of these taxes. Now indirect utility of generation $t-1$ is defined as

$$v_{t-1}^i(x_{it-1}, z_{it-1}, x_{it}, z_{it}, e_{it-1}, \tau_{et}, \tau_t) \equiv \max\{\widetilde{U}(c_{it-1}, \frac{z_{it-1}}{\omega_i}) + \delta\widetilde{U}(c_{it}, \frac{z_{it}}{\omega_i}) + \delta^2 U(c_{it+1}) \mid c_{it-1} + e_{it} \leq x_{it-1} + e_{it-1}, (1+\tau_t)(c_{it} + c_{it+1}) \leq x_{it} + (1-\tau_{et})e_{it}\}, \tag{8}$$

where for the moment we leave out taxes other than τ_{et}, τ_t. However, note that an inheritance tax τ_{et} in period t is equivalent to a bequest tax τ_{bt-1} in period $t-1$. (To see the relation formally, one has to interpret e_{it} as net bequests, write the budget constraints as $c_{it-1} + (1+\tau_{bt-1})e_{it} \leq x_{it-1} + e_{it-1}$, $c_{it} + c_{it+1} \leq x_{it} + e_{it}$ and set $\tau_{bt-1} = \tau_{et}/(1-\tau_{et})$). An important property of v_{t-1}^i is its recursive structure:

$$v_{t-1}^i(\cdot) = \max\{\widetilde{U}(c_{it-1}, \frac{z_{it-1}}{\omega_i}) + \delta v_t^i(\cdot) \mid c_{it-1} + e_{it} \leq x_{it-1} + e_{it-1}\} \tag{9}$$

with v_t^i being defined as

$$v_t^i(x_{it}, z_{it}, e_{it}, \tau_{et}, \tau_t) \equiv \max\{\widetilde{U}(c_{it}, \frac{z_{it}}{\omega_i}) + \delta U(c_{it+1}) \mid (1+\tau_t)(c_{it} + c_{it+1}) \leq x_{it} + (1-\tau_{et})e_{it}\}. \tag{10}$$

Before we investigate the extended model in detail, it is instructive to ask what is the welfare effect of τ_{et} and τ_t, if we still assume that generation $t-1$ have already made all decisions, in particular concerning bequests. That is, we assume that the bundles $(x_{it-1}, z_{it-1}), i = L, H$, are given (as well as e_{it-1}) and, moreover, that generation $t-1$ have already fixed their bequests $b_{it-1} = e_{it}$ (and own consumption c_{it-1}). The intertemporal social objective (2) is now extended to

$$\sum_{i=L,H} f_i(\beta^{-1} v_{t-1}^i(\cdot) + v_t^i(\cdot) + \beta U(c_{it+1})) \tag{2'}$$

with $f_L > f_H$ denoting welfare weights of the two dynasties, as before. (Note that the government is assumed to maximise social welfare (2') in period t, therefore generation $t-1$ is weighted by β^{-1}, but this clearly inessential.)

Though generation $t-1$ have made all their decisions, they are affected by the introduction of τ_{et} and τ_t, because they care for consumption of their descendants which is altered, if in period t additional redistribution is performed through the introduction of τ_{et}, τ_t, resp. As there is a higher weight on the low-able type in the social welfare function, one can expect that this redistribution has a positive effect, even if generation $t-1$ is included. Indeed, we find that the same formula as in Proposition 2b) applies. Let S_3 be the optimum value function of maximising (2') with respect to $(x_{it}, z_{it}), i = L, H$, subject to (6) and (7). We find

Proposition 3 *Given that generation $t-1$ cannot adjust their behaviour, the welfare effect of introducing a tax τ_{et} on inheritances or a tax τ_t on all expenditures of generation t reads*

$$\left.\frac{\partial S_3}{\partial \tau_{et}}\right|_{\tau_{bt}=\tau_{et}=\tau_t=0} = \left.\frac{\partial S_3}{\partial \tau_t}\right|_{\tau_{bt}=\tau_{et}=\tau_t=0} = \mu \frac{\partial v_t^H[L]}{\partial x_{Lt}}(e_{Ht} - e_{Lt}).$$

This effect is positive, given that within generation t inheritances received by the high-able individual are larger than those received by the low-able individual.

Proof. See Appendix. ■

Now we continue with an analysis of the effects of τ_{et} and τ_t on welfare of the three generations, if generation $t-1$ can adapt their behaviour. For this purpose, we consider a model where the government determines optimum nonlinear income taxes for the generations $t-1$ and t, for given tax rates $\tau_{et} = \tau_t = 0$, which means to find optimum bundles $(x_{it-1}, z_{it-1}), (x_{it}, z_{it})$, $i = L, H$.

In order to keep the structure of the problem as simple as possible, we avoid the implications of uncertainty concerning future abilities. Therefore, as already mentioned, we assume that within a dynasty abilities remain constant over generations, and this is known by the authority. Then, as an important consequence, the planner only has to observe the self-selection constraint for the first generation of the dynasties, later generations cannot mimic, because their abilities are known to be the same as those of their parents.[14]

On the other hand, as is usual in Ramsey-type dynamic problems, we also assume that the government can credibly commit not to change the taxes, which are determined in

[14] See also Golosov et al. 2007, Diamond 2007.

period $t-1$, in the following period t. Otherwise, as the solution of the planner's problem is not time consistent, individuals would expect re-optimisation in period t, which would change their behaviour.

Moreover, we assume that the government has no instrument to transfer resources over time (this is only performed within dynasties). As a consequence, separate resource constraints have to observed for the two periods. Let again β denote the social rate for discounting future generations' welfare. The optimisation problem of the planner reads:

$$\max_{\substack{(x_{it-1},z_{it-1})\\(x_{it},z_{it}),i=1,..,K}} \sum_{i=L,H} [f_i(v^i_{t-1}(\cdot)+\beta v^i_t(\cdot)+\beta^2 U(c_{it+1})] \tag{11}$$

$$\text{s.t.} \quad v^H_{t-1}(x_{Ht-1},z_{Ht-1},x_{Ht},z_{Ht},e_{Ht-1},\tau_{et},\tau_t) \geq \tag{12}$$

$$v^H_{t-1}(x_{Lt-1},z_{Lt-1},x_{Lt},z_{Lt},e_{Ht-1},\tau_{et},\tau_t),$$

$$\sum_{i=L,H} x_{it-1} \leq \sum_{i=L,H} z_{it-1} - g_{t-1}, \tag{13}$$

$$\sum_{i=L,H} x_{it} \leq \sum_{i=L,H} z_{it} + \tau_{et} \sum_{i=L,H} e_{it}(\cdot)+ \tag{14}$$

$$+\tau_t \sum_{i=L,H} (c_{it}(\cdot)+c_{it+1}(\cdot)) - g_t.$$

Let $S_4(\tau_{et},\tau_t)$ denote the optimum value of the above maximisation problem and μ the Lagrange multiplier of the self-selection constraint (12).

Proposition 4 *a) Given that generation $t-1$ can adjust their behaviour, the welfare effect of introducing an inheritance tax τ_{et} in period t reads*

$$\left.\frac{\partial S_4}{\partial \tau_{et}}\right|_{\tau_{et}=\tau_t=0} = \sum_{i=L,H} f_i(\beta\frac{\partial v^i_t}{\partial x_{it}} + \beta^2 U'_{it+1}\frac{\partial c^t_{it+1}}{\partial x_{it}})\frac{\partial \tilde{e}^{com}_{it}}{\partial \tau_{et}} + \mu\frac{\partial v^H_{t-1}[L]}{\partial x_{Lt-1}}(e_{Ht}[L]-e_{Lt}).$$

In general, the sign of this effect is ambiguous. The first term is negative, while the second term is positive, given weak separability between consumption and labour and if within generation $t-1$ the high-able individual is endowed with more inherited wealth than the low-able. If $\beta=0$, then the effect is positive.

b) The welfare effect of introducing a tax τ_t on all expenditures of generation t is the *same:*

$$\left.\frac{\partial S_4}{\partial \tau_t}\right|_{\tau_{et}=\tau_t=0} = \left.\frac{\partial S_4}{\partial \tau_{et}}\right|_{\tau_{et}=\tau_t=0}.$$

Proof. See Appendix. ■

In the formula of Proposition 4, $\widetilde{e}_{it} = e_{it}(1-\tau_{et})$ denotes bequests net of the inheritance tax and $\partial\widetilde{e}_{it}^{com}/\partial\tau_{et}$ is the compensated effect of τ_{et} on net bequests, left by generation $t-1$. $\partial c_{it+1}^{t}/\partial x_{it}$ denotes the effect of x_{it} on consumption of generation $t+1$, as determined by generation t (with given inheritances e_{it}).

It turns out that the condition which is decisive for the introduction of an inheritance tax in period t is analogous to that of Proposition 2a), which refers to an inheritance tax in period $t+1$ (that is, a tax τ_{bt} on generation t's bequests) in the model with two generations only. There is a negative term (the own compensated price effect $\partial\widetilde{e}_{it}^{com}/\partial\tau_{et}$), due to the distortion of the bequest decisions of generation $t-1$. The reduction of bequests now affects welfare of two subsequent generations t and $t+1$. On the other hand, there is a positive effect on the self-selection constraint, as before: given that in generation $t-1$ the high-able type has more initial wealth than the low-able, the former, when mimicking, will choose higher bequests than the latter, given weak separability of preferences. Then the introduction of an inheritance tax τ_{et} allows more redistribution.

Moreover, these welfare effects, found for the inheritance tax τ_{et}, are identical to those of an expenditure tax τ_t. This identity was already found above, in the models of Proposition 2b and 3, where the effect was unambigously positive, as the inheritances of generation t were taken exogenouly given, thus no distortion occured.

In Proposition 3, we found that both taxes τ_{et} and τ_t increase intertemporal social welfare of all three generations, if they are introduced at a point in time when generation $t-1$ cannot change their behaviour (but they are affected by the tax, and this is accounted for in the social welfare function). Hence, at first glance, one might expect that the welfare effect is the more positive, if generation $t-1$ can adapt to the taxes, as they will not react in a way which reduces own welfare. However, Proposition 4 tells us that the overall effect is now ambiguous. This discrepancy can be explained by observing the external effect associated with bequests. By adapting to either tax, the individuals of generation

$t-1$ minimise the loss of their own welfare (which includes welfare of future generations), but not the loss of social welfare. That is, they ignore the external effect arising with double-counting of future generations. If $\beta = 0$, the effect of τ_{et} and τ_t is unambiguously positive.

It should also be mentioned that the tax τ_{et} on inheritances of generation t is indeed equivalent to a tax τ_{bt-1} on bequests left by generation $t-1$ (as mentioned above), also from the perspective of the government. In particular, it is shown in the Appendix that nothing changes with Proposition 4, if it is assumed that the revenues from the tax run into the budget of generation $t-1$ (instead of t), as might be more appropriate for a bequest tax. Intertemporal transfer within a dynasty balances the shift of public resources.

In a final step of this section we turn to the analysis of a tax τ_{et-1} on initial wealth of the first generation and of a uniform tax τ_{t-1} on all expenditures of this generation (i.e. on own consumption c_{it-1} and bequests $b_{it-1} = e_{it}$), and show that their consequences also are analogous to those in the two-generations model. For this, we include taxes τ_{et-1} and τ_{t-1} in the indirect utility function (but neglect τ_{et} and τ_t):

$$v^i_{t-1}(x_{it-1}, z_{it-1}, x_{it}, z_{it}, e_{it-1}, \tau_{et-1}, \tau_{t-1}) \equiv \max\{\widetilde{U}(c_{it-1}, \frac{z_{it-1}}{\omega_i}) + \delta\widetilde{U}(c_{it}, \frac{z_{it}}{\omega_i}) + \tag{15}$$

$$+\delta^2 U(c_{it+1}) \mid (1+\tau_{t-1})(c_{it-1} + e_{it}) \leq x_{it-1} + (1-\tau_{et-1})e_{it-1}, c_{it} + c_{it+1} \leq x_{it} + e_{it}\}$$

Let, for $\tau_{et-1} = \tau_{t-1} = 0$, $S_5(\tau_{et-1}, \tau_{t-1})$ denote the optimum value of the maximisation of (11) (where v^i_{t-1} is defined in (15) and v^i_t is defined in (10) for $\tau_{et} = \tau_t = 0$) with respect to $(x_{it-1}, z_{it-1}), (x_{it}, z_{it})$, $i = L, H$, subject to the self-selection constraint

$$\begin{aligned} v^H_{t-1}(x_{Ht-1}, z_{Ht-1}, x_{Ht}, z_{Ht}, e_{Ht-1}, \tau_{et-1}, \tau_{t-1}) \geq \\ v^H_{t-1}(x_{Lt-1}, z_{Lt-1}, x_{Lt}, z_{Lt}, e_{Ht-1}, \tau_{et-1}, \tau_{t-1}) \end{aligned} \tag{16}$$

and to the resource constraints in period $t-1$ and t

$$\sum_{i=L,H} x_{it-1} \leq \sum_{i=L,H} [z_{it-1} + \tau_{et-1}e_{it-1} + \tau_{t-1}(c_{it-1}(\cdot) + e_{it}(\cdot))] - g_{t-1}, \tag{17}$$

$$\sum_{i=L,H} x_{it} \leq \sum_{i=L,H} z_{it} - g_t. \tag{18}$$

Proposition 5 *The welfare effect of introducing a tax τ_{et-1} on initial wealth or a tax τ_{t-1} on expenditures of generation $t-1$ reads*

$$\left.\frac{\partial S_5}{\partial \tau_{et-1}}\right|_{\tau_{et-1}=\tau_{t-1}=0} = \left.\frac{\partial S_5}{\partial \tau_{t-1}}\right|_{\tau_{et-1}=\tau_{t-1}=0} = \mu \frac{\partial v^H_{t-1}[L]}{\partial x_{Lt-1}}(e_{Ht-1} - e_{Lt-1}).$$

This effect is positive, if within generation $t-1$ the high-able individual has more inherited wealth than the low-able.

Proof. See Appendix. ■

Thus, as in the two-generations model (Proposition 2b), we again find that a tax on (fixed) initial wealth has an unambiguously positive effect on intertemporal welfare of all three generations, if within generation $t-1$ initial wealth of the high-able individual is larger than that of the low-able individual. Moreover, this tax is equivalent to a general tax on all expenditures of generation $t-1$. The increase in welfare is due to the additional redistribution performed through these taxes, as explained earlier. All other welfare consequences, in particular those on the descendant generations t and $t-1$, can be offset by an appropriate adaptation of the nonlinear income tax.[15]

5 Stochastic inheritances

As explained earlier, an objection against the models of Sections 3 and 4 could be that with a fixed one-to-one relation between abilities and inherited wealth it would be possible for the social planner to identify individuals according to their earning abilities by their inherited wealth and to impose a first-best tax. In reality, no tax authority follows this strategy, because the relation between inherited wealth and abilities is not fixed, but stochastic. In order to capture this issue, we now generalise the problem of the previous Section 4 by extending the number of individuals (or dynasties) to some $K \geq 2$ and by assuming a stochastic relationship between abilities and initial wealth. Let $\omega_1 < \omega_2 < ... < \omega_K$ be the earning abilities, again identical for the generations $t-1$ and t. The vector of initial wealth $e_{it-1} \geq 0$ of the individuals $i = 1, ..., K$ is now a random variable,

[15] Let us finally mention that the results of this Section can also be shown to hold in a model where each dynasty comprises N generations (not just three).

with $B \subseteq \mathbb{R}^{K+}$ as its supports and with joint distribution $F : B \rightarrow \mathbb{R}$. We assume that the aggregate amount $e^{ag}_{t-1} \equiv \sum\limits_{i=1}^{K} e_{it-1}$ is the same for any realisation of e_{it-1} and e^{ag}_{t-1} is known by the government. For any vector of realisations e_{it-1}, indirect utility v^i_{t-1} and v^i_t is defined as in (8) and (10), but for convenience we assume in this section that utility is linear in $t+1$-consumption, i.e. $U(c_{it+1}) = \gamma c_{it+1}$, with $\gamma > 0$.

In order to determine the optimum nonlinear income tax for given taxes τ_{et}, τ_t, the planner's problem is to find bundles (x_{it-1}, z_{it-1}), (x_{it}, z_{it}) which maximise expected aggregate welfare (with individual weights $f_1 > f_2 > ... > f_K$):

$$\max_{\substack{(x_{it-1}, z_{it-1}) \\ (x_{it}, z_{it}), i=1,..,K}} \int \sum_{i=1}^{K} f_i[(v^i_{t-1}(e_{it-1}, \cdot) + \beta v^i_t(e_{it}(e_{it-1}, \cdot), \cdot) + \\ + \beta^2 \gamma c_{it+1}(e_{it}(e_{it-1}, \cdot), \cdot)] dF \tag{19}$$

$$\text{s.t. } v^i_{t-1}(x_{it-1}, z_{it-1}, x_{it}, z_{it}, e_{it-1}, \tau_{et}, \tau_t) \geq \\ v^i_{t-1}(x_{i-1,t-1}, z_{i-1,t-1}, x_{i-1t}, z_{i-1t}, e_{it-1}, \tau_{et}, \tau_t), \; i = 2, ..., K, \tag{20}$$

$$\sum_{i=1}^{K} x_{it-1} \leq \sum_{i=1}^{K} z_{it-1} - g_{t-1}, \tag{21}$$

$$\sum_{i=1}^{K} x_{it} \leq \sum_{i=1}^{K} z_{it} + \tau_{et} \int \sum_{i=1}^{K} e_{it}(e_{it-1}, \cdot), \cdot) dF + \\ + \tau_t \int \sum_{i=1}^{K} [c_{it}(e_{it-1}, \cdot), \cdot) + c_{it+1}(e_{it}(e_{it-1} \cdot), \cdot)] dF - g_t. \tag{22}$$

Quasilinear preferences imply that different realisations of e_{it-1} have an income effect only on c_{it+1} (and e_{it}), while c_{it-1} and c_{it} are unaffected. In particular, if we denote by e^0_{it} and c^0_{it+1} inheritances and bequests, resp., of individual i of generation t, for the case $e_{it-1} = 0$, we have for actual values that $e_{it} = e^0_{it} + e_{it-1}$, $c_{it+1} = c^0_{it+1} + e_{it-1}(1 - \tau_{et})/(1 + \tau_t)$. It follows that the self-selection constraints (20) are independent of the particular realisation of e_{it-1}, because the same welfare effect $\gamma e_{it-1}(1 - \tau_{et})/(1 + \tau_t)$ occurs on both sides.[16] Moreover, as aggregate inheritances e^{ag}_{t-1} are taken as constant, the resource constraint

[16] As is well known, it suffices to consider only the self-selection constraints for adjacent individuals. Further, due to the order of the welfare weights, the self-selection constraints for less-able types do not bind in the optimum and can be neglected. We also neglect the second-order condition (i.e. the possibility of bunching).

(22) reduces to

$$\sum_{i=1}^{K} x_{it} \leq \sum_{i=1}^{K} z_{it} + \tau_{et}(\sum_{i=1}^{K} e_{it}^{0}(\cdot) + e_{t-1}^{ag}) + \tau_t \sum_{i=1}^{K} (c_{it}(\cdot) + c_{it+1}^{0}(\cdot) + e_{t-1}^{ag}(1-\tau_{et})/(1+\tau_t)).$$

Thus, the resource constraint (22) is also independent of the realisations e_{it-1}. Let $S_6(\tau_{et}, \tau_t)$ denote the optimum value of (19) - (22) and μ_i the Lagrangian multipliers associated with (20).

Proposition 6 *a) The welfare effect of an introduction of an inheritance tax τ_{et} reads*

$$\left.\frac{\partial S_6}{\partial \tau_{et}}\right|_{\tau_{et}=\tau_t=0} = \beta\gamma(\delta+\beta) \int \sum_{i=1}^{K} f_i \frac{\partial e_{it}(e_{it-1}, \cdot)}{\partial \tau_{et}} dF + \delta^2 \gamma \int \sum_{i=2}^{K} \mu_i (e_{it}[L] - e_{i-1t}) dF$$

In general, the sign of this effect is ambiguous. The first term is negative, while the second is positive, given weak separability between consumption and leisure and if within generation $t-1$ higher-able individuals have, on average, more initial wealth than lower-able.[17] *The effect is positive, if $\beta = 0$.*

b) The welfare effect of an introduction of an expenditure tax τ_t is the same:

$$\left.\frac{\partial S_6}{\partial \tau_{et}}\right|_{\tau_{et}=\tau_t=0} = \left.\frac{\partial S_6}{\partial \tau_t}\right|_{\tau_{et}=\tau_t=0}.$$

Proof. See Appendix. ■

With quasilinear preferences the uncompensated effect of the taxes is equal to the compensated effect and the marginal utility of net income is equal to $\delta^2\gamma$, $\delta\gamma$ and γ for generations $t-1$, t and $t+1$, resp. Moreover, $\partial c_{it+1}^{t} / \partial x_{it} = 1$, thus the formula in Proposition 6 is indeed analogous to that in Proposition 4.

[17] Weak separability implies an additive utility function, as described at the beginning of Section 4. It should be mentioned that this is compatible with our assumption of quasilinearity only if utility out of consumption for the generations $t-1$ and t is strictly concave (and different from $U(c_{t+1}) = \gamma c_{t+1}$). Otherwise, because of $\delta < 1$, no bequests would be chosen.

6 Conclusion

In this paper we have analysed the welfare effects of estate or inheritance taxation in a model, which accounts for the fact that initial wealth constitutes a second distinguishing characteristic of individuals, in addition to earning abilities. In the concluding Section we summarise our results and add some remarks:

Prior studies using an optimum-taxation framework for the analysis of bequest taxation have usually worked with a model in the tradition of Mirrlees (1971), which leaves out differences in initial wealth.[18] Perhaps, one might interpret this model as referring to a hypothetical original state of a society without those differences. But we think that this is not an adequate framework, if an economic appraisal of estate taxation is to be relevant for the current debate about such a tax. We are now in a world where differences in initial wealth - as a result of transfers over generations in the past - already exist. They should be recognized, and we have demonstrated that they matter for our understanding of the consequences of bequest taxation.

Indeed, prior studies did not find a case for taxing estates or inheritances for redistributive reasons, because they neglected differences in initial wealth, but concentrated on bequests as a specific way of using income. This lead them to the Atkinson-Stiglitz type of argument and even to the desirability of a Pigouvian subsidy, in order to correct for the external effect associated with gifts.

In contrast, by taking differences in initial wealth seriously, we were able to show that such a tax has a redistributive effect, which increases intertemporal social welfare, if initial wealth and earning abilities are positively correlated. This may explain why it is frequently regarded as enhancing equality of opportunity. The welfare-increasing effect is unambiguous, if a (proportional) tax on inheritances received from the parents is imposed at a point in time, when the preceding generation cannot react to the tax any more. Otherwise, a second, welfare-decreasing effect (familiar from above, calling for a subsidy) arises, because the preceding generation adapts to the tax in a way which ignores the positive external effect of bequests on later generations.

[18] With the exception of Brunner and Pech (2008).

In general, the sign of the total effect is ambiguous. The size of the second effect depends on the parameter β, which describes the social rate of discounting the welfare of future generations. From another perspective, β measures the extent of double-counting, as welfare of future generations is already accounted for in the utility function of the first generation, given their altruistic preferences. If β is set to zero, all taxes considered in the paper have only a positive, redistributive effect, irrespective of the timing of their introduction.

A particularly interesting result is that in our model a tax on inheritances received by individuals of some generation is completely equivalent to a tax on all their expenditures for own consumption and for their bequests left to the descendants. An adaptation of the optimum nonlinear income tax by the planner allows a compensation of the individuals such that these two taxes have identical consequences on the present, the later and on the previous generation.

In this paper, we studied the welfare consequences of introducing taxes on inheritances or expenditures, but we did not characterise optimum values. These obviously are found by balancing the distorting effect against the redistributive effect. In a broader view, too high tax rates are prevented by the reaction of individuals, who will attempt to conceal the tax base. This issue has been modelled in Brunner et.al (2010).

In a related paper, Brunner and Pech (2008) studied estate or inheritance taxation when bequest are motivated by joy-of-giving instead of altruism. That is, (net) bequests instead of consumption of future generations enter their utility function. Empirically, there is no clear-cut evidence, which of the two motives dominates actual decisions, probably a mixture of them (in combination with accidental bequests) applies (see, e.g., Arrondel and Laferrère 2001, Laitner and Juster 1996, Laitner and Ohlssen 2001 and Wilhelm 1996). The main consequence of the joy-of-giving motive is that bequeathing individuals only care about taxes directly related to their bequests, but do not care about future taxes that are imposed on expenditures of the descendant generation. This causes a difference between inheritance and expenditure taxation and makes the latter a preferable instrument.

Several questions remain for future research. One concerns the possibility of imposing a nonlinear tax on inheritances instead of a proportional one. A further interesting question is what are the welfare consequences of differing initial wealth, when taxes on inheritances or expenditures are introduced permanently, not in a single period only. For such an analysis one needs to deal with the problem of time inconsistency of the planner's solution, as in later periods social welfare can be improved through a new decision on tax rates.[19]

Appendix

Proof of Proposition 1

Let $\tau_{bt} = 0$. From the Lagrangian to the maximisation problem (2), (3) and (4), we derive the first-order conditions with respect to x_{Lt} and x_{Ht},

$$f_L \frac{\partial v_t^L}{\partial x_{Lt}} + \beta f_L U'_{Lt+1} \frac{\partial c_{Lt+1}}{\partial x_{Lt}} - \mu \frac{\partial v_t^H[L]}{\partial x_{Lt}} - \lambda \;=\; 0, \tag{A1}$$

$$f_H \frac{\partial v_t^H}{\partial x_{Ht}} + \beta f_H U'_{Ht+1} \frac{\partial c_{Ht+1}}{\partial x_{Ht}} + \mu \frac{\partial v_t^H}{\partial x_{Ht}} - \lambda \;=\; 0, \tag{A2}$$

where $U'_{it+1} \equiv dU/dc_{it+1}, i = L, H$. The symbol $[L]$ refers to a situation where the high-wage individual opts for the L-bundle (mimicking) and μ, λ denote the Lagrange multipliers corresponding to self-selection constraint (3) and to the resource constraint (4), resp. Using the Envelope Theorem, we get for the optimal value function $S_1(\tau_{bt})$ at $\tau_{bt} = 0$

$$\frac{\partial S_1}{\partial \tau_{bt}} = \sum_{i=L,H} f_i\left(\frac{\partial v_t^i}{\partial \tau_{bt}} + \beta U'_{it+1} \frac{\partial c_{it+1}}{\partial \tau_{bt}}\right) + \mu\left(\frac{\partial v_t^H}{\partial \tau_{bt}} - \frac{\partial v_t^H[L]}{\partial \tau_{bt}}\right) - \lambda(c_{Lt+1} + c_{Ht+1}). \tag{A3}$$

We have $\partial v_t^i / \partial \tau_{bt} = -c_{it+1} \partial v_t^i / \partial x_{it}$ from Roy's identity. Using this and (A1), (A2), multiplied by c_{Lt+1}, c_{Ht+1}, resp., we arrive at

$$\frac{\partial S_1}{\partial \tau_{bt}} = \beta \sum_{i=L,H} f_i U'_{it+1}\left(c_{it+1} \frac{\partial c_{it+1}}{\partial x_{it}} + \frac{\partial c_{it+1}}{\partial \tau_{bt}}\right) + \mu \frac{\partial v_t^H[L]}{\partial x_{Lt}}(c_{Ht+1}[L] - c_{Lt+1}). \tag{A4}$$

Finally, application of the Slutsky-equation gives us the formula of Proposition 1.

[19] A similar problem arises with the result (Chamley 1986, Judd 1985) that taxation of income from capital should be zero in the long run, which also relies on a technology, available for the government, to commit to taxes set in the first period.

Proof of Proposition 2

Throughout the proof, let $\tau_{bt} = \tau_{et} = \tau_t = 0$. The first-order conditions to the maximisation problem (2), (6) and (7) with respect to x_{Lt} and x_{Ht} are the same as (A1), (A2).

a) The formula of Proposition 2a is derived from (A1), (A2) in the same way as shown in the Proof of Proposition 1. The second term is positive, if $e_{Ht} > e_{Lt}$ (c_{it+1} is assumed to be a normal good).

b) By use of the Envelope theorem, we differentiate the optimum value S_2 with respect to τ_{et}, τ_t, resp., to obtain

$$\frac{\partial S_2}{\partial \tau_{et}} = \sum_{i=L,H} f_i\left(\frac{\partial v_t^i}{\partial \tau_{et}} + \beta U'_{it+1}\frac{\partial c_{it+1}}{\partial \tau_{et}}\right) + \mu\left(\frac{\partial v_t^H}{\partial \tau_{et}} - \frac{\partial v_t^H[L]}{\partial \tau_{et}}\right) + \lambda \sum_{i=L,H} e_{it}, \quad \text{(A5)}$$

$$\frac{\partial S_2}{\partial \tau_t} = \sum_{i=L,H} f_i\left(\frac{\partial v_t^i}{\partial \tau_t} + \beta U'_{it+1}\frac{\partial c_{it+1}}{\partial \tau_t}\right) + \mu\left(\frac{\partial v_t^H}{\partial \tau_t} - \frac{\partial v_t^H[L]}{\partial \tau_t}\right) + \quad \text{(A6)}$$
$$+\lambda \sum_{i=L,H}(c_{it} + c_{it+1}).$$

First, we use $\partial v_t^i/\partial \tau_{et} = -e_{it}\partial v_t^i/\partial x_{it}$ and $\partial c_{it+1}/\partial \tau_{et} = -e_{it}\partial c_{it+1}/\partial x_{it}$, together with (A1), (A2), multiplied by e_{Lt}, e_{Ht}, resp., in (A5), to obtain the formula of Proposition 2b.

Next, we rewrite the budget equation of a parent i as $c_{it} + c_{it+1} = m_{it}$, where $m_{it} \equiv (x_{it} + e_{it})/(1+\tau_t)$. We have $\partial m_{it}/\partial \tau_t = -(x_{it} + e_{it})/(1+\tau_2)^2 = -(c_{it} + c_{it+1})/(1+\tau_t)$, and $\partial c_{it+1}/\partial x_{it} = (\partial c_{it+1}/\partial m_{it})/(1+\tau_t))$. Thus, $\partial c_{it+1}/\partial \tau_t = (\partial c_{it+1}/\partial m_{it})\partial m_{it}/\partial \tau_t = -(c_{it} + c_{it+1})\partial c_{it+1}/\partial x_{it}$. Using this term and $\partial v_t^i/\partial \tau_t = -(c_{it} + c_{it+1})\partial v_t^i/\partial x_{it}$, as well as (A1), (A2), multiplied by $(c_{Lt} + c_{Lt+1})$ and $(c_{Ht} + c_{Ht+1})$, resp., we arrive at

$$\frac{\partial S_2}{\partial \tau_t} = \mu\frac{\partial v_t^H[L]}{\partial x_{Lt}}[(c_{Ht}[L] + c_{Ht+1}[L]) - (c_{Lt} + c_{Lt+1})]. \quad \text{(A7)}$$

Substituting the budget equation of individual H when mimicking and that of individual L, i.e. $c_{Ht}[L] + c_{Ht+1}[L] = x_{Lt} + e_{Ht}$ and $c_{Lt} + c_{Lt+1} = x_{Lt} + e_{Lt}$, resp., into (A7), gives us the formula of Proposition 2b.

Proof of Proposition 3

Let $\tau_{et} = \tau_t = 0$. The partial derivative of the optimum value $S_3(\tau_{et}, \tau_t)$ of the maximisation problem (2'), (6) and (7) with respect to τ_{et}, τ_t, resp., is found by differentiating the Lagrangian

$$\begin{aligned}\frac{\partial S_3}{\partial \tau_{et}} &= \sum_{i=L,H} f_i\left(\frac{1}{\beta}\frac{\partial v_{t-1}^i}{\partial \tau_{et}} + \frac{\partial v_t^i}{\partial \tau_{et}} + \beta U'_{it+1}\frac{\partial c_{it+1}}{\partial \tau_{et}}\right) + \\ &\quad + \mu\left(\frac{\partial v_t^H}{\partial \tau_{et}} - \frac{\partial v_t^H[L]}{\partial \tau_{et}}\right) + \lambda \sum_{i=L,H} e_{it},\end{aligned} \tag{A8}$$

$$\begin{aligned}\frac{\partial S_3}{\partial \tau_t} &= \sum_{i=L,H} f_i\left(\frac{1}{\beta}\frac{\partial v_{t-1}^i}{\partial \tau_t} + \frac{\partial v_t^i}{\partial \tau_t} + \beta U'_{it+1}\frac{\partial c_{it+1}}{\partial \tau_t}\right) + \\ &\quad + \mu\left(\frac{\partial v_t^H}{\partial \tau_t} - \frac{\partial v_t^H[L]}{\partial \tau_t}\right) + \lambda \sum_{i=L,H} (c_{it} + c_{it+1}),\end{aligned} \tag{A9}$$

with μ, λ as the Lagrange multipliers corresponding to (6) and (7), resp. Using the recursive structure (9) of v_{t-1}^i, together with the fact that $\widetilde{U}(c_{it-1}, z_{it-1}/\omega_i)$ and e_{it} are fixed (as generation $t-1$ cannot adapt their decisions anymore), we find that $\partial v_{t-1}^i/\partial \tau_{et} = \delta \partial v_t^i/\partial \tau_{et}$, $\partial v_{t-1}^i/\partial \tau_t = \delta \partial v_t^i/\partial \tau_t$ and $\partial v_{t-1}^i/\partial x_{it} = \delta \partial v_t^i/\partial x_{it}$. Applying these formulas, together with the first-order conditions of the maximisation problem (2'), (6) and (7) for x_{Lt} and x_{Ht}, i.e.,

$$\frac{1}{\beta} f_L \frac{\partial v_{t-1}^L}{\partial x_{Lt}} + f_L \frac{\partial v_t^L}{\partial x_{Lt}} + \beta f_L U'_{Lt+1}\frac{\partial c_{Lt+1}}{\partial x_{Lt}} - \mu \frac{\partial v_t^H[L]}{\partial x_{Lt}} - \lambda = 0, \tag{A10}$$

$$\frac{1}{\beta} f_H \frac{\partial v_{t-1}^H}{\partial x_{Ht}} + f_H \frac{\partial v_t^H}{\partial x_{Ht}} + \beta f_H U'_{Ht+1}\frac{\partial c_{Ht+1}}{\partial x_{Ht}} + \mu \frac{\partial v_t^H}{\partial x_{Ht}} - \lambda = 0, \tag{A11}$$

and proceeding as in the Proof of Proposition 2b, we obtain the formula in Proposition 3. For $\partial S_3/\partial \tau_t$ also the formulas above (A7) are used.

Proof of Proposition 4

Throughout the proof, let $\tau_{et} = \tau_t = 0$. The proof proceeds in several steps.

(i) First, we show that τ_{et} and τ_t have the same effect. The first-order conditions of the maximisation problem (11) - (14) for x_{Lt}, x_{Ht} read:

$$f_L[\frac{\partial v_{t-1}^L}{\partial x_{Lt}} + \beta\frac{\partial v_t^L}{\partial x_{Lt}}(1+\frac{\partial e_{Lt}}{\partial x_{Lt}}) + \beta^2 U'_{Lt+1}\frac{\partial c_{Lt+1}}{\partial x_{Lt}}] - \mu\frac{\partial v_{t-1}^H[L]}{\partial x_{Lt}} - \lambda_t = 0, \quad \text{(A12)}$$

$$f_H[\frac{\partial v_{t-1}^H}{\partial x_{Ht}} + \beta\frac{\partial v_t^H}{\partial x_{Ht}}(1+\frac{\partial e_{Ht}}{\partial x_{Ht}}) + \beta^2 U'_{Ht+1}\frac{\partial c_{Ht+1}}{\partial x_{Ht}}] + \mu\frac{\partial v_{t-1}^H}{\partial x_{Ht}} - \lambda_t = 0, \quad \text{(A13)}$$

with μ and λ_t denoting the Lagrange multiplier corresponding to (12) and (14), resp. Note that an increase in x_{it} influences the welfare position v_t^i of an individual i generation t directly (for given e_{it}), but also indirectly, because generation $t-1$ adapt bequests e_{it}.

The derivatives of the optimum value $S_4(\tau_{et}, \tau_t)$ with respect to τ_{et}, τ_t read

$$\begin{aligned}\frac{\partial S_4}{\partial \tau_{et}} &= \sum_{i=L,H} f_i(\frac{\partial v_{t-1}^i}{\partial \tau_{et}} + \beta\frac{\partial v_t^i}{\partial \tau_{et}} + \beta^2 U'_{it+1}\frac{\partial c_{it+1}}{\partial \tau_{et}}) + \qquad \text{(A14)}\\ &\quad +\mu(\frac{\partial v_{t-1}^H}{\partial \tau_{et}} - \frac{\partial v_{t-1}^H[L]}{\partial \tau_{et}}) + \lambda_t \sum_{i=L,H} e_{it}.\\ \frac{\partial S_4}{\partial \tau_t} &= \sum_{i=L,H} f_i(\frac{\partial v_{t-1}^i}{\partial \tau_t} + \beta\frac{\partial v_t^i}{\partial \tau_t} + \beta^2 U'_{it+1}\frac{\partial c_{it+1}}{\partial \tau_t}) + \qquad \text{(A15)}\\ &\quad +\mu(\frac{\partial v_{t-1}^H}{\partial \tau_t} - \frac{\partial v_{t-1}^H[L]}{\partial \tau_t}) + \lambda_t \sum_{i=L,H} (c_{it} + c_{it+1})\end{aligned}$$

Though not written explicitly, the taxes τ_{et}, τ_t, resp., influence welfare of generation t in two ways - directly and indirectly - as described above for x_{it}.[20] Thus, $\partial v_t^i/\partial\tau_{et} = (-e_{it} + \partial e_{it}/\partial\tau_{et})\partial v_t^i/\partial x_{it}$. Moreover, we have $\partial v_{t-1}^i/\partial\tau_{et} = -e_{it}\partial v_{t-1}^i/\partial x_{it}$. Using these expressions, together with the first-order conditions (A12) and (A13) multiplied by e_{Lt}, e_{Ht}, resp., (A14) can be transformed to

$$\begin{aligned}\frac{\partial S_4}{\partial \tau_{et}} &= \sum_{i=L,H} f_i[\beta\frac{\partial v_t^i}{\partial x_{it}}(\frac{\partial e_{it}}{\partial \tau_{et}} + e_{it}\frac{\partial e_{it}}{\partial x_{it}}) + \beta^2 U'_{it+1}(\frac{\partial c_{it+1}}{\partial \tau_{et}} + e_{it}\frac{\partial c_{it+1}}{\partial x_{it}})] + \qquad \text{(A16)}\\ &\quad +\mu\frac{\partial v_{t-1}^H[L]}{\partial x_{Lt}}(e_{Ht}[L] - e_{Lt}).\end{aligned}$$

Further, we have $\partial v_{t-1}^i/\partial\tau_t = -(c_{it} + c_{it+1})\partial v_{t-1}^i/\partial x_{it}$ and, due to the direct and indirect effect, $v_t^i/\partial\tau_t = [-(c_{it} + c_{it+1}) + \partial e_{it}/\partial\tau_t]\partial v_t^i/\partial x_{it}$. Using these expressions, together with the first-order conditions (A12), (A13), multiplied by $(c_{Lt} + c_{Lt+1})$, $(c_{Ht} +$

[20] Note that both effects are behind $\partial c_{it+1}/\partial\tau_{et}$ and $\partial c_{it+1}/\partial\tau_t$ as well, see (A18) below.

c_{Ht+1}) resp., in (A15), we obtain

$$\begin{aligned}\frac{\partial S_4}{\partial \tau_t} &= \sum_{i=L,H} f_i[\beta \frac{\partial v_t^i}{\partial x_{it}}(\frac{\partial e_{it}}{\partial \tau_t} + (c_{it} + c_{it+1})\frac{\partial e_{it}}{\partial x_{it}}) + \\ &\quad + \beta^2 U'_{it+1}(\frac{\partial c_{it+1}}{\partial \tau_t} + (c_{it} + c_{it+1})\frac{\partial c_{it+1}}{\partial x_{it}})] + \\ &\quad + \mu \frac{\partial v_{t-1}^H[L]}{\partial x_{Lt}}[(c_{Ht}[L] + c_{Ht+1}[L]) - (c_{Lt} + c_{Lt+1})].\end{aligned} \tag{A17}$$

By eliminating e_{it} from the two budget constraints in (8), one gets

$$c_{it-1} + \frac{1+\tau_t}{1-\tau_{et}}(c_{it} + c_{it+1}) \leq x_{it-1} + e_{it-1} + \frac{x_{it}}{1-\tau_{et}}.$$

Let $p \equiv (1+\tau_t)/(1-\tau_{et})$ and $B_{t-1} \equiv x_{it-1} + e_{it-1} + x_{it}/(1-\tau_{et})$, then at $\tau_{et} = \tau_t = 0$

$$\begin{aligned}\frac{\partial c_{is}}{\partial \tau_t} &= \frac{\partial c_{is}}{\partial p}, \qquad s = t-1, t, t+1 \\ \frac{\partial c_{is}}{\partial \tau_{et}} &= \frac{\partial c_{is}}{\partial p} + \frac{\partial c_{is}}{\partial B_{t-1}} x_{it}, \qquad s = t-1, t, t+1\end{aligned}$$

which together with $\partial c_{is}/\partial B_{t-1} = \partial c_{is}/\partial x_{it}$ gives us $\partial c_{is}/\partial \tau_{et} = \partial c_{is}/\partial \tau_t + x_{it}\partial c_{is}/\partial x_{it}$. Using this equality for $s = t+1$, together with $e_{it} = c_{it}+c_{it+1}-x_{it}$ (the budget equation for period t at $\tau_{et} = \tau_t = 0$), one finds immediately that $\partial c_{it+1}/\partial \tau_{et} + e_{it}\partial c_{it+1}/\partial x_{it}$ in (A16) is equal to $\partial c_{it+1}/\partial \tau_t + (c_{it} + c_{it+1})\partial c_{it+1}/\partial x_{it}$ in (A17). Moreover, observing from the budget equation $e_{it} = x_{it-1}+e_{it-1}-c_{it-1}$ for period $t-1$ that $\partial e_{it}/\partial \tau_{et} = -\partial c_{it-1}/\partial \tau_{et}$ and $\partial e_{it}/\partial \tau_t = -\partial c_{it-1}/\partial \tau_t$, the same reasoning gives us $\partial e_{it}/\partial \tau_{et} = \partial e_{it}/\partial \tau_t + x_{it}\partial e_{it}/\partial x_{it}$ and hence equality of the corresponding bracket-terms in (A16) and (A17). Finally, we conclude from the period-t budget equations of individidal L and individual H, when mimicking, that $c_{Ht}[L] + c_{Ht+1}[L] - (c_{Lt} + c_{Lt+1}) = e_{Ht}[L] - e_{Lt}$. Altogether, we have shown that the right-hand sides of (A16) andf (A17) are identical, that is, the taxes τ_{et} and τ_t have completely the same effect. In the following we only refer to the effect of τ_{et}.

Let $c^t_{it+1}(\cdot)$ denote child consumption decided by generation t, for given inheritances e_{it} (and $x_{it}, z_{it}, \tau_{et}, \tau_t$). Clearly, if e_{it} is appropriate, c^t_{it+1} is equal to child consumption

c_{it+1} decided by generation $t-1$, due to the recursive structure of utility:

$$c_{it+1}(x_{it-1}, x_{it}, z_{it-1}, z_{it}, e_{it-1}, \tau_{et}, \tau_t) = c^t_{it+1}(x_{it}, z_{it}, e_{it}(\cdot), \tau_{et}, \tau_t), \tag{A18}$$

with $e_{it}(\cdot)$ having the same arguments as $c_{it+1}(\cdot)$. Thus, $\partial c_{it+1}/\partial \tau_{et} = (\partial c^t_{it+1}/\partial x_{it})(-e_{it} + \partial e_{it}/\partial \tau_{et})$ and $\partial c_{it+1}/\partial x_{it} = \partial c^t_{it+1}/\partial x_{it} + (\partial c^t_{it+1}/\partial x_{it})\partial e_{it}/\partial x_{it}$ (note that $\partial c^t_{it+1}/\partial e_{it} = \partial c^t_{it+1}/\partial x_{it}$). Substituting these expressions into (A16) gives

$$\begin{aligned} \frac{\partial S_4}{\partial \tau_{et}} &= \sum_{i=L,H} f_i(\beta \frac{\partial v^i_t}{\partial x_{it}} + \beta^2 U'_{it+1} \frac{\partial c^t_{it+1}}{\partial x_{it}})[\frac{\partial e_{it}}{\partial \tau_{et}} + e_{it}\frac{\partial e_{it}}{\partial x_{it}}] + \\ &\quad + \mu \frac{\partial v^H_{t-1}[L]}{\partial x_{Lt}}(e_{Ht}[L] - e_{Lt}). \end{aligned} \tag{A19}$$

(ii) Next, we show that $\partial e_{it}/\partial \tau_{et} + e_{it}\partial e_{it}/\partial x_{it}$ in the square brackets in (A19) is, at $\tau_{et} = 0$, equal to the own compensated price effect $\partial \widetilde{e}_{it}/\partial \tau_{et}$, which is negative. (Remember that $\widetilde{e}_{it} = e_{it}(1-\tau_{et})$ denotes bequests net of the inheritance tax.) To do so, we make use of the recursive structure of indirect utility (compare (9) in the text), which allows us to reformulate the maximisation problem of individual i of generation $t-1$ as

$$v^i_{t-1}(\cdot) = \max\{u(c_{it-1}, \frac{z_{it-1}}{\omega_i}) + \delta v^i_t(\cdot) \mid c_{it-1} + \frac{\widetilde{e}_{it}}{1-\tau_{et}} \leq x_{it-1} + e_{it-1}\}. \tag{A20}$$

This is a standard textbook problem and we can apply the Slutzky equation directly for $\widetilde{e}_{it}$, to get

$$\frac{\partial \widetilde{e}^{com}_{it}}{\partial \tau_{et}} = \frac{\partial \widetilde{e}_{it}}{\partial \tau_{et}} + \frac{\widetilde{e}_{it}}{(1-\tau_{et})^2}\frac{\partial \widetilde{e}_{it}}{\partial x_{it-1}}, \tag{A21}$$

knowing that the expenditure function has its standard properties with the compensated own-price effect being negative, i.e. $\partial \widetilde{e}^{com}_{it}/\partial \tau_{et} < 0$.

Further, by use of $\widetilde{e}_{it} \equiv (1-\tau_{et})e_{it}$, we find that $\partial \widetilde{e}_{it}/\partial \tau_{et} = -e_{it} + \partial e_{it}/\partial \tau_{et}$ and $\partial \widetilde{e}_{it}/\partial x_{it-1} = \partial e_{it}/\partial x_{it-1}$. Moreover, $\partial e_{it}/\partial x_{it-1} = 1 + \partial e_{it}/\partial x_{it}$ (differentiate the budget equation $c_{it-1} + e_{it} = x_{it-1} + e_{it-1}$ with respect to x_{it-1},x_{it}, resp., i.e., $\partial c_{it-1}/\partial x_{it-1} + \partial e_{it}/\partial x_{it-1} = 1$ and $\partial c_{it-1}/\partial x_{it} + \partial e_{it}/\partial x_{it} = 0$ and use $\partial c_{it-1}/\partial x_{it-1} = \partial c_{it-1}/\partial x_{it}$).

Substituting these expressions into (A21) gives us

$$\frac{\partial \widetilde{e_{it}^{com}}}{\partial \tau_{et}} = -e_{it} + \frac{\partial e_{it}}{\partial \tau_{et}} + e_{it}(1 + \frac{\partial e_{it}}{\partial x_{it}}) = \frac{\partial e_{it}}{\partial \tau_{et}} + e_{it}\frac{\partial e_{it}}{\partial x_{it}}. \tag{A22}$$

(iii) Finally we show that the welfare effects found in Proposition 4 do not change when the revenues from τ_{et} (or τ_t) are assumed to run into the resource constraint of period $t-1$ instead of that of period t. In this case, the resource constraints are modified to

$$\sum_{i=L,H} x_{it-1} \leq \sum_{i=L,H} z_{it-1} + \tau_{et} \sum_{i=L,H} e_{it} + \tau_t \sum_{i=L,H} (c_{it}(\cdot) + c_{it+1}(\cdot)) - g_{t-1}, \tag{A23}$$

$$\sum_{i=L,H} x_{it} \leq \sum_{i=L,H} z_{it} - g_t. \tag{A24}$$

One observes immediately that the derivatives of the optimum value function of maximizing (11) subject to (A23), (A24) and to the self-selection contraint (12) with respect to τ_{et}, τ_t are the same as (A14), (A15), if λ_t is replaced by λ_{t-1} (the multiplier associated with (A23)). Moreover, the F.O.C.'s with respect to x_{Lt}, x_{Ht} are the same as (A12), (A13), while the F.O.C.'s with respect to x_{Lt-1}, x_{Ht-1} read:

$$f_L[\frac{\partial v_{t-1}^L}{\partial x_{Lt-1}} + \beta\frac{\partial v_t^L}{\partial x_{Lt}}\frac{\partial e_{Lt}}{\partial x_{Lt-1}} + \beta^2 U'_{Lt+1}\frac{\partial c_{Lt+1}}{\partial x_{Lt-1}}] - \mu\frac{\partial v_{t-1}^H[L]}{\partial x_{Lt-1}} - \lambda_{t-1} = 0 \tag{A25}$$

$$f_H[\frac{\partial v_{t-1}^H}{\partial x_{Ht-1}} + \beta\frac{\partial v_t^H}{\partial x_{Ht}}\frac{\partial e_{Ht}}{\partial x_{Ht-1}} + \beta^2 U'_{Ht+1}\frac{\partial c_{Ht+1}}{\partial x_{Ht-1}}] + \mu\frac{\partial v_{t-1}^H}{\partial x_{Ht-1}} - \lambda_{t-1} = 0 \tag{A26}$$

We have $\partial v_{t-1}^i/\partial x_{it-1} = \partial v_{t-1}^i/\partial x_{it}$, $\partial v_t^i/\partial x_{it-1} = \partial v_t^i/\partial x_{it}$, $\partial e_{it}/\partial x_{it-1} = 1 + \partial e_{it}/\partial x_{it}$ and $\partial c_{it+1}/\partial x_{it-1} = \partial c_{it+1}/\partial x_{it}$. Using these expressions, (A25) can be transformed to

$$f_L[\frac{\partial v_{t-1}^L}{\partial x_{Lt}} + \beta\frac{\partial v_t^L}{\partial x_{Lt}}(1 + \frac{\partial e_{Lt}}{\partial x_{Lt}}) + \beta^2 U'_{Lt+1}\frac{\partial c_{Lt+1}}{\partial x_{Lt}}] - \mu\frac{\partial v_{t-1}^H[L]}{\partial x_{Lt}} - \lambda_{t-1} = 0. \tag{A27}$$

Comparing (A12) and (A27), it follows that $\lambda_t = \lambda_{t-1}$. Inspection of the foregoing proof shows that the results remain valid.

Proof of Proposition 5

Note first that for the maximisation problem (11), (16) - (18) the first-order conditions look the same as those of the foregoing problem (11) - (14), as in either case all taxes $\tau_{et-1}, \tau_{t-1}, \tau_{et}, \tau_t$ are assumed to be zero. We use the Envelope Theorem to derive

$$\begin{aligned}\frac{\partial S_5}{\partial \tau_{et-1}} &= \sum_{i=L,H} f_i(\frac{\partial v_{t-1}^i}{\partial \tau_{et-1}} + \beta \frac{\partial v_t^i}{\partial \tau_{et-1}} + \beta^2 U'_{it+1} \frac{\partial c_{it+1}}{\partial \tau_{et-1}}) + \\ &\quad + \mu(\frac{\partial v_{t-1}^H}{\partial \tau_{et-1}} - \frac{\partial v_{t-1}^H[L]}{\partial \tau_{et-1}}) + \lambda_{t-1} \sum_{i=L,H} e_{it-1}\end{aligned} \tag{A28}$$

We have $\partial v_{t-1}^i / \partial \tau_{et-1} = -e_{it-1} \partial v_{t-1}^i / \partial x_{it-1}$ and $\partial v_t^i / \partial \tau_{et-1} = (\partial v_t^i / \partial x_{it}) \partial e_{it} / \partial \tau_{et-1}$, $\partial e_{it} / \partial \tau_{et-1} = -e_{it-1} \partial e_{it} / \partial x_{it-1}$ and $\partial c_{it+1} / \partial \tau_{et-1} = -e_{it-1} \partial c_{it+1} / \partial x_{it-1}$. Using these expressions and (A25), (A26), multiplied by e_{Lt-1}, e_{Ht-1}, resp., in (A28), gives us the formula of Proposition 5.

Further we find that

$$\begin{aligned}\frac{\partial S_5}{\partial \tau_{t-1}} &= \sum_{i=L,H} f_i(\frac{\partial v_{t-1}^i}{\partial \tau_{t-1}} + \beta \frac{\partial v_t^i}{\partial \tau_{t-1}} + \beta^2 U'_{it+1} \frac{\partial c_{it+1}}{\partial \tau_{t-1}}) + \\ &\quad \mu(\frac{\partial v_{t-1}^H}{\partial \tau_{t-1}} - \frac{\partial v_{t-1}^H[L]}{\partial \tau_{t-1}}) + \lambda_{t-1} \sum_{i=L,H} (c_{it-1} + e_{it})\end{aligned} \tag{A29}$$

We have $\partial v_{t-1}^i / \partial \tau_{t-1} = -(c_{it-1} + e_{it}) \partial v_{t-1}^i / \partial x_{it-1}$ and $\partial v_t^i / \partial \tau_{t-1} = (\partial v_t^i / \partial x_{it}) \partial e_{it} / \partial \tau_{t-1}$. Using this, the formulas (A25), (A26) multiplied by $(c_{Lt-1} + e_{Lt}), (c_{Ht-1} + e_{Ht})$, resp., and the fact that $\partial Z / \partial \tau_{t-1} = (c_{it-1} + e_{it}) \partial Z / \partial x_{t-1}, Z = c_{it}, c_{it+1}, e_{it}$ from inspection of the combined budget constraint gives us

$$\frac{\partial S_5}{\partial \tau_{t-1}} = \mu \frac{\partial v_1^H[L]}{\partial x_1} [(c_{Ht-1}[L] + e_{Ht}[L]) - (c_{Lt-1} + e_{Lt})], \tag{A30}$$

which is equal to the formula of Proposition 5 (use the budget equations for period $t-1$).

Proof of Proposition 6

a) The first-order-conditions with respect to x_{it} of (19) - (22) read, at $\tau_{et} = \tau_t = 0$:

$$\int f_i[\frac{\partial v^i_{t-1}}{\partial x_{it}} + \beta(\frac{\partial v^i_t}{\partial x_{it}} + \frac{\partial v^i_t}{\partial e_{it}}\frac{\partial e_{it}}{\partial x_{it}}) + \beta^2\gamma\frac{\partial c_{it+1}}{\partial x_{it}}]dF + \\ + \mu_i\frac{\partial v^i_{t-1}}{\partial x_{it}} - \mu_{i+1}\frac{\partial v^{i+1}_{t-1}[L]}{\partial x_{it}} - \lambda_t = 0, \quad i = 1, ...K, \tag{A31}$$

where $\mu_i, i = 2, ..., K$ and λ_t denote the Lagrangian variables referring to the self-selection constraints (20) (μ_1 and μ_{K+1} are set to zero), and to the resource constraint (22), resp. $[L]$ indicates mimicking, as before. Note that, as in the Proof of Proposition 4, both - the direct and the indirect - effects of x_{it} on the welfare positions of generation t are written explicitly. The Envelope theorem implies

$$\frac{\partial S_6}{\partial \tau_{et}} = \int \sum_{i=1}^{K} f_i[\frac{\partial v^i_{t-1}}{\partial \tau_{et}} + \beta\frac{\partial v^i_t}{\partial \tau_{et}} + \beta^2\gamma\frac{\partial c_{it+1}}{\partial \tau_{et}}]dF + \\ + \sum_{i=2}^{K}\mu_i(\frac{\partial v^i_{t-1}}{\partial \tau_{et}} - \frac{\partial v^i_{t-1}[L]}{\partial \tau_{et}}) + \lambda_t\sum_{i=1}^{K} e_{it} \tag{A32}$$

Due to quasilinear utility, we have at $\tau_{et} = \tau_t = 0$: $\partial v^i_{t-1}/\partial x_{it} = \delta^2\gamma$, $\partial v^i_t/\partial x_{it} = \delta\gamma$, $\partial e_{it}/\partial x_{it} = 0$ and $\partial c_{it+1}/\partial x_{it} = 1$. Using this, (A31) is transformed to

$$-f_i[\delta^2\gamma + \beta\delta\gamma + \beta^2\gamma] = \mu_i\delta^2\gamma - \mu_{i+1}\delta^2\gamma - \lambda_t, \tag{A33}$$

which is independent of e_{it-1}. Multiplying both sides of (A33) with arbitrary e_{it} and summing up we get for the expected value

$$-\int \sum_{i=1}^{K} e_{it} f_i[\delta^2\gamma + \beta\delta\gamma + \beta^2\gamma]dF = \int \sum_{i=1}^{K} e_{it}\delta^2\gamma(\mu_i - \mu_{i+1})dF - \\ -\lambda_t \int \sum_{i=1}^{K} e_{it} dF \tag{A34}$$

Quasilinear utility implies $\partial v^i_{t-1}/\partial \tau_{et} = -e_{it}\delta^2\gamma$, $\partial v^i_{t-1}[L]/\partial \tau_{et} = -e_{it}[L]\delta^2\gamma$, $\partial v^i_t/\partial \tau_{et} = \delta\gamma(-e_{it} + \partial e_{it}/\partial \tau_{et})$ and $\partial c_{it+1}/\partial \tau_{et} = -e_{it} + \partial e_{it}/\partial \tau_{et}$ (observe the direct and indirect effects of τ_{et} on v^i_t and c_{it+1}). Using these terms in (A32) and substituting (A34) into

(A32) gives (note that $\sum_{i=1}^{K} e_{it}$ is independent of the realisations of e_{it-1}, by assumption)

$$\begin{aligned}\frac{\partial S_6}{\partial \tau_{et}} &= \int \sum_{i=1}^{K} f_i \beta\gamma(\delta+\beta)\frac{\partial e_{it}}{\partial \tau_{et}} dF + \sum_{i=2}^{K} \mu_i \delta^2 \gamma(-e_{it} + e_{it}[L]) + \\ &+ \int \sum_{i=1}^{K} e_{it} \delta^2 \gamma(\mu_i - \mu_{i+1}) dF\end{aligned} \tag{A35}$$

Finally, we know that $-e_{it}+e_{it}[L]$ is independent of the realisation of e_{it-1}, thus the second term on the RHS of (A35) can be written as expected value $\int \sum_{i=2}^{K} \mu_i \delta^2 \gamma(-e_{it} + e_{it}[L]) dF$, which gives us the formula in Proposition 6a.

b) The effect of τ_t is

$$\begin{aligned}\frac{\partial S_6}{\partial \tau_t} = \int \sum_{i=1}^{K} f_i [\frac{\partial v_{t-1}^i}{\partial \tau_t} + \beta \frac{\partial v_t^i}{\partial \tau_t} + \beta^2 \gamma \frac{\partial c_{it+1}}{\partial \tau_t}] dF + \\ + \sum_{i=2}^{K} \mu_i (\frac{\partial v_{t-1}^i}{\partial \tau_t} - \frac{\partial v_{t-1}^i[L]}{\partial \tau_t}) + \lambda_t \sum_{i=1}^{K} e_{it}.\end{aligned} \tag{A36}$$

Similar to the procedure above, we can multiply both sides of (A33) with $c_{it} + c_{it+1}$ and get for the expected value

$$\begin{aligned}-\int \sum_{i=1}^{K} (c_{it} + c_{it+1}) f_i [\delta^2\gamma + \beta\delta\gamma + \beta^2\gamma] dF = \\ \int \sum_{i=1}^{K} (c_{it} + c_{it+1}) \delta^2 \gamma(\mu_i - \mu_{i+1}) dF - \lambda_t \int \sum_{i=1}^{K} (c_{it} + c_{it+1}) dF.\end{aligned} \tag{A37}$$

Quasilinear preferences imply $\partial v_{t-1}^i / \partial \tau_t = -\delta^2 \gamma (c_{it} + c_{it+1})$, $\partial v_{t-1}^i[L]/\partial \tau_t = -\delta^2 \gamma (c_{it}[L] + c_{it+1}[L])$, $\partial v_t^i / \partial \tau_t = \delta\gamma(-c_{it} - c_{it+1} + \partial e_{it}/\partial \tau_t)$ and $\partial c_{it+1}/\partial \tau_t = -c_{it} - c_{it+1} + \partial e_{it}/\partial \tau_t$. Substituting these terms, together with (A37), into (A36), (A36) can be written as

$$\begin{aligned}\frac{\partial S_6}{\partial \tau_t} = \int \sum_{i=1}^{K} \beta\gamma(\delta+\beta) \frac{\partial e_{it}}{\partial \tau_t} dF + \int \sum_{i=2}^{K} \mu_i \delta^2 \gamma(-c_{it} - c_{it+1} + c_{it}[L] + \\ + c_{it+1}[L]) dF + \int \sum_{i=1}^{K} (c_{it} + c_{it+1}) \delta^2 \gamma (\mu_i - \mu_{i+1}) dF\end{aligned} \tag{A38}$$

Similar transformations of (A38) as in a) lead to the formula in Proposition 6b. Note that in case of mimicking the $i+1$-individual receives net income x_{it}, therefore the budget equation for period t, $c_{it} + c_{it+1} = x_{it} + e_{it}$, implies $c_{i+1t}[L] + c_{i+1,t+1}[L] - c_{it} - c_{it+1} = e_{i+1t}[L] - e_{it}$. Moreover, with quasilinear utility, $\partial e_{it}/\partial\tau_{et}$ and $\partial e_{it}/\partial\tau_t$ are identical to the compensated effects and these two are already known to be equal from the proof of Proposition 4.

References

Arrondel, L. and A. Laferrère (2001), Taxation and wealth transmission in France, *Journal of Public Economics 79*, 3-33.

Atkinson, A. B. and J. E. Stiglitz (1976), The design of tax structure: direct versus indirect taxation, *Journal of Public Economics 6*, 55-75.

Blumkin, T. and E. Sadka (2003), Estate taxation with intended and accidental bequests, *Journal of Public Economics 88*, 1-21.

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

Brunner, J.K., P. Eckerstorfer and S. Pech (2010), Optimal taxes on wealth and expenditures in the presence of tax evasion, *mimeo*, University of Linz.

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

Chamley, C. (1986), Optimal taxation of capital income in general equilibrium with infinite lives, *Econometrica*, Vol. 54, 607-622.

Corlett, W.J. and D. C. Hague (1953), Complementarity and the excess burden of taxation, *Review of Economic Studies,* Vol. 21, 21-30.

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

Cremer, H., Pestieau, P. and J.-Ch. Rochet (2003), Capital income taxation when inherited wealth is not observable, *Journal of Public Economics*, Vol. 87., 2475-2490.

Cremer, H., Pestieau, P. (2006), Wealth transfer taxation: a survey of the theoretical literature, in: Kolm, S.-Ch., Ythier, J. M. (Eds.). *Handbook of Economics of Giving, Altruism and Reciprocity*, Vol. 2. Amsterdam: North-Holland, 1107-1134.

Diamond, P. (2007), Comment to Golosov, M., Tsyvinski, A. and I. Werning (2007), New Dynamic Public Finance: A User's Guide, in: Acemoglu, D., Rogoff, K. and M. Woodford (2007) (Eds.). *NBER Macroeconomics Annual 2006,* Vol. 21.

Díaz-Giménez, J., Quadrini, V., Ríos-Rull, J.-V., Budriá Rodriguez, S.B., (2002), Updated Facts on U.S. Distributions of Earnings, Income and Wealth, *Federal Reserve Bank of Minneapolis Review*, Vol. 26, No. 3, 2-35.

Fahri, E. and I. Werning (2008), Progressive estate taxation, *mimeo.*

Gale W. G. and J. K. Scholz (1994), Intergenerational Transfers and the Accumulation of Wealth, *Journal of Economic Perspectives 8(4)*, 145-160.

Golosov, M., Tsyvinski, A. and I. Werning (2007), New Dynamic Public Finance: A User's Guide, in: Acemoglu, D., Rogoff, K. and M. Woodford (2007) (Eds.). *NBER Macroeconomics Annual 2006,* Vol. 21.

Judd, K. L. (1985), Redistributive taxation in a simple perfect foresight model, *Journal of Public Economics*, Vol. 28, 59-83.

Laitner J.. and F. T. Juster (1996), New Evidence on Altruism: A Study of TIAAF-CREF Retirees, *American Economic Review 86 (4)*, 893-908.

Laitner, J. and H. Ohlsson (2001), Bequest motives: a comparison of Sweden and the United States, *Journal of Public Economics 79*, 205-236.

Mirrlees, J. (1971), An exploration in the theory of optimum income taxation, *Review of Economic Studies 38*, 175-208.

Saez, E. (2002), The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. *Journal of Public Economics 83*, 217-230.

Seade, J. (1982), On the sign of the optimum marginal income taxation, *Review of Economic Studies 49*, 637-643.

Wilhelm, M.O. (1996), Bequest Behavior and the Effect of Heirs' Earnings: Testing the Altruistic Model of Bequest, *American Economic Review 86 (4)*, 874-892.