Department of Economics

Johannes Kepler University of Linz

Optimal Taxation of Income and Bequests

by

Johann K. Brunner*)

Working Paper No. 9722

July 1997

Johannes Kepler University of Linz
Department of Economics
Altenberger Strasse 69
A-4040 Linz - Auhof, Austria
www.economics.uni-linz.ac.at

*) corresponding author:
johann.brunner@jku.at
phone +43 (0)70 2468 -8248, -9821 (fax)

I. Introduction

A standard result in optimum taxation theory says that if preferences are appropriately separable in leisure and the bundle of consumption goods, it suffices to tax wage income, no consumption good needs to be taxed. This result was first derived in the representative-consumer model of the Ramsey type with linear taxes, where it requires implicit separability (of which weak separability together with unitary expenditure elasticities of all goods is a special case. See, e. g., Deaton 1981). More interesting, however, is the case of an economy with differing individuals, where a redistributive target arises in addition to efficiency. If individuals differ in their ability to earn income, but are equal otherwise, and if preferences are weakly separable, then an optimal strategy for the government is to impose a nonlinear tax on wages and let consumption untaxed (Deaton 1981). As a straightforward implication one concludes that, given a utility function, which is weakly separable in leisure and consumption in different periods, and given an optimal tax on wages, taxation of interest income is harmful (Atkinson and Stiglitz 1980, p. 442). Ordover and Phelps (1979) showed this more generally in a growth model with overlapping-generations, where capital formation (and interest income) comes from the saving for second-period consumption according to the life-cycle motive.

In this paper we study the question of how bequests should be treated in an optimal tax structure. Most countries have in fact established a tax on bequests, which appears mainly to be motivated by redistributive reasons. However, with just the same argument many people plead for a tax on interest income, which economic reasoning seems to disprove, as was mentioned above. Therefore, a detailed analysis of the role of a tax on bequests in a model with differing individuals appears important.[1] In doing this, one has to ask for the reason why individuals leave bequests. As is well-known, several motives can be distinguished: unintended bequests due to ignorance of own life-time; strategic bequests in order to receive care from descendants; purely altruistic bequests. Obviously, those in the first category are not valued as such by the decision-making individual, hence their supply seems to be independent of taxation, which clearly makes them a preferential object of taxation.

[1] Atkinson and Stiglitz 1980, ch. 3, consider the effect of a tax on bequests in a single-consumer model.

The point of optimal taxation theory is, however, to take into account the reaction of individuals on the taxes imposed by the authority. Such reactions can be expected for bequests falling into the second and third category mentioned above, and these will be considered in the present study. We incorporate these motives simply by assuming that bequests appear as an argument in the utility function of the testator, without analyzing further whether a strategic motive or altruism lies behind.[2] Thus, in principle, bequests are considered in the same way as consumption of goods, and one might draw the same conclusion as above, saying that no tax should be imposed if weak separability holds.

However, this conclusion rests on the assumption that individuals differ only in their earning abilities. Such an assumption (which appears to be rather unrealistic in any case) makes no sense in an economy with bequests: Persons with differing abilities, thus differing incomes, will also leave unequal bequests to their descendants. Therefore we have to consider an extended model, where individuals may differ in two characteristics: earning abilities and inherited wealth. In such a model the above-mentioned result does not hold any more, because the redistributive target becomes more important. Indeed, it will be shown that under the assumption that bequests are a normal good a tax on them is desirable, in addition to the optimal nonlinear income tax, if the social welfare function favours redistribution strongly enough to outweigh the distorting consequences. The reason is that such a tax represents a substitute for a first-best tax on endowments.

For the aim of comparison, in the following section we formulate a simple version of the standard model of optimal income and commodity taxation of two individuals with differing earning abilities, and show why commodity taxation is undesirable, given weak separability. In Section III the model will be adapted in order to incorporate bequests, and the main results will be stated. Section IV deals with an extension to differentiated tax rates, Section V contains concluding remarks.

II. A simple model of optimal income and commodity taxation

For the aim of comparison, we first analyze the structure of direct and indirect taxation in a simple model. Let two individuals be given, whose identical preferences over two consumption goods c, d

[2] A formulation where utility of the heir appears in the testator's utility function (see, e. g. Barro 1974) might be more suitable for the description of pure altruism. However, the formulation chosen in this paper allows a more

and labor time l are described by a concave utility function u(c,d,l) with partial derivatives u_c, $u_d > 0$, $u_l < 0$. The individuals differ in their earning abilities $w_1 < w_2$, which can, by appropriate scaling, be taken as their respective wage rates.

We use a mixed primal - dual approach. For any pair of prices p_c, p_d for c and d, resp., we consider the Hicksian composite commodity x, and introduce for every individual the utility function $v^i(x, z, p_c, p_d) \equiv \max \{u(c,d,z/w_i) \mid p_c c + p_d d = x\}$. In this formulation, z denotes gross income or efficient labor supply, hence z/w is the labor time for an individual with wage rate w_i, necessary to supply (and earn) z. The composite commodity x represents net income. Furthermore, we define the slope of an indifference curve with respect to x and z as $\sigma^i(x, z, p_c, p_d) \equiv -v^i_z(x, z, p_c, p_d) / v^i_x(x, z, p_c, p_d)$. We assume the condition of agent monotonicity

AM: $\sigma^1(x, z, p_c, p_d) > \sigma^2(x, z, p_c, p_d)$, for any x, z and p_c, p_d.

This standard assumption in optimal income taxation theory (see, e. g., Seade 1982) requires that, for any given prices p_c, p_d and for any bundle x, z, the marginal rate of substitution between gross income and composite consumption is lower for the more able individual. An immediate consequence of this condition is that for any income tax function t(z) depending on gross income, the more able individual chooses no lower value of gross (and net) income than the less able (see Brunner 1989, p. 26f). This seems to be a mild assumption (it is implied by normality of net income), which ensures that redistribution of income (which is observable) from the top to the bottom goes into the right direction, namely from high to low ability (which itself is unobservable).

Let τ_c, τ_d be unit taxes on the consumption goods and $c_i(x_i, z_i, p_c + \tau_c, p_d + \tau_d)$, $d_i(x_i, z_i, p_c + \tau_c, p_d + \tau_d)$, for i = 1, 2, be the demand functions of the individuals. Using, for simplicity, a utilitarian social welfare function with weights $f_1 \geq f_2$, the authority faces a problem of tax design, which we formulate in two steps. First we take the commodity tax rates τ_c, τ_d as given and determine the optimum nonlinear income tax as the solution of:

---

convenient treatment without being too specific.

(1) $$\max_{x_i, z_i} \; f_1 v^1(x_1, z_1, p_c + \tau_c, p_d + \tau_d) + f_2 v^2(x_2, z_2, p_c + \tau_c, p_d + \tau_d)$$

s. t.

(2) $$x_1 + x_2 \le z_1 + z_2 + \tau_c(c_1 + c_2) + \tau_d(d_1 + d_2) - p_g g,$$

(3) $$v^2(x_2, z_2, p_c + \tau_c, p_d + \tau_d) \ge v^2(x_1, z_1, p_c + \tau_c, p_d + \tau_d),$$

(4) $$x_i, z_i \ge 0.$$

In the resource constraint (2) (where, for shortness, the arguments of the demand functions c_i and d_i are dropped) we have assumed a linear production function: p_c units of effective labor are used to produce one unit of c, the same for d. This gives $p_c(c_1 + c_2) + p_d(d_1 + d_2) \le z_1 + z_2 - p_g g$, where $p_g g$ denotes the resource requirement of the state. Adding $\tau_c(c_1 + c_2) + \tau_d(d_1 + d_2)$ on both sides leads to (2). (3) is the so-called self-selection constraint (Stiglitz 1982), which ensures incentive compatibility: The values of x_i, z_i, i = 1, 2, determined by the state, must have the property that the more able individual does not prefer to earn gross (and net) income of the less able, which he could do with less effort than the latter. (In principle, such a condition should also be formulated for the less able, but with the redistributive target expressed by the utilitarian objective, one can show that it is automatically fulfilled by the optimal solution. See Brunner 1989, p. 190f.[3]) If (3) is satisfied, then the government can construct a (nonlinear) tax function (say, a step function), such that the less (more) able will just choose the bundle x_1, z_1 (x_2, z_2, resp.).

Introducing the Lagrange variables λ and μ, associated with (2) and (3), resp., we derive the first-order conditions for the maximization of (1) as[4]

(5) $$f_1 \frac{\partial v^1}{\partial x_1} - \lambda + \lambda\tau_c \frac{\partial c_1}{\partial x_1} + \lambda\tau_d \frac{\partial d_1}{\partial x_1} - \mu \frac{\partial v^2}{\partial x_1} = 0,$$

[3] Moreover, one can show that with a utilitarian objective, (3) is binding if u is strictly concave or $f_1 > f_2$. This will be assumed in the following.

(6) $$f_1 \frac{\partial v^1}{\partial z_1} + \lambda + \lambda\tau_c \frac{\partial c_1}{\partial z_1} + \lambda\tau_d \frac{\partial d_1}{\partial z_1} - \mu \frac{\partial v^2}{\partial z_1} = 0 \text{ (or < 0 and } z_1 = 0\text{)},$$

(7) $$f_2 \frac{\partial v^2}{\partial x_2} - \lambda + \lambda\tau_c \frac{\partial c_2}{\partial x_2} + \lambda\tau_d \frac{\partial d_2}{\partial x_2} + \mu \frac{\partial v^2}{\partial x_2} = 0,$$

(8) $$f_2 \frac{\partial v^2}{\partial z_2} + \lambda + \lambda\tau_c \frac{\partial c_2}{\partial z_2} + \lambda\tau_d \frac{\partial d_2}{\partial z_2} + \mu \frac{\partial v^2}{\partial z_2} = 0,$$

Let $S(\tau_c,\tau_d)$ be the value of the objective, given the optimum solution of (1) - (4). Our main interest is, whether commodity taxation is desirable, in addition to the income tax. For this aim, in the second step we analyze $\partial S/\partial\tau_c$ and $\partial S/\partial\tau_d$ at $\tau_c = \tau_d = 0$. Applying the envelope theorem, we have

(9) $$\frac{\partial S}{\partial \tau_c} = f_1 \frac{\partial v^1}{\partial \tau_c} + f_2 \frac{\partial v^2}{\partial \tau_c} + \lambda(c_1 + c_2 + \tau_c(\frac{\partial c_1}{\partial \tau_c} + \frac{\partial c_2}{\partial \tau_c}) + \tau_d(\frac{\partial d_1}{\partial \tau_c} + \frac{\partial d_2}{\partial \tau_c})) + \mu(\frac{\partial v^2}{\partial \tau_c} - \frac{\partial v^2[1]}{\partial \tau_c}),$$

(10)

$$\frac{\partial S}{\partial \tau_d} = f_1 \frac{\partial v^1}{\partial \tau_d} + f_2 \frac{\partial v^2}{\partial \tau_d} + \lambda(\tau_c(\frac{\partial c_1}{\partial \tau_d} + \frac{\partial c_2}{\partial \tau_d}) + d_1 + d_2 + \tau_d(\frac{\partial d_1}{\partial \tau_d} + \frac{\partial d_2}{\partial \tau_d})) + \mu(\frac{\partial v^2}{\partial \tau_d} - \frac{\partial v^2[1]}{\partial \tau_d})$$

.

In (9) and (10), the symbol [1] indicates that the derivative of v^2 is taken at x_1, z_1. By Roy's Lemma, we have $\partial v^i / \partial\tau_c = -c_i\, \partial v^i / \partial x_i$, $\partial v^i / \partial\tau_d = -d_i\, \partial v^i / \partial x_i$. Moreover, we can compute from (5) and (7), at $\tau_c = \tau_d = 0$:

(11) $$\lambda = f_1 \frac{\partial v^1}{\partial x_1} - \mu \frac{\partial v^2}{\partial x_1} \text{ and } \lambda = f_2 \frac{\partial v^2}{\partial x_2} + \mu \frac{\partial v^2}{\partial x_2}.$$

---

[4] For shortness, we disregard the possibility of corner solutions except the one with $z_1 = 0$.

Using (11), in turn, we can transform (9) to[5]

$$\frac{\partial S}{\partial \tau_c} = -f_1 c_1 \frac{\partial v^1}{\partial x_1} - f_2 c_2 \frac{\partial v^2}{\partial x_2} + f_1 c_1 \frac{\partial v^1}{\partial x_1} - \mu c_1 \frac{\partial v^2}{\partial x_1} + f_2 c_2 \frac{\partial v^2}{\partial x_2} + \mu c_2 \frac{\partial v^2}{\partial x_2} - \mu c_2 \frac{\partial v^2}{\partial x_2} + \mu c_2[1] \frac{\partial v^2}{\partial x_1}$$

and further to

(12) $$\frac{\partial S}{\partial \tau_c} = \mu \frac{\partial v^2}{\partial x_1}(c_2[1] - c_1).$$

In the same way, we get from (10)

(13) $$\frac{\partial S}{\partial \tau_d} = \mu \frac{\partial v^2}{\partial x_1}(d_2[1] - d_1).$$

With these formulas we can proof

Result 1: Given the optimum nonlinear income tax, if the utility function is weakly separable in leisure and the consumption goods, then no commodity should be taxed, i. e., $\tau_c = \tau_d = 0$. Otherwise, the good which is less complementary with labor should be taxed.

Proof: As μ is a positive multiplier and $\partial v^2/\partial x$ is positive as well, the signs of $\partial S / \partial \tau_c$ and $\partial S / \partial \tau_d$ depend on $c_2[1] - c_1$ and $d_2[1] - d_1$, resp. Consider $c_2[1]$, which denotes what the more able individual would demand of good c, would his income be that of the less able, i. e., would he earn gross income z_1 and net income x_1. Obviously, the difference between $c_2[1]$ and c_1 comes from the fact that the more able individual would have to work less for the same income. Now, if preferences are weakly separable, the marginal rate of substitution between c and d is independent of labor supply, therefore $c_2[1] = c_1$ and $d_2[1] = d_1$, which means $\partial S / \partial \tau_c = \partial S / \partial \tau_d = 0$ and no good should be taxed. Otherwise, a positive tax rate

[5] Note that $\partial v^2[1] / \partial \tau_c = -c_2[1]\, \partial v^2 / \partial x_1$, where $c_2[1]$ denotes consumption of c by individual two if he had

should be imposed on the good, whose demand increases with leisure. Obviously, as income is fixed, demand for the other good decreases with leisure, it should be subsidized. QED.

Thus we have reestablished, in our simple model, the standard result of optimal commodity taxation with differing individuals.[6] In the present paper, we are mainly interested in the optimal taxation of interest income and bequests. The first of these questions can be answered within the above model by interpreting the two goods as consumption in two different periods, with $p_d = p_c/(1 + r)$, where r is the interest rate. Then the above result tells us that in case of weak separability there should be no (linear) tax on interest income.

We can also reproduce the well-known properties of the optimal income tax easily:

Result 2: Assume that commodity taxes are zero. Then the optimal marginal income tax rate is positive for the less able individual and zero for the more able.

Proof: Collecting terms and dividing (8) by (7), we find, with $\tau_c = \tau_d = 0$,

$$-\frac{\partial v^2}{\partial z_2} / \frac{\partial v^2}{\partial x_2} = 1,$$

while (5) and (6) give

$$(14) \qquad -\frac{\partial v^1}{\partial z_1} / \frac{\partial v^1}{\partial x_1} = \frac{\lambda - \mu\, \partial v^2 / \partial z_1}{\lambda + \mu\, \partial v^2 / \partial x_1}.$$

Due to AM, we have $-\frac{\partial v^1}{\partial z_1} / \frac{\partial v^1}{\partial x_1} > -\frac{\partial v^2}{\partial z_1} / \frac{\partial v^2}{\partial x_1}$, moreover $\lambda > 0$, $\mu > 0$, $\frac{\partial v^2}{\partial z_1} < 0$, therefore (14) can hold only if $(\lambda - \mu\frac{\partial v^2}{\partial z_1})/(\lambda + \mu\frac{\partial v^2}{\partial x_1}) < 1$. QED.

---

income x_1.

III. Taxation of bequests

Now we modify the model in order to analyze the taxation of bequests. Analogous to the approach indicated above for the case of interest taxation, we simply interpret good d as bequests, which produce utility for the testator, be it for altruistic or strategic reasons, as discussed in the introduction. The main new element occuring in this section is that bequests go to the descendants as inherited wealth, and constitute a second differentiating factor, in addition to earning ability. This turns out to be an important fact.

Moreover, bequests build up a link between different generations, hence we have to modify the static model of the foregoing Section into a dynamic one. We assume that individuals live for one period, where they earn income, consume c and bequeath d, with prices normalized to $p_c = p_d = 1$. We indicate periods (and generations) by t. As in Section II we introduce the primal - dual utility function $v_t^i(x_t, z_t, 1+\tau_{ct}, 1+\tau_{dt}, d_{it-1}) \equiv \max \{u(c_t, d_t, z_t/w_i) \mid (1+\tau_{ct})c_t + (1+\tau_{dt})d_t = x_t + d_{it-1}(1+r)\}$. Here r denotes the interest rate, which is assumed constant over time. For simplicity, wage rates are assumed constant as well, that is, we do not consider the influence of capital formation on labor productivity.[7]

A specific question in the present framework is how abilities and inheritances are correlated in the course of time. We make the straightforward assumption that the descendants have the same ability as their parents and receive what the latter bequeath. Certainly, this is a rather pronounced formulation which brings the distributive issue to the point. However, if one regards the two types of individuals as representatives of two distinct classes of society, between which (nearly) no intermarriage takes place, then the assumption does not seem to be unrealistic.[8]

---

[6] That commodities are taxed differently in accordance with complementarity/substitutability with labor is sometimes called the Corlett-Hague rule. (Corlett and Hague 1953)

[7] In other words, in every period we have the same linear production function depending on (bequeathed) capital and labor.

[8] In a sense we could think of two dynasties. These are usually assumed to make decisions with an infinite time horizon. In our model the tax authority has the power to interfere whenever a new generation appears.

τ_{dt} denotes the tax on bequests in period t, individuals are assumed to care about net bequests d_{it}.[9] What makes this problem conceptually different from that of Section II is that inheritances enter the budget condition. In the initial period, inheritances d_{i0} are taken as given, later on, they are determined by individual decisions. For a derivation of results it is necessary to formulate assumptions on the distribution of the d_0. As appears natural, we assume that initial endowment is not smaller for the more able individual than for the less able:

BD: $d_{10} \le d_{20}$.

Moreover, we have to extend the condition of agent monotonicity. We do this in an obvious way:

AMD: $\sigma^1(x, z, p_c, p_d, d_1) > \sigma^2(x, z, p_c, p_d, d_2)$, for any x and z and $d_1 \le d_2$.

As $\partial^2 v^i / \partial x_{it}^2 \le 0$, by concavity of u, AMD implies that the marginal utility of net income does not decrease too fast or that the distance between d_1 and d_2 is not too large, compared to the distance between w_1 and w_2.

As we have now a dynamic model, an intertemporal norm is required for the social decision. We assume again that the social welfare function is utilitarian, with weights $f_1 \ge f_2$ within a generation, and with diminishing weights of later generations, according to a social rate of time preference $\gamma \ge 0$. For a finite-time horizon T we arrive at the following formulation, where we take the series of τ_{ct}, τ_{dt} as given and normalize p_{gt} to one as well:

$$(15)\qquad \max_{x_{it}, z_{it}} \sum_{t=1}^{T} (1+\gamma)^{-t+1} (f_1 v_t^1(x_{1t}, z_{1t}, 1+\tau_{ct}, 1+\tau_{dt}, d_{1t-1}) + f_2 v_t^2(x_{2t}, z_{2t}, 1+\tau_{ct}, 1+\tau_{dt}, d_{2t-1}))$$

s. t.

$$(16)\qquad x_{1t} + x_{2t} \le z_{1t} + z_{2t} + \tau_{ct}(c_{1t} + c_{2t}) + \tau_{dt}(d_{1t} + d_{2t}) - g_t, \qquad t = 1, ..., T$$

[9] Note that if individuals care about gross bequests, a tax on these would be lump-sum in our model.

(17) $$v_t^2(x_{2t}, z_{2t}, 1+\tau_{ct}, 1+\tau_{dt}, d_{2t-1}) \geq v_t^2(x_{1t}, z_{1t}, 1+\tau_{ct}, 1+\tau_{dt}, d_{2t-1}), \quad t = 1, ..., T$$

(18) $$x_{it}, z_{it} \geq 0 .$$

We have again assumed that only the self-selection condition for individual two is binding in the optimum. One observes that via bequests the values of the variables in earlier periods influence the decisions of all forthcoming generations. We abbreviate this link by considering only two generations, i. e., T = 2.

The first-order conditions regarding the variables in period 1 are given in Appendix A. Similar to the procedure in Section II, the optimum value of the social objective is denoted by S and its derivative $\partial S / \partial \tau_{d1}$, at $\tau_{d1} = \tau_{c1} = 0$ is derived as (see Appendix A):

(19) $$\frac{\partial S}{\partial \tau_{d1}} = \frac{1}{1+\gamma}\left(f_1 \frac{\partial v_2^1}{\partial d_{11}} \frac{\partial d_{11}^h}{\partial \tau_{d1}} + f_2 \frac{\partial v_2^2}{\partial d_{21}} \frac{\partial d_{21}^h}{\partial \tau_{d1}}\right) + \mu_1 \frac{\partial v_1^2}{\partial x_{11}}(d_{21}[1] - d_{11}) + \mu_2\left(\frac{\partial v_2^2}{\partial d_{21}} - \frac{\partial v_2^2[1]}{\partial d_{21}}\right) \frac{\partial d_{21}^h}{\partial \tau_{d1}}.$$

where the upper index h denotes Hicksian demand and μ_1 and μ_2 are the Lagrange variables referring to (17) for t = 1, 2.

The three terms occuring in (19) have a straightforward interpretation: The first one shows the net influence of the bequest tax, whose revenue is distributed to the individuals via the income tax, on the inheritances (and, thus, on welfare) of the following generation. Note that own compensated price effects are always negative. The second term in brackets is familiar from the foregoing section, it refers to the difference in bequests, whether the more or the less able individual receives gross (net) income z_{11} (x_{11}). Its sign depends on substitutability/complementarity of bequests to labor time and on the difference between d_{20} and d_{10}. The third term expresses the influence of τ_{d1} on the self-selection constraint of the following period. Its sign is undetermined in general.

The derivation of definite results requires the restriction to a specific type of preferences. In case that the initial endowments are identical, we can draw a conclusion similar to Result 1:

Result 3: Let $d_{10} = d_{20}$. If the utility function $u(c_t,d_t,l_t)$ is of the form $U(\varphi(c_t, d_t) + \psi(l_t))$, where φ is linear homogeneous, U is concave and ψ is strictly concave with $\psi' < 0$ (i. e., u is quasilinear in net income and leisure, thus also weakly separable[10]), then no tax on bequests is optimal. In fact, bequests should be subsidized.

Proof: We consider the derivative of the intertemporal social objective with respect to τ_{d1}, at $\tau_{d1} = 0$, as expressed by (19). With weak separability and because of $d_{10} = d_{20}$ we have $d_{21}[1] = d_{11}$, thus the second term in (19) is zero. Moreover, $\partial v_2^2 / \partial d_{21} = (1+r)\partial v_2^2 / \partial x_{22} = (1+r)BU'(\varphi(c_{22}, d_{22}) + \psi(l_{22}))$, where $B = \partial\varphi / \partial x$ is a constant, due to linear homogeneity of φ. In the same way, $\partial v_2^2[1] / \partial d_{21} = (1+r)BU'(\varphi(c_{22}[1], d_{22}[1]) + \psi(l_{12}w_1 / w_2))$. The self-selection constraint is binding, therefore $\varphi(c_{22}, d_{22}) + \psi(l_{22}) = \varphi(c_{22}[1], d_{22}[1]) + \psi(l_{12}w_1 / w_2)$, which means that the third term in (19) is zero as well. Finally, we have already noted that the first term is negative, due to the negativity of own compensated price effects. QED.

The motivation for a subsidy on bequests obviously comes from the fact that they create twofold utility: for the testator and for the heir. It must be stressed, however, that this result is strongly based on the assumption of equal initial endowments. Closer inspection shows that this is not a reasonable assumption: If bequests are a normal good, then one can expect that more able individuals will bequeath more to their descendants than less able, which means that endowments will not be equal in the following generations. Indeed, for homogeneous φ one derives $d_t = \beta_d(\tau_{dt})(x_{it} + d_{it-1}(1+r))$, where $0 < \beta_d(\tau_d) < 1$ and $\beta'_d < 0$. A steady state, where $d_{it} = d_{it-1} = d_i$, is characterized by $d_i = x_i\beta_d(\tau_{dt})/(1-\beta_d(\tau_{dt})(1+r))$.

In view of this, we turn now to the case that unequal capital endowments occur in period one, as a result of bequests of foregoing generations. We ask, how this additional differentiating factor, besides differing earning abilities, influences the conclusion concerning the desirability of a tax on bequests in the first generation.[11] On this question we formulate two results which indicate the essential points of the problem.

Result 4: Let $d_{10} < d_{20}$ and let the utility function be of the form as in Result 3. If bequests are a normal good and the compensated demand for bequests is completely inelastic, then a tax on bequests is desirable.

Proof: We consider again formula (19). The third term in brackets is zero, as was shown in the proof of Result 3. Completely inelastic compensated demand means that the first term is zero. Finally, normality of bequests and the larger endowment of individual two imply $d_{21}[1] > d_{11}$. As $\mu_1 > 0$ and the marginal utility of net income is positive, we find that the derivative (19) is positive. QED.

Result 5: Let the utility function be of the form as in Result 3 and $U(x) = x$. If f_1 is sufficiently larger than f_2 and d_{10} is sufficiently smaller than d_{20}, then a tax on bequests is desirable.

Proof: In Appendix B it is shown that with the specification of u, at $\tau_{d1} = 0$, (19) reads

$$(20) \qquad \frac{\partial S}{\partial \tau_{d1}} = \frac{1}{1+\gamma} B(0)^2 (1+r) \rho_d^{'}(0)[f_1(x_{11} + d_{10}(1+r)) + f_2(x_{21} + d_{20}(1+r))] + \mu_1 B(0) \beta_d(0)(d_{20} - d_{10})(1+r),$$

where $B(0) > 0$, $\rho_d^{'}(0) < 0$ are constants. Moreover

$$(21) \qquad \mu_1 = \frac{f_1 - f_2}{2}\left(1 + \frac{(1+r)\beta_d(0)}{1+\gamma}\right)$$

[10] Quasilinear preferences are frequently used in optimal taxation theory. See, e. g., Weymark 1987.

[11] In a static model, one would suggest the introduction of a tax on endowments as a first-best measure for redistribution. In our dynamic model this would only be possible in the starting period. We concentrate on the

and not only μ_1, but also x_{11} and x_{21} are independent of d_{10} and d_{20}. In general, the sign of $\partial S / \partial \tau_{d1}$ is indeterminate, as the first term in (20) is negative while the second is positive. However, using (21) in (20), one observes that an increase of d_{20} changes $\partial S / \partial \tau_{d1}$ by $-f_2C+(f_1-f_2)D$, where C and D are positive constants. Thus, for sufficiently large f_1 this change is positive, then if d_{20} increases sufficiently, (20) become eventually positive, which means that a tax on bequests is desirable. QED.

Remark: As mentioned above, the difference in initial endowments d_{10} and d_{20} arises as the result of bequests of foregoing generations, ultimately it is determined by the difference in abilities w_1 and w_2. In Appendix C an example is provided where, given a steady-state with $\tau_{d1} = 0$, the introduction of a tax on bequests is desirable, according to the considerations of Result 5.

The Results 4 and 5 may be interpreted as theoretical underpinnings of a widely held view on the tax on bequests, namely that it represents an effective means for redistribution and is necessary because people differ in initial endowments. However, the tax causes an adverse substitution effect. Result 4 states that if this effect can be neglected, then a tax on bequests improves social welfare. Otherwise the redistributive motive and the initial differences have to be large enough to outweigh the adverse effect of the tax (Result 5).

Finally, we discuss how the existence of bequests changes the optimum marginal tax rate on wage income.

Result 6: Assume that the utility function is of the form as in Result 3 and that the taxes on consumption and bequests are zero. Then the optimal marginal income tax rate is negative for the more able individual.

Proof: Dividing (A4) by (A3) of Appendix A, after collecting terms, one gets for the more able individual (note $\partial d_{21}/\partial z_{21} = 0$ and $\partial v_2^2 / \partial d_{21} = \partial v_2^2[1] / \partial d_{21}$ for quasilinear utility)

---

more realistic measure, namely to impose a bequest tax, which is a tax on endowments of the following generation,

(22) $$-\frac{\partial v_1^2}{\partial z_{21}} / \frac{\partial v_1^2}{\partial x_{21}} = \lambda_1 / (\lambda_1 - \frac{1}{1+\gamma} f_2 \frac{\partial v_2^2}{\partial d_{21}} \frac{\partial d_{21}}{\partial x_{21}}).$$

The right-hand side of (22) is larger than one, if bequests are normal. QED.

Remark: The analogous expression for the less able individual is, from (A1) and (A2):

(23) $$-\frac{\partial v_1^1}{\partial z_{11}} / \frac{\partial v_1^1}{\partial x_{11}} = (\lambda_1 - \mu_1 \frac{\partial v_1^2}{\partial z_{11}}) / (\lambda_1 + \mu_1 \frac{\partial v_1^2}{\partial x_{11}} - \frac{1}{1+\gamma} f_1 \frac{\partial v_2^1}{\partial d_{11}} \frac{\partial d_{11}}{\partial x_{11}})$$

Without the last term in the denominator, the right-hand side of (23) would be lower than one, as was shown in the proof of Result 2. Thus, with this term subtracted, no definite statement on the marginal tax rate can be given.

The reason, why the existence of bequests calls for a reduction of the marginal tax rates, compared to the situation in Section II, again comes from the fact that bequests create twofold utility: for the testator and for the heir. Therefore, an incentive to work more, which leads to increased income and bequests, is efficient.

IV. Differing tax rates

It is instructive to analyse how the situation changes if differing tax rates τ_{dt}^i, i = 1, 2 may be imposed on the bequests of the two types of individuals. To model this we return to the optimization problem (15) - (18). In the objective (15) the tax rates τ_{dt}^i instead of τ_{dt} have to appear as the arguments of the respective utility functions v_t^i, analogously the term $\tau_{dt}(d_{1t} + d_{2t})$ has to be substituted by $\tau_{dt}^1 d_{1t}$ + $\tau_{dt}^2 d_{2t}$ in the constraint (16). The crucial point arises with the self-selection constraint (17): which tax rate is relevant for the more able individual if he would be in the less able's position? We consider two cases, assuming again a quasilinear utility function, as introduced in Result 3:

---

but is not a first-best instrument.

1. In the first case the individuals have identical inherited wealth. Then, if the more able receives gross and net income of the less able, bequests of both are the same and there is no possibility for the authority to impose differing tax rates. As a consequence, τ_{dt}^1 has to be inserted instead of τ_{dt} on the right-hand side of (17). Straightforward derivations analogous to those in Appendix A, lead to (at $\tau_{dt}^1 = \tau_{dt}^2 = 0$)

$$\text{(24)} \qquad \frac{\partial S}{\partial \tau_{d1}^1} = \frac{1}{1+\gamma} f_1 \frac{\partial v_2^1}{\partial d_{11}} \frac{\partial d_{11}^h}{\partial \tau_{d1}^1} + \mu_1 \frac{\partial v_1^2}{\partial x_{11}} (d_{21}[1] - d_{11}),$$

$$\text{(25)} \qquad \frac{\partial S}{\partial \tau_{d1}^2} = \frac{1}{1+\gamma} f_2 \frac{\partial v_2^2}{\partial d_{21}} \frac{\partial d_{21}^h}{\partial \tau_{d1}^2} + \mu_2 \left(\frac{\partial v_2^2}{\partial d_{21}} - \frac{\partial v_2^2[1]}{\partial d_{21}}\right) \frac{\partial d_{21}^h}{\partial \tau_{d1}^2}.$$

We know that $\partial d_{i1}^h / \partial \tau_{d1} < 0$, that the expression containing μ_2 is zero due to quasilinearity of u and that $d_{21}[1] = d_{11}$ because of weak separability and equal endowments. It follows that both tax rates should be negative. The reason for this - perhaps unexpected - conclusion is familiar from Result 3: if individuals differ only in abilities, in general no other tax is desirable in addition to the income tax, not even for redistributive reasons. However, bequests create twofold utility and should be subsidized therefore.

2. In the second case the individuals have unequal inherited wealth, the more able will bequeath more, even if he is in the less able's position. This makes it possible to impose differing tax rates, which formally means that τ_{dt}^2 has to be inserted instead of τ_{dt} on the right-hand side of (17). We get (at $\tau_{dt}^1 = \tau_{dt}^2 = 0$):

$$\text{(26)} \qquad \frac{\partial S}{\partial \tau_{d1}^1} = \frac{1}{1+\gamma} f_1 \frac{\partial v_2^1}{\partial d_{11}} \frac{\partial d_{11}^h}{\partial \tau_{d1}^1} - \mu_1 \frac{\partial v_1^2}{\partial x_{11}} d_{11},$$

$$\text{(27)} \qquad \frac{\partial S}{\partial \tau_{d1}^2} = \frac{1}{1+\gamma} f_2 \frac{\partial v_2^2}{\partial d_{21}} \frac{\partial d_{21}^h}{\partial \tau_{d1}^2} + \mu_1 \frac{\partial v_1^2}{\partial x_{11}} d_{21}[1] + \mu_2 \left(\frac{\partial v_2^2}{\partial d_{21}} - \frac{\partial v_2^2[1]}{\partial d_{21}}\right) \frac{\partial d_{21}^h}{\partial \tau_{d1}^2}.$$

Now we arrive at the expected result that the tax rate should be negative for the less able and is indeterminate for the more able (where the tendency is more towards a positive rate, compared to the situation in Result 5). The fact that the individuals can be discerned with respect to a second characteristic, in addition to ability, makes it possible to employ a second tax instrument, besides the income tax, in a differentiated way.[12]

V. Concluding remarks

In this paper we have studied the role of a tax on bequests and have compared it with that of a tax on consumption goods and on interest income, in an optimal income tax system with individuals differing in earning ability. While it is known that in such a system there is no role for a consumption tax nor for an interest tax, given weak separability of leisure, a tax on bequests turns out to be desirable, if there is a strong enough motive for redistribution. The reason for this result is that with bequests an additional differentiating factor must be taken into account: More able individuals earn higher income and will thus save and bequeath more to their descendants. Then capital endowments will differ in later generations, even if they were equal at the "beginning of the world". With differing endowments, the task of redistribution cannot be performed by an income tax alone, but requires taxation of bequests.

Ideally, one would like to redistribute capital endowments in a lump-sum way, but this is impossible, because any redistributive measure is recognized by the bequeathing generation and distorts its decision. Hence, a tax on bequests represents a second-best instrument, whose adverse effects have to be weighted against its redistributive performance, as was illustrated by Result 5.

In contrast, in the usual overlapping-generations models, first-period savings are used up for consumption in the period of retirement. Every following generation starts with zero initial capital endowments (capital is in the hands of the retired), thus individuals differ only in earning abilities. In such a case and with weakly separable preferences, it suffices to redistribute via an optimal nonlinear tax on wage income (see, e. g., Ordover and Phelps 1979).

[12] It should be mentioned, however, that in principle in our model the tax authority could infer individual abilities from bequests (which are assumed discernible in case 2.) and, thus, turn to first-best taxation. But this theoretical

The model, which offered the framework for the analysis of bequest taxation, was based on a primal-dual approach. The formulation of such a model should be of value in a more general context, because it seems to represent the simplest way of studying the optimal nonlinear income tax (with its specific incentive compatibility constraint) and linear commodity taxes simultaneously.

possibility is never applied in practice.