The Scandinavian
Journal of Economics

Scand. J. of Economics 114(4), 1368–1392, 2012
DOI: 10.1111/j.1467-9442.2012.01725.x

Optimal Taxation of Bequests in a Model with Initial Wealth*

Johann K. Brunner

University of Linz, A-4040 Linz, Austria
johann.brunner@jku.at

Susanne Pech

University of Linz, A-4040 Linz, Austria
susanne.pech@jku.at

Abstract

We formulate an optimal-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. The introduction of the inheritance tax either increases intertemporal social welfare or has an ambiguous effect, depending on whether the external effect related to altruism is accounted for in the social objective.

Keywords: Expenditure tax; inheritance tax; intergenerational wealth transfers

JEL classification: *H*21; *H*24

I. Introduction

The taxation of estates or inheritances is still an issue that is intensively discussed in tax policy. There are strong movements in many countries to abolish the tax on bequests (in fact, it has recently been repealed in Sweden and Austria), because it is considered to be both immoral (called a "death tax") and adverse to savings.[1] However, proponents primarily stress that it has a redistributive effect – they regard the tax as an instrument for increasing "equality of opportunity". The existence of such controversial views might be the consequence of deep-seated ideological differences, but it might also be attributed to the missing evidence offered by economists as to the effects of a bequest tax.

*We are grateful to two anonymous referees for many helpful comments. This project was supported by the Austrian Science Fund (NRN Austrian Center for Labor Economics and the Analysis of the Welfare State).

[1] However, such a tax still exists in most European countries. In the US, the estate tax expired in 2010, but it was reintroduced for the years 2011 and 2012 (with an exemption at $5 million and a maximum marginal tax rate at 35 percent).

In this paper, we aim to provide new evidence about 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 that 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-theoretical analysis of estate taxation, which is based on the optimal-taxation model in the tradition of Mirrlees (1971) and which only concentrates on the differences in earning abilities. In such a restricted framework, redistribution is best performed via an income tax alone, and there is no need for any indirect tax (Atkinson and Stiglitz, 1976). Consequently, there is no role for a tax on bequests either, because leaving bequests can be seen as a specific way of spending income, similar to the consumption of goods.[2] Even a subsidy on bequests might be considered desirable, if we take into account the view that bequests (or, more generally, gifts) create a twofold utility – for the donor as well as for the donee (i.e., if a positive externality is attributed to leaving bequests, which calls for a Pigouvian subsidy).[3]

However, the situation is fundamentally different if we introduce into the model the fact that the individuals of some generation are already endowed with (differing) initial wealth, as a result of bequests left by their parents. Then, individuals differ in two characteristics – earning abilities and initial wealth – and the aim of our paper is to show that this does matter when determining the welfare effect of inheritance taxation.

To our knowledge, in the prior research, there has been no attempt to provide such an analysis. Some authors have discussed 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 have assumed that bequests are unobservable, and they have analyzed to what extent other taxes can be designed as surrogates. In contrast, we model bequests as being taxable (as is labor income), because this is the assumption on which actual tax systems rely.[4]

[2] To be precise, this result follows if preferences are weakly separable between consumption and leisure. Otherwise, the complementarity or substitutability of some consumption good with leisure plays a role (Corlett and Hague, 1953). Saez (2002) has considered heterogeneity in tastes and he has argued, 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 we analyze its consequences.

[3] For example, see Blumkin and Sadka (2003) or Farhi and Werning (2010).

[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.

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 (see Blumkin and Sadka, 2003, among others) and it implies, in particular, a precise rule of how estates are allocated to the members of the subsequent generation; on their deaths, parents leave all their wealth to their own children.

We begin our analysis with a simple model of two generations, each consisting of two individuals. Those in the parent generation are endowed with given initial wealth and they have differing earning abilities. They use their labor income, together with initial wealth, for their own consumption and for leaving bequests to their immediate descendants. The planner determines an optimal non-linear tax on labor income and considers, in addition, the introduction of a proportional tax on bequests or inheritances.[5] We show that a (specific) tax on bequests left by the parent generation has an ambiguous effect on intertemporal social welfare. On the one hand, there is the above-mentioned effect that a subsidy on bequests increases intertemporal social welfare if the external effect is observed (i.e., if there is "double counting" of the welfare effect of bequests);[6] otherwise this effect disappears. On the other hand, there is also an argument for a tax for redistributive reasons, because taxing the bequests of the parents means indirectly taxing the 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 increases welfare, provided that high-ability individuals have larger initial wealth than low-ability individuals.

Moreover, we show that, for obvious reasons, a direct tax on the given inheritances of the parent generation is definitely desirable, because it allows for more redistribution than the optimal labor income tax alone, if the high-ability individual also has more inherited wealth. It has no adverse effects on the 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 (i.e., on their consumption as well as on the bequests they leave to their descendants). Both taxes are equivalent, although the tax on initial wealth is a lump-sum tax while the expenditure tax is not.

[5] For our main point (i.e., the consequences of unequal initial wealth), it is inessential whether we work with a proportional (as do Blumkin and Sadka 2003) or a non-linear tax (Farhi and Werning 2010).

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

In a next step, we account explicitly for the fact that the parent generation inherited its initial wealth from the previous generation. That is, we introduce an earlier third generation of altruistic 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). The social planner determines optimal non-linear labor income taxes for these two generations, knowing that the abilities remain the same within a dynasty[7] and being able to credibly commit to not changing taxes in the following periods.

We find that the equivalence of a tax on the inheritances of the parent generation with a tax on all expenditures of this generation still hold, even if their effect on the previous generation is taken into account. The overall effect on the intertemporal social welfare of these taxes is ambiguous – there is a welfare-increasing redistributive effect if the initial wealth of the grandparents increases with earning abilities. However, these taxes also induce a reduction of bequests, which diminishes their positive external effect, and thus welfare if the externality is observed by the planner.

As mentioned above, it is the positive correlation between inherited wealth and abilities that is crucial for the result that inheritance taxation extends the scope of redistribution. There is some evidence that human capital income is positively correlated with inheritances (Masson and Pestieau, 1997).[8] However, in many studies, a high intergenerational correlation of wealth (for an overview, see Davies and Shorrocks, 2000) and, to a lower extent, of labor income (for an overview, see Solon, 1999, 2002) has been found. These findings are consistent with a transmission process in which higher-ability individuals bequeath both more wealth and a higher earning ability to their descendants.

The plan of the paper is as follows. In Section II, we consider bequest taxation in the standard optimal-taxation model with two generations, extended by the existence of initial wealth. In Section III, we incorporate the consequences on a previous generation, and in Section IV we provide concluding remarks.

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

[8] In particular, there is evidence for a positive correlation between labor earnings and wealth (Díaz-Giménez *et al.*, 2011), a substantial part of which is known to result from inheritances (e.g., Gale and Scholz, 1994; Davies and Shorrocks, 2000).

II. Basic Model with Two Generations

We begin with the simplest version of a model of dynasties, similar to that in Blumkin and Sadka (2003) or in section II of Farhi and Werning (2010). There are only two dynasties, and each comprises two generations (a parent and a child). Each generation, consisting of two individuals, is assumed to live for one period. The children (generation $t+1$) do not work at all; they live on inheritances.[9] However, the parents (generation t) do work; they differ in earning abilities $\omega_L < \omega_H$. By working l_{it} units of time, each parent $i = L, H$ earns gross income $z_{it} = \omega_i l_{it}$ and net income x_{it}, $i = L, H$, which are spent on own consumption c_{it} and bequests b_{it}. Each parent has a single child to whom they leave all their bequests. Thus, b_{it} is equal to child consumption c_{it+1}.

As mentioned in Section I, an important feature of our model is that we allow for the fact that parents are endowed with (differing) initial wealth $e_{it} \geq 0$. Thus, a second distinguishing characteristic is introduced, in addition to the earning abilities. The initial wealth of each individual i of generation t is a result of previous intergenerational transfers, but it is taken to be exogenous in this section. In Section III, the bequest decision of the foregoing generation is considered in detail.

The identical preferences of the parents are characterized by pure altruism and they can be described by the concave utility function $u(c_{it}, c_{it+1}, l_{it})$, strictly increasing with c_{it} and c_{it+1}, strictly decreasing with l_{it}. Child consumption is assumed to be a normal good, and it enters the utility function just like own consumption. The utility $U(c_{it+1})$ of the child depends only on own consumption c_{it+1}, with $U : \mathbb{R} \to \mathbb{R}$ being strictly concave and increasing. In the next section, when we introduce a third generation, we require additive separability with respect to generations and we write the parent's utility as $\widetilde{U}(c_{it}, l_{it}) + \delta U(c_{it+1})$, where $0 < \delta \leq 1$ is a discount factor, usually interpreted as representing the degree of altruism.[10]

As is usual in Mirrlees-type models, we assume that the government cannot observe abilities. Hence, it imposes – as a second-best instrument – a non-linear tax on the labor income of the working generation t. Moreover, it can use three proportional taxes: a tax τ_{et} on the initial (inherited)

[9] The assumption of non-working children is made for ease of exposition, and it does not affect the results. Moreover, the results of this paper can be shown to hold analogously in a model of dynasties with an arbitrary number of descendant generations.

[10] In an even more specific version, additivity is also assumed between consumption and labor with utility out of consumption being the same for the parent and the child: $u(c_{it}, c_{it+1}, l_{it}) = U(c_{it}) + \delta U(c_{it+1}) - h(l_{it})$, where $h : \mathbb{R} \to \mathbb{R}$, strictly convex and increasing, describes the disutility of labor.

wealth of generation t, a tax τ_{bt} on the bequests left by generation t (i.e., on the inheritances of generation $t+1$), and a tax τ_t on all expenditures of the parents (i.e., a uniform tax rate on their own consumption c_{it} and bequests $b_{it} = c_{it+1}$). An important assumption in this section and in Section III is that the government does not use information on any of the (observable) amounts, on which these indirect taxes are based, to identify individuals according to their earning ability. That is, if there is a fixed and publicly known relation between inheritances (or bequests or expenditures) and abilities (in particular, a positive one, as we suggest later), then the government could, in principle, infer the ability types from the reported amounts e_{it} (or b_{it} or $c_{it} + b_{it}$) and apply a differentiated lump-sum tax as a first-best instrument. Our assumption that the government does not follow such a strategy is in accordance with the actual behavior of tax authorities, which is probably because, in reality, the relation between these characteristics is stochastic and allows no such identification. In an extended version of this paper, we show that the simplifying assumption of a fixed relation between the observable inherited wealth and the unobservable abilities can be dropped and that results analogous to those derived below also hold in a model with a stochastic relation, given that preferences are restricted to being quasi-linear.[11] For the tax on expenditures, it is also be possible to invoke the argument that the total expenditures of an individual are unobservable and that only single purchases of consumption goods can be observed. Note that in any study of indirect taxes in a Mirrlees-type model, it is assumed that individuals cannot be identified by their consumption demand.

For a given net income x_{it}, gross income z_{it}, initial wealth e_{it} and tax rates τ_{bt}, τ_{et} and τ_t, let the indirect utility of a parent for a general utility function u be defined by

$$v_t^i(x_{it}, z_{it}, e_{it}, \tau_{bt}, \tau_{et}, \tau_t) \equiv \max \left\{ u\left(c_{it}, c_{it+1}, \frac{z_{it}}{\omega_i}\right) \middle| (1+\tau_t)[c_{it} + (1+\tau_{bt})c_{it+1}] \leq x_{it} + (1-\tau_{et})e_{it} \right\}. \quad (1)$$

Here, we assume that either τ_{bt} or τ_t exists. First, we take the taxes on bequests, initial wealth, and expenditures to be fixed at $\tau_{bt} = \tau_{et} = \tau_t = 0$, and we consider a benevolent government that imposes an optimal nonlinear income tax in order to maximize the welfare of the two generations. This is equivalent to determining two bundles (x_{Lt}, z_{Lt}) and (x_{Ht}, z_{Ht}), subject to a self-selection constraint and a resource constraint. With a social discount factor $\beta \geq 0$ and the required government resources g_t, the

[11] The extended version of this paper is available from the authors on request.

problem can be written as

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

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

$$\begin{aligned} x_{Lt} + x_{Ht} \leq z_{Lt} + z_{Ht} + \tau_{et} \sum_{i=L,H} e_{it} + \tau_{bt} \sum_{i=L,H} c_{it+1} \\ + \tau_t \sum_{i=L,H} (c_{it} + c_{it+1}) - g_t. \end{aligned} \tag{4}$$

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.[12] Note that, for $\beta = 0$, the social objective (2) is equal to the welfare of the parent generation (which includes the welfare of the descendants). For $\beta > 0$, the welfare of the descendants is also included separately, which implies double counting, as mentioned in Section I.[13]

As a first step, we ask how the introduction of a tax on bequests affects social welfare. Let $S_1(\tau_{bt}, \tau_{et}, \tau_t)$ be the optimal value of the foregoing problem (2)–(4) and let μ be the Lagrange multiplier of equation (3). We find the following.

Proposition 1. *The welfare effect of introducing a tax τ_{bt} on bequests left by generation t is*

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

[12] We further assume that agent monotonicity (Seade, 1982) is fulfilled for general preferences $u(c_{it}, c_{it+1}, l_{it})$, that is, $-(\partial v_t^L/\partial z_{Lt})/(\partial v_t^L/\partial x_{Lt}) > -(\partial v_t^H/\partial z_{Ht})/(\partial v_t^H/\partial x_{Ht})$ at any admissible (x, z).

[13] Although double counting might appear as the appropriate procedure in a utilitarian framework, there is no consensus among economists as to whether the social objective should allow for it or not (e.g., Cremer *et al.*, 2003, p. 2488). A major argument against double counting is the following. By implying a subsidy for bequests (see the remark after Proposition 1), which is financed by a tax on the parent generation, double counting effectively suggests pure redistribution in favor of the descendants, at the expense of the parents. It induces transfers (bequests) above the level desired by the parents themselves, but these transfers do not create a net gain (no potential Pareto improvement, as stressed by Milgrom, 1993), in contrast to the increased (subsidized) consumption of a good that causes a positive external effect.

In general, the sign of this effect is ambiguous. The first term is negative if $\beta > 0$. The second term can have either sign. With weak separability of preferences between consumption and labor, it is zero if individuals have identical initial wealth, but it is positive if the high-ability individual is endowed with more inherited wealth than the low-ability individual.

Proof: See the Appendix.

In the equation of Proposition 1, $U'_{it+1} \equiv dU/dc_{it+1}$ for $i = L, H$. The superscript "com" denotes compensated demand and $[L]$ refers to "mimicking" (i.e., a situation where the high-wage individual opts for the bundle designed for the low-wage individual).

Thus, for the special case where the individuals have zero (and hence also equal) initial wealth, as assumed in the models of Blumkin and Sadka (2003) and Farhi and Werning (2010), we find the familiar result that, with the mild assumption of weak separability of preferences between consumption and labor, a subsidy on bequests increases welfare. Weak separability, together with $e_{Lt} = e_{Ht}$, implies that the second term in the equation of Proposition 1 (i.e., the effect on the self-selection constraint) is zero, because the difference between bequests left by type H, when mimicking type L, and the bequests of type L, is zero: $c_{Ht+1}[L] = c_{Lt+1}$.[14] However, 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 optimal non-linear income tax is a sufficient instrument for redistribution within a generation, and that there is no role for a tax on a specific good. The particular issue in the present model is that the good bequests (i.e., 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 internalize a positive external effect is desirable.[15] Indeed, if β is zero (no double counting), we have the Atkinson–Stiglitz outcome.

However, the essential insight of Proposition 1 is that this result no longer holds if we allow explicitly for differing initial wealth. Then, even for

[14] Mimicking by type H means that type H chooses the same bundle of net and gross incomes as type L. Because initial wealth is also the same, the only difference between the two types is in labor supply (i.e., type H can earn the same gross income with less working time). However, because of weak separability, this does not influence the decision of how to spend the net income.

[15] In a model with an optimal non-linear tax on bequests, Farhi and Werning (2010) show that this tax is progressive (i.e., the marginal subsidy is lower for high-ability individuals).

weak separability, we have $c_{Ht+1}[L] > c_{Lt+1}$, if $e_{Ht} > e_{Lt}$. That is, because bequests are assumed to be a normal good, high-ability individuals – even when mimicking – will leave more bequests to their descendants if they are endowed with more initial wealth. In this case, the introduction of a tax on bequests gives slack to the self-selection constraint, and thus more redistribution via the income tax becomes possible. To find some intuition for this result, we consider the effect of an introduction of some small $\Delta\tau_{bt}$, which raises tax revenues $\Delta\tau_{bt}(c_{Lt+1} + c_{Ht+1})$. Compensating the individuals by an increase in net income $\Delta x_{it} = \Delta\tau_{bt}c_{it+1}$ would leave the welfare of both individuals approximately unchanged (note that the deadweight loss is of second order). However, because $\Delta x_{Lt} < \Delta x_{Ht}$, if $e_{Lt} < e_{Ht}$, this procedure makes mimicking less attractive, and thus it allows further redistribution of net income by increasing Δx_{Lt} and decreasing Δx_{Ht}, thereby increasing social welfare. This positive welfare effect counteracts the consequence of double counting, which calls for a subsidy; the overall welfare effect is ambiguous and depends on the magnitude of both effects.

Next, we focus on the two alternative tax instruments τ_{et} and τ_t and we determine their effects on social welfare.

Proposition 2. *The welfare effect of introducing a tax τ_{et} on initial wealth or a tax τ_t on the expenditures of the parent generation t reads as*

$$\left.\frac{\partial S_1}{\partial \tau_{et}}\right|_{\tau_{bt}=\tau_{et}=\tau_t=0} = \left.\frac{\partial S_1}{\partial \tau_t}\right|_{\tau_{bt}=\tau_{et}=\tau_t=0} = \mu \frac{\partial v_t^H[L]}{\partial x_{Lt}}(e_{Ht} - e_{Lt}).$$

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

Proof: See the Appendix.

In contrast to Proposition 1, 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 non-linear income tax. It is interesting to observe that the effects of τ_{et} and τ_t are identical, although the former is clearly a lump-sum tax while the latter is distorting, because expenditures are endogenous. The positive welfare effect of either tax comes from a relaxation of the self-selection constraint, and the intuition is similar to that presented above, by considering the additional redistribution of income made possible by the introduction of a small $\Delta\tau_{et}$ instead of $\Delta\tau_{bt}$. Moreover, note that in the equation of Proposition 2, the social discount factor β does not appear (i.e., the positive effect occurs whether or not there is double counting).

III. Considering the Previous Generation

After having demonstrated that the (differing) initial wealth of a generation t provides a rationale for bequest taxation, we now introduce into our model the fact that this initial wealth occurs as a result of bequests left by the previous generation $t-1$. In other words, we no longer take initial wealth e_{Lt}, e_{Ht} to be exogenous, but we incorporate the decisions of generation $t-1$ and we analyze the welfare consequences of the taxes τ_{et}, τ_t, and τ_{bt}, given that generation $t-1$ is aware of these taxes.

Again, we assume that pure altruism motivates the bequest decisions of generation $t-1$. That is, this generation cares for its own activities as well as for the activities of the following generations. As mentioned in Section II, we assume from now on that utility is additively separable with respect to generations. Hence, the utility function of an individual i of generation $t-1$ is $\widetilde{U}(c_{it-1}, l_{it-1}) + \delta\widetilde{U}(c_{it}, l_{it}) + \delta^2 U(c_{it+1})$, 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 also assume that generation $t-1$ consists of two types of individuals, with abilities ω_L and ω_H, respectively. Moreover, each individual knows that the descendant, to whom all bequests are left, has the same earning ability. The members of generation $t-1$ have initial wealth e_{Lt-1}, e_{Ht-1}. Thus, the indirect utility of generation $t-1$ is defined as

$$
\begin{aligned}
&v^i_{t-1}(x_{it-1}, z_{it-1}, x_{it}, z_{it}, e_{it-1}, \tau_{bt}, \tau_{et}, \tau_t) \\
&\quad \equiv \max\left\{ \widetilde{U}\left(c_{it-1}, \frac{z_{it-1}}{\omega_i}\right) + \delta\widetilde{U}\left(c_{it}, \frac{z_{it}}{\omega_i}\right) + \delta^2 U(c_{it+1}) \mid c_{it-1} + e_{it} \right. \\
&\qquad\qquad \left. \le x_{it-1} + e_{it-1}, (1+\tau_t)\left[c_{it} + (1+\tau_{bt})c_{it+1}\right] \le x_{it} + (1-\tau_{et})e_{it} \right\},
\end{aligned} \tag{5}
$$

where not both τ_t and τ_{bt} exist. 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, it is necessary to introduce $\hat{e}_{it}$ as bequests net of the tax τ_{bt-1} and to write the budget constraints as $c_{it-1} + (1+\tau_{bt-1})\hat{e}_{it} \le x_{it-1} + e_{it-1}$, $(1+\tau_t)(c_{it} + c_{it+1}) \le x_{it} + \hat{e}_{it}$. If we set $\tau_{bt-1} = \tau_{et}/(1-\tau_{et})$ in the relation $(1+\tau_{bt-1})\hat{e}_{it} = e_{it}$, we obtain the equivalence of the two budget constraints.) Similarly, of course, τ_{bt} is equivalent to an inheritance tax

τ_{et+1}. The indirect utility of generation t is defined accordingly as[16]

$$\begin{aligned} & v_t^i(x_{it}, z_{it}, e_{it}, \tau_{bt}, \tau_{et}, \tau_t) \\ & \equiv \max\left\{\widetilde{U}\left(c_{it}, \frac{z_{it}}{\omega_i}\right) + \delta U(c_{it+1}) \mid (1+\tau_t)\left[c_{it} + (1+\tau_{bt})c_{it+1}\right]\right. \\ & \qquad \left. \leq x_{it} + (1-\tau_{et})e_{it}\right\}, \end{aligned} \tag{6}$$

while the utility of generation $t+1$ is again given by $U(c_{it+1})$.

The government determines the optimal non-linear income taxes for generations $t-1$ and t, which is equivalent to finding the optimal bundles (x_{it-1}, z_{it-1}), (x_{it}, z_{it}), $i = L, H$, for given tax rates τ_{et}, τ_t, and τ_{bt} (possibly zero). 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, the abilities remain constant over generations and that this is known by the authority. Then, the planner only has to observe the self-selection constraint for the first generation of the dynasties; mimicking cannot occur in later generations, because their abilities are known to be the same as those of their parents (see also Diamond, 2007; Golosov *et al.*, 2007).

However, as is usual in Ramsey-type dynamic problems, we also assume that the government can credibly commit to not changing the taxes that are determined in period $t-1$, in the following period t. Otherwise, because the solution of the planner's problem is not time-consistent, individuals would expect reoptimization in period t, which would change their behavior.

Moreover, we assume that the government can transfer resources over time, when determining the gross- and net-income positions of the individuals in periods $t-1$ and t. Consequently, only one resource constraint has to be observed for the two periods.[17] Let β again denote the social rate for discounting the welfare of future generations. The optimization problem of the planner to determine the non-linear income tax in periods $t-1$ and t for given τ_{et}, τ_t, and τ_{bt} is (with the government resource requirements g_{t-1} and g_t)

$$\max_{(x_{it-1}, z_{it-1})(x_{it}, z_{it}), i=L,H} \sum_{i=L,H} f_i\left[v_{t-1}^i(\cdot) + \beta v_t^i(\cdot) + \beta^2 U(c_{it+1})\right] \tag{7}$$

[16] Note the recursive structure: $v_{t-1}^i(\cdot) = \max\{\widetilde{U}(c_{it-1}, z_{it-1}/\omega_i) + \delta v_t^i(\cdot) \mid c_{it-1} + e_{it} \leq x_{it-1} + e_{it-1}\}$.

[17] It can be shown that nothing changes if separate resource constraints for the two periods are introduced without any public instrument of transferring resources across generations. By adapting bequests e_{it} appropriately, dynasties choose their preferred intertemporal allocation for given taxes τ_{et}, τ_t, or τ_{bt}, independently of whether tax revenues are used by the government to increase net incomes in periods $t-1$ or t.

$$\text{s.t.}\ \ v^H_{t-1}(x_{Ht-1}, z_{Ht-1}, x_{Ht}, z_{Ht}, e_{Ht-1}, \tau_{et}, \tau_t, \tau_{bt}) \geq v^H_{t-1}(x_{Lt-1}, z_{Lt-1}, x_{Lt}, z_{Lt}, e_{Ht-1}, \tau_{et}, \tau_t, \tau_{bt}), \tag{8}$$

$$\sum_{i=L,H}(x_{it-1} + x_{it}) \leq \sum_{i=L,H}(z_{it-1} + z_{it}) + \tau_{et}\sum_{i=L,H} e_{it}(\cdot) + \tau_t \sum_{i=L,H}[c_{it}(\cdot) + c_{it+1}(\cdot)] + \tau_{bt}\sum_{i=L,H} c_{it+1}(\cdot) - g_{t-1} - g_t. \tag{9}$$

Our first result generalizes the close relation between the taxes τ_{et} and τ_t, as illustrated in Proposition 2. The equivalence continues to hold if their effect on the previous generation is taken into account, and it extends to general (non-zero) tax rates. Let $S_2(\tau_{et}, \tau_t, \tau_{bt})$ denote the optimal value of the maximization problem (7)–(9) and let $\tau_{bt} = 0$.

Proposition 3. *For any given $\tau_{et}, \tau_t \geq 0$*

$$\frac{\partial S_2}{\partial \tau_{et}} = \frac{\partial S_2}{\partial \tau_t}\frac{1+\tau_t}{1-\tau_{et}}.$$

Proof: See the Appendix.

Given any pair of non-negative tax rates τ_{et}, τ_t, the welfare effect of a marginal increase of either tax is essentially the same, up to the term $(1+\tau_t)/(1-\tau_{et})$, which arises because the first represents a deduction, while the second makes expenditures more expensive. This equivalence means that the expenditure tax τ_t, although its base comprises the initial wealth as well as the labor income of generation t, effectively falls on initial wealth alone, if it is combined with an optimal tax on labor income. Its potential impact on labor income can be completely offset by a reduction of the optimal labor income tax.

In particular, at $\tau_{et} = \tau_t = 0$, the welfare effect of both taxes is identical. We provide an equation for this in terms of τ_{et}, thereby extending Proposition 2.

Proposition 4. *The welfare effect of introducing a tax τ_{et} on the inheritances of generation t is*

$$\left.\frac{\partial S_2}{\partial \tau_{et}}\right|_{\tau_{et}=\tau_t=\tau_{bt}=0} = \sum_{i=L,H} f_i\left(\beta\frac{\partial v^i_t}{\partial x_{it}} + \beta^2 U'_{it+1}\frac{\partial c^t_{it+1}}{\partial x_{it}}\right)\frac{\partial \widetilde{e}^{\text{com}}_{it}}{\partial \tau_{et}} + \mu\frac{\partial v^H_{t-1}[L]}{\partial x_{Lt-1}}(e_{Ht}[L] - e_{Lt}). \tag{10}$$

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

Proof: See the Appendix.

In the equation of Proposition 4, $\widetilde{e}_{it} = e_{it}(1-\tau_{et})$ denotes the bequests net of the inheritance tax and $\partial \widetilde{e}_{it}^{\text{com}}/\partial \tau_{et}$ is the compensated effect of τ_{et} on net bequests, left by generation $t-1$. Here, $\partial c_{it+1}^{t}/\partial x_{it}$ denotes the effect of x_{it} on the consumption of generation $t+1$, as determined by generation t (with given inheritances e_{it}). Moreover, μ indicates the Lagrange multiplier of the self-selection constraint (8).

We find that the condition that is decisive for the introduction of an inheritance tax in period t is analogous to that of Proposition 1, which refers to a tax τ_{bt} on the bequests of generation t in the model with only two generations. The first term, associated with the own compensated price effect $\partial \widetilde{e}_{it}^{\text{com}}/\partial \tau_{et}$, is negative; the net bequests of generation $t-1$ decrease. Thus, their external effect is reduced, which now affects the welfare of the two subsequent generations t and $t+1$, as is visible from the expression in brackets. However, there is a positive effect on the self-selection constraint as before. Given that, in generation $t-1$, the high-ability type has more initial wealth than the low-ability type, the former, when mimicking, will choose higher bequests than the latter, given the weak separability of preferences. Then, the introduction of an inheritance tax τ_{et} allows more redistribution.[18]

Given that the introduction of an inheritance tax increases social welfare (i.e., equation (10) is positive), we can ask for a characterization of the optimal tax rate. For this, we have to observe that with increasing τ_{et} a negative deadweight-loss effect arises, which is of second order and therefore comes in with $\tau_{et} > 0$. The optimal value τ_{et}^{*} occurs where the positive welfare effect balances the deadweight-loss effect. Formally, setting $\partial S_2/\partial \tau_{et} = 0$, for $\tau_t = \tau_{bt} = 0$, gives us (see the Appendix)

$$\frac{\tau_{et}^{*}}{1-\tau_{et}^{*}} = -\frac{1}{\lambda}\frac{W(\tau_{et}^{*})}{\sum_{i=L,H}(\partial \tilde{e}_{it}^{\text{com}}/\partial \tau_{et})}. \tag{11}$$

[18] Note that the negative compensated effect arises, because it is assumed that the taxes τ_{et} and τ_t, although imposed on the parent generation t, are already known when the grandparent generation $t-1$ makes its decisions. We can think of a scenario such that the taxes τ_{et} and τ_t are imposed at a point in time when the decisions of generation $t-1$ have already been made. In this case, only the positive term occurs, as in Proposition 2.

Here, λ denotes the Lagrange multiplier referring to the resource constraint (9) and $W(\tau_{et}^*)$ is the marginal welfare effect of τ_{et}, as shown on the right-hand side (rhs) of equation (10), but evaluated at the optimal τ_{et}^*. Note that on the rhs of equation (11), the denominator contains the negative own compensated price reactions. The larger (in absolute value) these are, the larger the marginal deadweight loss and the lower the optimal tax rate. Moreover, in the optimum, the welfare effect W must still be positive and $\tau_{et}^* < 1$ (clearly, with τ_{et} approaching 1, the effective price of bequests goes to infinity).

Finally, we focus on a tax τ_{bt} on bequests (instead of a tax τ_{et} on inheritances) of generation t, assuming that generation $t-1$ is already aware of this tax. We find the following characterization of its welfare effect.

Proposition 5. *The welfare effect of introducing a tax τ_{bt} on bequests left by generation t is*

$$\left.\frac{\partial S_2}{\partial \tau_{bt}}\right|_{\tau_{et}=\tau_t=\tau_{bt}=0} = \sum_{i=L,H} f_i \left(\beta \frac{\partial v_t^i}{\partial x_{it}} \frac{\partial \tilde{e}_{it}^{\text{com}}}{\partial \tau_{bt}} + \beta^2 U'_{it+1} \frac{\partial c_{it+1}^{\text{com}}}{\partial \tau_{bt}}\right) + \mu \frac{\partial v_{t-1}^H[L]}{\partial x_{Lt}} \left(c_{Ht+1}[L] - c_{Lt+1}\right).$$

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

Proof: See the Appendix.

The compensated effect of τ_{bt} on the (net) inheritances $\tilde{e}_{it}$ and $\tilde{e}_{it+1} = c_{it+1}$ determines how the distortion of the bequest decision affects the welfare of generations t and $t+1$. It is negative for both, but clearly more accentuated for the taxed good c_{it+1} alone than for the sum $c_{it} + c_{it+1}$ (note that $\partial \tilde{e}_{it}^{\text{com}}/\partial \tau_{bt} = \partial c_{it}^{\text{com}}/\partial \tau_{bt} + \partial c_{it+1}^{\text{com}}/\partial \tau_{bt}$), because some substitution takes place from c_{it+1} to c_{it}. However, τ_{bt} also causes the already familiar positive redistributive effect related to the difference in demand for c_{it+1} between the mimicking and the mimicked individual.[19] Combining Propositions 4 and 5, we find that the marginal welfare effect of an inheritance tax imposed simultaneously on the two generations t and $t+1$,

[19] If the introduction of τ_{bt} is announced too late for generation $t-1$ (but not for generation t) to adapt its behavior, then we return to the problem studied in Proposition 1.

introduced at a point in time such that generation $t - 1$ can adapt its behavior, is also ambiguous, if double counting ($\beta > 0$) is included in the objective.

The equations of Propositions 4 and 5 are related to the work of Cremer *et al.* (2003), who have studied the welfare effect of a capital income tax in a dynamic model, where individuals differ in initial wealth, as a consequence of bequests left by the previous generation because of a joy-of-giving motive. Bequests are untaxed in this model because of unobservability, and it is shown that the capital income tax can instead be used as an instrument for redistribution, given that initial wealth is positively correlated with abilities. The optimal tax rate arises when this positive redistributive effect is balanced by the negative welfare effect, because of the distortion of life-cycle saving. In a related work, Boadway *et al.* (2000) established the redistributive role of a tax on capital income, when the differences in initial wealth arise from accidental (unobservable) bequests.

In contrast, in this paper, our aim is to study how a tax on bequests affects the intergenerational transmission of wealth itself, and to work out the equivalence of a tax on the bequests of some generation with a tax on all expenditures of the subsequent generation. Assuming a model where generations are linked via altruistic preferences and where, within the first generation, there are differences in initial wealth, we find equations for the welfare effect of these taxes, depending on the point in time when they are imposed.[20]

As mentioned in Section I, it is an open question to what extent wealth transfers are observable, and thus whether the taxes on them can actually be enforced. Perhaps we can take the very existence of these taxes as an indication of a widely held view that enforceability is given to a sufficient degree. In a related work (Brunner *et al.*, 2010), we have analyzed the role of taxes on (inherited) wealth and expenditures in an optimal tax system if only partial observability is assumed (i.e., if we allow for tax evasion).

IV. Conclusion

In this paper, we have analyzed the welfare effects of estate or inheritance taxation in a model that accounts for the fact that initial wealth constitutes a second distinguishing characteristic of individuals, in addition to earning abilities. This extension of the optimal-taxation model is essential if an economic appraisal of estate taxation is to be relevant for the current debate on such a tax. We are now in a world where differences in initial

[20] Note that in Cremer *et al.* (2003), a 100 percent tax on inheritances (given that they were observable) would be optimal, which causes no adverse incentive effect because of their specific modeling of the wealth-transfer technology.

wealth already exist, as a result of transfers over generations in the past. These should be recognized, and we have demonstrated that they matter for our understanding of the consequences of a tax on intergenerational wealth transfers.

We have been able to show that, in the extended model, an inheritance tax has a redistributive effect, which increases intertemporal social welfare if initial wealth and earning abilities are positively correlated. This might explain why the inheritance tax is frequently regarded as an enhancing equality of opportunity. The welfare-increasing effect is unambiguous if inheritances are considered to be exogenous, which might be interpreted as meaning that the tax is introduced at a point in time when the preceding generation can no longer react to the tax. Otherwise, a second, welfare-decreasing effect arises, because the preceding generation adapts to the tax in a way that ignores the positive external effect of bequests on later generations.

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, because the 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, the introduction of all taxes considered in the paper has only a positive, redistributive effect, irrespective of the reaction of the bequeathing generation.

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

In a related paper, Brunner and Pech (2012) have studied estate or inheritance taxation when bequests are motivated by joy-of-giving instead of altruism. That is, the (net) bequests instead of the consumption of future generations enter their utility function. Empirically, there is no clear-cut evidence as to which of the two motives dominates actual decisions; it is probable that a mixture of them (in combination with accidental bequests) applies (e.g., Laitner and Juster, 1996; Wilhelm, 1996; Arrondel and Laferrère, 2001; Laitner and Ohlsson, 2001). The main consequence of the joy-of-giving motive is that the bequeathing individuals only care about the taxes directly related to their bequests, but they do not care about future taxes that are imposed on the expenditures of the descendant generation. This causes a difference between inheritance and expenditure taxation and it makes the latter a preferable instrument, in contrast to the equivalence result found in the present study.

Appendix: Proofs of the Propositions

Proof of Proposition 1

Throughout the proof, let $\tau_{bt} = \tau_{et} = \tau_t = 0$. From the Lagrangian to the maximization problem (2)–(4), we derive the first-order conditions with respect to x_{Lt} and x_{Ht}:

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

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

Here, $U'_{it+1} \equiv dU/dc_{it+1}, i = L, H,$ and $v_t^H[L] \equiv v_t^H(x_{Lt}, z_{Lt}, \cdot)$, and μ and λ denote the Lagrange multipliers corresponding to equations (3) and (4), respectively. Using the envelope theorem, for the optimal value function $S_1(\tau_{bt}, \tau_{et}, \tau_t)$, we obtain

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

We have $\partial v_t^i/\partial \tau_{bt} = -c_{it+1} \partial v_t^i/\partial x_{it}$ and $\partial v_t^H[L]/\partial \tau_{bt} = -c_{Ht+1}[L]$ $\partial v_t^H[L]/\partial x_{Lt}$ from Roy's identity. Using these terms and equations (A1) and (A2), multiplied by c_{Lt+1} and c_{Ht+1}, respectively, in equation (A3), we arrive at

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

Finally, the application of the Slutsky equation $\partial c_{it+1}^{\text{com}}/\partial \tau_{bt} = \partial c_{it+1}/\partial \tau_{bt} + c_{it+1} \partial c_{it+1}/\partial x_{it}$ gives us the equation of Proposition 1. Here, $\partial c_{it+1}^{\text{com}}/\partial \tau_{bt} < 0$ holds, because the compensated own-price effect is always negative. Given the weak separability of preferences between consumption and labor, we have $c_{Ht+1}[L] = c_{Lt+1}$, if $e_{Ht} = e_{Lt}$, which implies that the second term on the rhs of equation (A4) is zero; if $e_{Ht} > e_{Lt}$, then $c_{Ht+1}[L] > c_{Lt+1}$ (remember that c_{it+1} is assumed to be a normal good).

Proof of Proposition 2

Throughout the proof, again let $\tau_{bt} = \tau_{et} = \tau_t = 0$. The derivatives of the optimal value $S_1(\tau_{bt}, \tau_{et}, \tau_t)$ with respect to τ_{et} and τ_t are

$$\begin{aligned}\frac{\partial S_1}{\partial \tau_{et}} = &\sum_{i=L,H} f_i \left(\frac{\partial v_t^i}{\partial \tau_{et}} + \beta U'_{it+1} \frac{\partial c_{it+1}}{\partial \tau_{et}} \right) \\ &+\mu \left(\frac{\partial v_t^H}{\partial \tau_{et}} - \frac{\partial v_t^H[L]}{\partial \tau_{et}} \right) + \lambda \sum_{i=L,H} e_{it},\end{aligned} \tag{A5}$$

$$\begin{aligned}\frac{\partial S_1}{\partial \tau_t} = &\sum_{i=L,H} f_i \left(\frac{\partial v_t^i}{\partial \tau_t} + \beta U'_{it+1} \frac{\partial c_{it+1}}{\partial \tau_t} \right) \\ &+\mu \left(\frac{\partial v_t^H}{\partial \tau_t} - \frac{\partial v_t^H[L]}{\partial \tau_t} \right) + \lambda \sum_{i=L,H} (c_{it} + c_{it+1}).\end{aligned} \tag{A6}$$

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

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 $\partial m_{it}/\partial \tau_t = -(x_{it} + e_{it})/(1 + \tau_t)^2 = -(c_{it} + c_{it+1})/(1 + \tau_t)$ 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 expression and $\partial v_t^i/\partial \tau_t = -(c_{it} + c_{it+1}) \partial v_t^i/\partial x_{it}$, $\partial v_t^H[L]/\partial \tau_t = -(c_{Ht}[L] + c_{Ht+1}[L]) \partial v_t^H[L]/\partial x_{Lt}$, as well as equations (A1) and (A2), multiplied by $(c_{Lt} + c_{Lt+1})$ and $(c_{Ht} + c_{Ht+1})$, respectively, in equation (A6) we arrive at

$$\frac{\partial S_1}{\partial \tau_t} = \mu \frac{\partial v_t^H[L]}{\partial x_{Lt}} \left[(c_{Ht}[L] + c_{Ht+1}[L]) - (c_{Lt} + c_{Lt+1}) \right]. \tag{A7}$$

If we use the budget equation of the mimicking and mimicked individuals (i.e., $c_{Ht}[L] + c_{Ht+1}[L] = x_{Lt} + e_{Ht}$ and $c_{Lt} + c_{Lt+1} = x_{Lt} + e_{Lt}$) in equation (A7), we obtain the equation of Proposition 2.

Proof of Proposition 3

Throughout the proof, let $\tau_{bt} = 0$. From the Lagrangian to the maximization problem (7)–(9), we derive the first-order conditions with respect to x_{Lt}

and x_{Ht}:

$$f_L \left\{ \frac{\partial v_{t-1}^L}{\partial x_{Lt}} + \beta \frac{\partial v_t^L}{\partial x_{Lt}} \left[1 + (1-\tau_{et}) \frac{\partial e_{Lt}}{\partial x_{Lt}} \right] + \beta^2 U'_{Lt+1} \frac{\partial c_{Lt+1}}{\partial x_{Lt}} \right\} - \mu \frac{\partial v_{t-1}^H [L]}{\partial x_{Lt}} + \lambda \left[-1 + \tau_{et} \frac{\partial e_{Lt}}{\partial x_{Lt}} + \tau_t \left(\frac{\partial c_{Lt}}{\partial x_{Lt}} + \frac{\partial c_{Lt+1}}{\partial x_{Lt}} \right) \right] = 0; \tag{A8}$$

$$f_H \left\{ \frac{\partial v_{t-1}^H}{\partial x_{Ht}} + \beta \frac{\partial v_t^H}{\partial x_{Ht}} \left[1 + (1-\tau_{et}) \frac{\partial e_{Ht}}{\partial x_{Ht}} \right] + \beta^2 U'_{Ht+1} \frac{\partial c_{Ht+1}}{\partial x_{Ht}} \right\} + \mu \frac{\partial v_{t-1}^H}{\partial x_{Ht}} + \lambda \left[-1 + \tau_{et} \frac{\partial e_{Ht}}{\partial x_{Ht}} + \tau_t \left(\frac{\partial c_{Ht}}{\partial x_{Ht}} + \frac{\partial c_{Ht+1}}{\partial x_{Ht}} \right) \right] = 0. \tag{A9}$$

Here, μ and λ denote the Lagrange multipliers corresponding to equations (8) and (9), respectively. Note that an increase in x_{it} influences the welfare position v_t^i of an individual i of generation t directly (for given e_{it}), but also indirectly, because generation $t-1$ adapts bequests e_{it}. The indirect effect is seen in equations (A8) and (A9) as $(\partial v_t^i / \partial e_{it}) \partial e_{it} / \partial x_{it} = (1-\tau_{et})(\partial v_t^i / \partial x_{it}) \partial e_{it} / \partial x_{it}$.

The derivatives of the optimal value $S_2(\tau_{et}, \tau_t, \tau_{bt})$ with respect to τ_{et} and τ_t are

$$\frac{\partial S_2}{\partial \tau_{et}} = \sum_{i=L,H} f_i \left(\frac{\partial v_{t-1}^i}{\partial \tau_{et}} + \beta \frac{\partial v_t^i}{\partial \tau_{et}} + \beta^2 U'_{it+1} \frac{\partial c_{it+1}}{\partial \tau_{et}} \right) + \mu \left(\frac{\partial v_{t-1}^H}{\partial \tau_{et}} - \frac{\partial v_{t-1}^H [L]}{\partial \tau_{et}} \right) + \lambda \sum_{i=L,H} \left[e_{it} + \tau_{et} \frac{\partial e_{it}}{\partial \tau_{et}} + \tau_t \left(\frac{\partial c_{it}}{\partial \tau_{et}} + \frac{\partial c_{it+1}}{\partial \tau_{et}} \right) \right], \tag{A10}$$

$$\frac{\partial S_2}{\partial \tau_t} = \sum_{i=L,H} f_i \left(\frac{\partial v_{t-1}^i}{\partial \tau_t} + \beta \frac{\partial v_t^i}{\partial \tau_t} + \beta^2 U'_{it+1} \frac{\partial c_{it+1}}{\partial \tau_t} \right) + \mu \left(\frac{\partial v_{t-1}^H}{\partial \tau_t} - \frac{\partial v_{t-1}^H [L]}{\partial \tau_t} \right) + \lambda \sum_{i=L,H} \left[\tau_{et} \frac{\partial e_{it}}{\partial \tau_t} + c_{it} + c_{it+1} + \tau_t \left(\frac{\partial c_{it}}{\partial \tau_t} + \frac{\partial c_{it+1}}{\partial \tau_t} \right) \right]. \tag{A11}$$

Although not written explicitly, the taxes τ_{et} and τ_t, respectively, influence the welfare of generation t in two ways – directly and indirectly – as described above for x_{it}.[21] Thus, $\partial v_t^i/\partial\tau_{et} = [-e_{it} + (1-\tau_{et})\partial e_{it}/\partial\tau_{et}]\,\partial v_t^i/\partial x_{it}$. Moreover, we have $\partial v_{t-1}^i/\partial\tau_{et} = -e_{it}\partial v_{t-1}^i/\partial x_{it}$ and $\partial v_{t-1}^H[L]/\partial\tau_{et} = -e_{Ht}[L]\partial v_{t-1}^H[L]/\partial x_{Lt}$. Using these terms, together with the first-order conditions (A8) and (A9) multiplied by e_{Lt} and e_{Ht}, respectively, equation (A10) can be transformed to

$$\begin{aligned}\frac{\partial S_2}{\partial\tau_{et}} &= \sum_{i=L,H} f_i\left[\beta\frac{\partial v_t^i}{\partial x_{it}}(1-\tau_{et})\Delta_{\tau_{et}}^{e_{it}} + \beta^2 U'_{it+1}\Delta_{\tau_{et}}^{c_{it+1}}\right]\\ &\quad+\mu\frac{\partial v_{t-1}^H[L]}{\partial x_{Lt}}(e_{Ht}[L]-e_{Lt}) + \lambda\sum_{i=L,H}\left[\tau_{et}\Delta_{\tau_{et}}^{e_{it}} + \tau_t\left(\Delta_{\tau_{et}}^{c_{it}} + \Delta_{\tau_{et}}^{c_{it+1}}\right)\right].\end{aligned} \tag{A12}$$

Here, we use the abbreviation $\Delta_{\tau_{et}}^{Z_i} \equiv \partial Z_i/\partial\tau_{et} + e_{it}\partial Z_i/\partial x_{it}$, $Z_i = e_{it}, c_{it}, c_{it+1}$.

Furthermore, we have $\partial v_{t-1}^i/\partial\tau_t = -(c_{it}+c_{it+1})\partial v_{t-1}^i/\partial x_{it}$, $\partial v_{t-1}^H[L]/\partial\tau_t = -(c_{Ht}[L] + c_{Ht-1}[L])\partial v_{t-1}^H[L]/\partial x_{Lt}$, and, because of the direct and indirect effects, $v_t^i/\partial\tau_t = [-(c_{it}+c_{it+1}) + (1-\tau_{et})\partial e_{it}/\partial\tau_t]\partial v_t^i/\partial x_{it}$. Using these expressions, together with the first-order conditions (A8) and (A9), multiplied by $(c_{Lt}+c_{Lt+1})$ and $(c_{Ht}+c_{Ht+1})$, respectively, in equation (A11), we obtain

$$\begin{aligned}\frac{\partial S_2}{\partial\tau_t} &= \sum_{i=L,H} f_i\left[\beta\frac{\partial v_t^i}{\partial x_{it}}(1-\tau_{et})\Delta_{\tau_t}^{e_{it}} + \beta^2 U'_{it+1}\Delta_{\tau_t}^{c_{it+1}}\right]\\ &\quad+\mu\frac{\partial v_{t-1}^H[L]}{\partial x_{Lt}}\left[c_{Ht}[L] + c_{Ht+1}[L] - (c_{Lt}+c_{Lt+1})\right]\\ &\quad+\lambda\sum_{i=L,H}\left[\tau_{et}\Delta_{\tau_t}^{e_{it}} + \tau_t\left(\Delta_{\tau_t}^{c_{it}} + \Delta_{\tau_t}^{c_{it+1}}\right)\right].\end{aligned} \tag{A13}$$

Here, we use the definition $\Delta_{\tau_t}^{Z_i} \equiv \partial Z_i/\partial\tau_t + (c_{it}+c_{it+1})\partial Z_i/\partial x_{it}$, $Z_i = e_{it}, c_{it}, c_{it+1}$.

Next, we show that $\Delta_{\tau_{et}}^{Z_i} = \Delta_{\tau_t}^{Z_i}(1+\tau_t)/(1-\tau_{et})$ for $Z_i = e_{it}, c_{it}, c_{it+1}$. We eliminate e_{it} from the two budget constraints in equation (5) to obtain the combined budget equation $c_{it-1} + (c_{it}+c_{it+1})(1+\tau_t)/(1-\tau_{et}) = x_{it-1} + e_{it-1} + x_{it}/(1-\tau_{et})$. Let $p_t \equiv (1+\tau_t)/(1-\tau_{et})$ and $B_{it-1} \equiv x_{it-1} + e_{it-1} + x_{it}/(1-\tau_{et})$, then

$$\frac{\partial c_{is}}{\partial\tau_t} = \frac{\partial c_{is}}{\partial p_t}\frac{1}{1-\tau_{et}}, \quad s = t-1, t, t+1,$$

$$\frac{\partial c_{is}}{\partial\tau_{et}} = \frac{\partial c_{is}}{\partial p_t}\frac{1+\tau_t}{(1-\tau_{et})^2} + \frac{\partial c_{is}}{\partial B_{it-1}}\frac{x_{it}}{(1-\tau_{et})^2}, \quad s = t-1, t, t+1.$$

[21] Note that both effects are also behind $\partial c_{is}/\partial\tau_{et}$ and $\partial c_{is}/\partial\tau_t$, $s = t, t+1$; see the beginning of *Proof of Proposition 4*.

Together with $\partial c_{is}/\partial B_{it-1} = (1-\tau_{et})\partial c_{is}/\partial x_{it}$ (and the definition of p_t), this gives us $\partial c_{is}/\partial \tau_{et} = (\partial c_{is}/\partial \tau_t)p_t + (\partial c_{is}/\partial x_{it})x_{it}/(1-\tau_{et})$. Using this, together with the budget equation $e_{it} = p_t(c_{it} + c_{it+1}) - x_{it}/(1-\tau_{et})$ for period t, in $\Delta^{c_{is}}_{\tau_{et}} = \partial c_{is}/\partial \tau_{et} + e_{it}\partial c_{is}/\partial x_{it}$, we obtain $\Delta^{c_{is}}_{\tau_{et}} = p_t \Delta^{c_{is}}_{\tau_t}$. Moreover, observing from the budget equation $e_{it} = x_{it-1} + e_{it-1} - c_{it-1}$ for period $t-1$ that $\partial e_{it}/\partial \tau_t = -\partial c_{it-1}/\partial \tau_t$, $\partial e_{it}/\partial \tau_{et} = -\partial c_{it-1}/\partial \tau_{et}$, and $\partial e_{it}/\partial x_{it} = -\partial c_{it-1}/\partial x_{it}$, the same reasoning gives us $\partial e_{it}/\partial \tau_{et} = p_t \partial e_{it}/\partial \tau_t + (\partial e_{it}/\partial x_{it})x_{it}/(1-\tau_{et})$, and hence $\Delta^{e_{it}}_{\tau_{et}} = p_t \Delta^{e_{it}}_{\tau_t}$.

Finally, we conclude from the period-t budget equations of individual L and individual H, when mimicking, that $e_{Ht}[L] - e_{Lt} = p\,[c_{Ht}[L] + c_{Ht+1}[L] - (c_{Lt} + c_{Lt+1})]$. Substituting this expression, together with $\Delta^{Z_i}_{\tau_{et}} = p_t \Delta^{Z_i}_{\tau_t}$ for $Z_i = e_{it}, c_{it}, c_{it+1}$, into equation (A12), by comparing it with equation (A13) we immediately find that $\partial S_2/\partial \tau_{et} = p_t \partial S_2/\partial \tau_t$.

Proof of Proposition 4

Let $c^t_{is}(\cdot)$, $s = t, t+1$, denote (own and 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_{is} is equal to consumption c_{is} decided by generation $t-1$, because of the recursive structure of utility (see footnote 16), that is, $c_{is}(x_{it-1}, x_{it}, z_{it-1}, z_{it}, e_{it-1}, \tau_{et}, \tau_t, \tau_{bt}) = c^t_{is}[x_{it}, z_{it}, e_{it}(\cdot), \tau_{et}, \tau_t, \tau_{bt}]$, $s = t, t+1$, with $e_{it}(\cdot)$ having the same arguments as $c_{is}(\cdot)$. Thus,

$$\frac{\partial c_{is}}{\partial x_{it}} = \frac{\partial c^t_{is}}{\partial x_{it}} + \frac{\partial c^t_{is}}{\partial e_{it}}\frac{\partial e_{it}}{\partial x_{it}} = \frac{\partial c^t_{is}}{\partial x_{it}}\left[1 + (1-\tau_{et})\frac{\partial e_{it}}{\partial x_{it}}\right],$$
$$\frac{\partial c_{is}}{\partial \tau_{et}} = \frac{\partial c^t_{is}}{\partial \tau_{et}} + \frac{\partial c^t_{is}}{\partial e_{it}}\frac{\partial e_{it}}{\partial \tau_{et}} = \frac{\partial c^t_{is}}{\partial x_{it}}\left[-e_{it} + (1-\tau_{et})\frac{\partial e_{it}}{\partial \tau_{et}}\right];$$

note that we use $\partial c^t_{is}/\partial e_{it} = (1-\tau_{et})\partial c^t_{is}/\partial x_{it}$. Substituting these expressions into $\Delta^{c_{is}}_{\tau_{et}} = \partial c_{is}/\partial \tau_{et} + e_{it}\partial c_{is}/\partial x_{it}$, we obtain

$$\Delta^{c_{is}}_{\tau_{et}} = \frac{\partial c^t_{is}}{\partial x_{it}}(1-\tau_{et})\left(\frac{\partial e_{it}}{\partial \tau_{et}} + e_{it}\frac{\partial e_{it}}{\partial x_{it}}\right) = \frac{\partial c^t_{is}}{\partial x_{it}}(1-\tau_{et})\Delta^{e_{it}}_{\tau_{et}}. \quad \text{(A14)}$$

Next, we again use the recursive structure of indirect utility, which allows us to reformulate the maximization problem of individual i of generation $t-1$ as

$$v^i_{t-1}(\cdot) = \max\left\{\widetilde{U}\left(c_{it-1}, \frac{z_{it-1}}{\omega_i}\right) + \delta v^i_t(\cdot) \mid c_{it-1} + \frac{\widetilde{e}_{it}}{1-\tau_{et}} \le x_{it-1} + e_{it-1}\right\}, \quad \text{(A15)}$$

where $\widetilde{e}_{it} \equiv (1-\tau_{et})e_{it}$ denotes bequests net of the inheritance tax. The maximization in (A15) is a standard textbook problem, and we can apply the Slutsky equation directly for $\widetilde{e}_{it}$, to obtain

$$\frac{\partial \widetilde{e}_{it}^{\text{com}}}{\partial \tau_{et}} = \frac{\partial \widetilde{e}_{it}}{\partial \tau_{et}} + \frac{\widetilde{e}_{it}}{(1-\tau_{et})^2}\frac{\partial \widetilde{e}_{it}}{\partial x_{it-1}}, \tag{A16}$$

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

In a further step, we show that $\partial \widetilde{e}_{it}^{\text{com}}/\partial \tau_{et}$ is equal to the term $(1-\tau_{et})\Delta_{\tau_{et}}^{e_{it}}$. To do so, we use $\widetilde{e}_{it} = (1-\tau_{et})e_{it}$ to derive that $\partial \widetilde{e}_{it}/\partial \tau_{et} = -e_{it} + (1-\tau_{et})\partial e_{it}/\partial \tau_{et}$ and $\partial \widetilde{e}_{it}/\partial x_{it-1} = (1-\tau_{et})\partial e_{it}/\partial x_{it-1}$. Moreover, we have $\partial e_{it}/\partial x_{it-1} = 1 + (1-\tau_{et})\partial e_{it}/\partial x_{it}$; for this, we have to differentiate the budget equation $c_{it-1} + e_{it} = x_{it-1} + e_{it-1}$ for period $t-1$ with respect to x_{it-1} and x_{it}, respectively (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 we use $\partial c_{it-1}/\partial x_{it-1} = (\partial c_{it-1}/\partial x_{it})/(1-\tau_{et})$. Using these terms in equation (A16), we find

$$\begin{aligned}\frac{\partial \widetilde{e}_{it}^{\text{com}}}{\partial \tau_{et}} &= -e_{it} + (1-\tau_{et})\frac{\partial e_{it}}{\partial \tau_{et}} + e_{it}\left[1 + (1-\tau_{et})\frac{\partial e_{it}}{\partial x_{it}}\right] \\ &= (1-\tau_{et})\left(\frac{\partial e_{it}}{\partial \tau_{et}} + e_{it}\frac{\partial e_{it}}{\partial x_{it}}\right) = (1-\tau_{et})\Delta_{\tau_{et}}^{e_{it}}.\end{aligned} \tag{A17}$$

Using equations (A14) and (A17), we can rewrite equation (A12) as

$$\begin{aligned}\frac{\partial S_2}{\partial \tau_{et}} = \sum_{i=L,H} & f_i\left(\beta\frac{\partial v_t^i}{\partial x_{it}} + \beta^2 U'_{it+1}\frac{\partial c_{it+1}^t}{\partial x_{it}}\right)\frac{\partial \tilde{e}_{it}^{\text{com}}}{\partial \tau_{et}} \\ &+ \mu\frac{\partial v_{t-1}^H[L]}{\partial x_{Lt}}(e_{Ht}[L] - e_{Lt}) \\ &+ \lambda\sum_{i=L,H}\left[\frac{\tau_{et}}{1-\tau_{et}} + \tau_t\left(\frac{\partial c_{it}^t}{\partial x_{it}} + \frac{\partial c_{it+1}^t}{\partial x_{it}}\right)\right]\frac{\partial \tilde{e}_{it}^{\text{com}}}{\partial \tau_{et}}.\end{aligned} \tag{A18}$$

Finally, we set τ_{et} and τ_t in equation (A18) equal to zero in order to obtain equation (10) in Proposition 4.

Because the compensated own-price effect is negative (i.e., $\partial \widetilde{e}_{it}^{\text{com}}/\partial \tau_{et} < 0$), the first term on the rhs of equation (10) is negative. Given the weak separability of preferences between consumption and labor, individual H, when mimicking, will leave more bequests than individual L (i.e., $e_{Ht}[L] > e_{Lt}$, if $e_{Ht-1} > e_{Lt-1}$), which implies that the second term in equation (10) is positive.

Equation (11): Characterization of τ_{et}^*

We make use of Proposition 3, which states that the welfare effects of τ_{et} and τ_t are equivalent (one of the taxes suffices), and we set τ_t equal to zero. Obviously, the welfare optimum (and thus the optimal tax rate τ_{et}^*) occurs when $\partial S_2/\partial\tau_{et} = 0$. Hence, it is obtained by setting the rhs of equation (A18) equal to zero, that is,

$$\sum_{i=L,H} f_i \left(\beta \frac{\partial v_t^i}{\partial x_{it}} + \beta^2 U'_{it+1} \frac{\partial c_{it+1}^t}{\partial x_{it}}\right) \frac{\partial \tilde{e}_{it}^{\mathrm{com}}}{\partial \tau_{et}} + \mu \frac{\partial v_{t-1}^H[L]}{\partial x_{Lt}} (e_{Ht}[L] - e_{Lt}) + \frac{\lambda \tau_{et}^*}{1-\tau_{et}^*} \sum_{i=L,H} \frac{\partial \tilde{e}_{it}^{\mathrm{com}}}{\partial \tau_{et}} = 0.$$

By using the abbreviation $W(\tau_{et}^*) \equiv \sum_{i=L,H} f_i(\beta\partial v_t^i/\partial x_{it} + \beta^2 U'_{it+1}\partial c_{it+1}^t/\partial x_{it})\partial\tilde{e}_{it}^{\mathrm{com}}/\partial\tau_{et} + \mu(\partial v_{t-1}^H[L]/\partial x_{Lt})(e_{Ht}[L] - e_{Lt})$ and by simple rearrangement, we obtain equation (11) in the text.

Proof of Proposition 5

The derivative of the optimal value $S_2(\tau_{et}, \tau_t, \tau_{bt})$ with respect to τ_{bt} (at $\tau_{et} = \tau_t = \tau_{bt} = 0$) is

$$\frac{\partial S_2}{\partial \tau_{bt}} = \sum_{i=L,H} f_i \left(\frac{\partial v_{t-1}^i}{\partial \tau_{bt}} + \beta \frac{\partial v_t^i}{\partial \tau_{bt}} + \beta^2 U'_{it+1} \frac{\partial c_{it+1}}{\partial \tau_{bt}}\right) + \mu \left(\frac{\partial v_{t-1}^H}{\partial \tau_{bt}} - \frac{\partial v_{t-1}^H[L]}{\partial \tau_{bt}}\right) + \lambda \sum_{i=L,H} c_{it+1}. \tag{A19}$$

Again, as in the proof of Proposition 3, the welfare of generation t is influenced directly and indirectly by τ_{bt}, and thus $\partial v_t^i/\partial\tau_{bt} = (-c_{it+1} + \partial e_{it}/\partial\tau_{bt})\partial v_t^i/\partial x_{it}$. Moreover, we have $\partial v_{t-1}^i/\partial\tau_{bt} = -c_{it+1}\partial v_{t-1}^i/\partial x_{it}$ and $\partial v_{t-1}^H[L]/\partial\tau_{bt} = -c_{Ht+1}[L]\partial v_{t-1}^H[L]/\partial x_{Lt}$. Using these terms, together with the first-order conditions (A8) and (A9) at $\tau_{et} = \tau_t = \tau_{bt} = 0$, multiplied by c_{Lt+1}, c_{Ht+1}, respectively, equation (A19) can be transformed to the equation in Proposition 5, where we use $\partial\tilde{e}_{it}^{\mathrm{com}}/\partial\tau_{bt} = c_{it+1}\partial\tilde{e}_{it}/\partial x_{it} + \partial\tilde{e}_{it}/\partial\tau_{bt}$ and $\partial c_{it+1}^{\mathrm{com}}/\partial\tau_{bt} = c_{it+1}\partial c_{it+1}/\partial x_{it} + \partial c_{it+1}/\partial\tau_{bt}$.

The term $\partial c_{it+1}^{\mathrm{com}}/\partial\tau_{bt} < 0$ is the negative compensated own-price effect on c_{it+1}, as decided by generation $t-1$ (we use the combined budget equation $c_{it-1} + c_{it} + (1+\tau_{bt})c_{it+1} = x_{it-1} + e_{it-1} + x_{it}$ at $\tau_{et} = \tau_t = 0$ and we apply the Slutsky equation). Furthermore, $\partial\tilde{e}_{it}^{\mathrm{com}}/\partial\tau_{bt}$ is the compensated demand effect on net bequests $\tilde{e}_{it}$ (equal to e_{it}, because $\tau_{et} = 0$). The fact that this effect is also negative follows from $\partial\tilde{e}_{it}^{\mathrm{com}}/\partial\tau_{bt} = \partial c_{it}^{\mathrm{com}}/\partial\tau_{bt} + \partial c_{it+1}^{\mathrm{com}}/\partial\tau_{bt}$, together with the homogeneity of compensated demand (i.e.,

$\partial c_{it-1}^{\rm com}/\partial \tau_{bt} + \partial c_{it}^{\rm com}/\partial \tau_{bt} + (1+\tau_{bt})\partial c_{it+1}^{\rm com}/\partial \tau_{bt} = 0$ at $\tau_{bt} = 0$), and from the fact that for additive preferences all compensated cross price effects are positive (except in pathological cases; see Barten and Böhm, 1982, p. 425).

Given the weak separability of preferences between consumption and labor, and given $e_{Ht-1} > e_{Lt-1}$, the last term in the equation is positive for the analogous reason stated in the proof of Proposition 4.

References

Arrondel, L. and Laferrère, A. (2001), Taxation and Wealth Transmission in France, *Journal of Public Economics 79*, 3–33.

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

Barten, A. P. and Böhm, V. (1982), Consumer Theory, in K. J. Arrow and M. D. Intriligator (eds.), *Handbook of Mathematical Economics*, Vol. II, North Holland, Amsterdam, 381–429.

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. (2012), Optimal Taxation of Wealth Transfers When Bequests are Motivated by Joy of Giving, *The B.E. Journal of Economic Analysis & Policy*, *12*, Article 8.

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.

Corlett, W. J. and Hague, D. C. (1953), Complementarity and the Excess Burden of Taxation, *Review of Economic Studies 21*, 21–30.

Cremer, H. and Pestieau, P. (2006), Wealth Transfer Taxation: A Survey of the Theoretical Literature, in S.-Ch. Kolm and J. M. Ythier (eds.), *Handbook of the Economics of Giving, Altruism and Reciprocity*, Vol. 2., North-Holland, Amsterdam, 1107–1134.

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.

Davies, J. B. and Shorrocks, A. F. (2000), The Distribution of Wealth, in A. B. Atkinson and F. Bourguignon (eds.), *Handbook of Income Distribution*, Vol. 1. North-Holland, Amsterdam, 605–675.

Diamond, P. (2007), Comment, in D. Acemoglu, K. Rogoff, and M. Woodford (eds.), *NBER Macroeconomics Annual 2006*, Vol. 21, MIT Press, Cambridge, MA, 365–379.

Díaz-Giménez, J., Glover, A., and Ríos-Rull, J.-V. (2011), Facts on US Distributions of Earnings, Income and Wealth: 2007 Update, *Federal Reserve Bank of Minneapolis Quarterly Review 34*, 2–31.

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

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

Golosov, M., Tsyvinski, A., and Werning, I. (2007), New Dynamic Public Finance: A User's Guide, in D. Acemoglu, K. Rogoff, and M. Woodford (eds.), *NBER Macroeconomics Annual 2006*, Vol. 21, MIT Press, Cambridge, MA, 317–363.

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

Laitner, J. and Ohlsson, H. (2001), Bequest Motives: A Comparison of Sweden and the United States, *Journal of Public Economics 79*, 205–236.

Masson, A. and Pestieau, P. (1997), Bequests Motives and Models of Inheritance: A Survey of the Literature, in G. Erreygers and T. Vandevelde (eds.), *Is Inheritance Legitimate?*, Springer, Berlin, 54–88.

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*, Elsevier, Amsterdam, 417–435.

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 Tax, *Review of Economic Studies 49*, 637–643.

Solon, G. (1999), Intergenerational Mobility in the Labor Market, in O. C. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, Vol. 3A, North-Holland, Amsterdam, 1761–1800.

Solon, G. (2002), Cross-Country Differences in Intergenerational Earnings Mobility, *Journal of Economic Perspectives 16*, 59–66.

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

First version submitted April 2010;
final version received July 2011.