J Popul Econ (2003) 16:37–54

Journal of Population Economics

Family size and optimal income taxation

Helmuth Cremer[1], **Arnaud Dellis**[2], **Pierre Pestieau**[3]

[1] IDEI et GREMAQ, Université de Toulouse, Allée de Brienne 21, 31000 Toulouse, France (Fax: +33-5-61128637; e-mail: helmut@cict.fr)
[2] Department of Economics, Cornell University, Ithaca, NY607255-2818, USA (e-mail: ard23@cornell.edu)
[3] CREPP, Université de Liège, CORE and Delta, 7 Bvd du Rectorat, 4000 Liège, Belgium (+32-43663106, e-mail: p.pestieau@ulg.ac.be)

Received: 4 September 2000/Accepted: 17 May 2001

Abstract. This paper studies the role of family size in the design of optimal income taxation. We consider a second best setting where the government observes the number of children and the income of the parents but not their productivity. With a *linear tax* schedule the marginal tax rate is shown to decrease with the number of children, while the relationship between the demogrant and family size appears to be ambiguous. With two ability levels, optimal *non-linear income tax* implies zero marginal tax rates for the higher ability parents; low ability parents have positive marginal tax rates that decrease with family size.

JEL classification: J13, H21, H23

Key words: Family size, optimal taxation

1. Introduction

In most countries, families with children benefit from income tax breaks and family allowances. Putting aside the objective of fostering or discouraging fertility, there are two main rationales behind these measures. One is to achieve some horizontal equity, namely, to compensate families for children-related expenses.[1] Another rationale pertains to vertical equity and more specifically poverty alleviation. Child benefits tend to have a strong effect on poverty particularly in countries where families with children have few resources. They allow for low income families with children to be kept above poverty line in many countries.[2]

All correspondence to Helmuth Cremer. We would like to thank Alessandro Cigno, Maurice Marchand and two referees for helpful comments. *Responsible editor:* Alessandro Cigno.

Table 1. Family allowances in Europe

	Family size	Children age	Family income
Germany	+	=	=
Austria	=	+	=
Belgium	+	+	=
Denmark	=	+	=
Spain	=	=	Ceiling
Finland	+	=	=
France	+	+	=
Greece	+	=	−
Ireland	+	=	=
Italy	+	=	−
Luxembourg	+	+	=
Netherlands	+	+	=
Portugal	=	=	−
United Kingdom	−	=	=
Sweden	=	=	=

Source: MISSOC (1998)
"+" means that per child benefit increases with the variable, "−" means that it decreases and "=" that it is independent.

These two objectives, vertical and horizontal equity, can conflict. Compensating high income households for having children requires a higher subsidy than compensating low income households for the same reason. Most tax breaks have that feature as they aim at restoring some horizontal fairness at the expense of vertical redistribution.

In general, child benefits are independent of income; they can be differentiated according to family size and age of children. In Table 1, the pattern of variation is summarized for countries in the European Union. In any case, comparison of benefits has to take account of the relation with the tax systems. In fact both child benefits and tax allowances are part of the income tax system which depends not only on the level of income but also on family size.

This is the view we adopt in this paper. We try to design an optimal tax system that maximizes a utilitarian social welfare function and takes into account variable family size. The setting is one of imperfect information and thus of second best maximization. The government observes each household's size and income but cannot sort out the two sources of income, earnings ability and labor supply. In such a framework, we want to see whether or not the tax schedule is affected by the number of children and if so, how. Is it through tax allowances or through child benefits? To do so, we first consider the case of a linear income tax schedule with the tax rate and the lump-sum transfer varying with the family size. We then turn to a non linear income tax.

Throughout this paper, we assume that the number of children varies across families but that it is exogenous. Endogenous fertility, more precisely fertility depending on economic variables, could have two implications: it would call for public policy aimed at affecting population growth and it would lead to complex interactions between social and tax policy on the one hand and fertility decisions on the other.[3]

Another assumption that is basic to our analysis is that children welfare depends on their parents decisions. Children are not financially autonomous and if some of them are not well treated by their parents because of lack of resources or lack of altruism, the only way they can be helped by public authorities is through their parents' decisions.

The objective of the social planner is utilitarian, namely, the sum of utilities of parents and children. As long as parents weight their children's welfare the same way as the social planner, the problem is rather standard. It becomes different if weights differ. For example, if parents weight their children's utility less than their own and if the social planner insist on equal weights, then we have a typical agency problem. With instruments limited to income taxation, we have no effective way to secure that children are well-treated. This agency problem could induce the government to resort to policies directly aimed at children. By assuming a single consumption good, we cannot rely on indirect taxation that could foster child-specific goods rather than parents-specific goods; see Cigno and Pettini (2001).

Anticipating the main results, we show that with a *linear tax* schedule, the marginal tax rate decreases with the number of children while the relationship between the demogrant and family size appears to be ambiguous. In other words, a tax break for the presence of children is theoretically better grounded than family allowances. We also show that with two ability levels optimal *non-linear income tax* implies zero marginal tax rates for the higher ability parents; low ability parents have positive marginal tax rates which decrease with family size.[4]

In deriving these results, our main concern was to stay within the tradition of optimal income taxation theory[5] and to keep the presentation rather simple. To do so, we make a number of simplifying assumptions (quasi linear preferences, independent distributions of productivities and family size, etc.). These restrictions admittedly limit the generality and surely the applicability of our conclusions.

2. The model

Consider a society consisting of parents and dependent children. A parent is characterized by a productivity level, w_i, and by a number of children n_j. The parent's utility depends on his own consumption, x, that of his children, c, and on his labor supply, L. It is given by:

$$u_{ij} = u(r_{ij} - n_j c_{ij} - h(L_{ij})) + \gamma n_j u(c_{ij}) \tag{1}$$

where ij means that the parent has an ability w_i $(i = 1, \dots, m)$ and a number n_j $(j = 1, \dots, s)$ of children. The function u is strictly concave; r_{ij} is disposable (after tax) income; h is the monetary disutility of work, a strictly convex and increasing function; and γ denotes the factor of altruism.

The parent's consumption net of the cost of effort is:[6]

$$x_{ij} = r_{ij} - n_j c_{ij} - h(L_{ij}).$$

For the time being, we assume $\gamma = 1$. For given r and L, c_{ij} is chosen so that:

$$u'(r_{ij} - n_j c_{ij} - h(L_{ij})) = u'(c_{ij}),$$

which yields

$$x_{ij} = c_{ij} = \frac{r_{ij} - h(L_{ij})}{1 + n_j}. \tag{2}$$

We can now define the "indirect" utility function:

$$v_{ij}(r_{ij}, y_{ij}) = (1 + n_j)u\left(\frac{r_{ij} - h(y_{ij}/w_i)}{(1 + n_j)}\right), \tag{3}$$

where $y_{ij} = w_i L_{ij}$ is before tax income.

The information structure is as follows. The tax administration observes the number of children, n_j, and before tax income y_{ij}. After tax income r_{ij} is then of course also observable, but its allocation between parent's consumption, x, and children's consumption, c, is private information. Finally, labor supply, L_{ij}, and ability, w_i, are not observable. The information structure thus resembles that used in traditional optimal taxation models, except that we have an *observable* source of heterogeneity, namely family size, in addition to the traditional adverse selection variable (unobservable ability). The tax function can then be conditioned on this observable variable and we have

$$r_{ij} = y_{ij} - T_j(y_{ij}), \tag{4}$$

where $T_j(y_{ij})$ is the tax schedule applied to families of size n_j. We now turn to the determination of the optimal tax schedules $T_j(y_{ij})$, $j = 1, \dots, s$. First, we shall assume that T_j is restricted to be linear. Then, we shall consider a setting in which T_j is restricted solely by the information available to the tax administration; this is the general income tax problem.

3. The optimal linear income tax

We use a linear income tax schedule specified by

$$T_j(y_{ij}) = t_j y_{ij} - a_j,$$

with marginal tax rate, t_j, and demogrant, a_j, varying with the number of children. Then, the indirect utility function (3) can be redefined as:

$$V_{ij}(t_j, a_j) = (1 + n_j)u\left(\frac{(1 - t_j)w_i L_{ij} + a_j - h(L_{ij})}{1 + n_j}\right),$$

where $L_{ij} = L_{ij}((1 - t_j)w_i)$ is the labor supply function with net (after tax) wage as sole argument; with the utility function specified by (1), labor supply does not depend on a_j. The government maximizes a utilitarian welfare function given by

$$W(t_1, \dots, t_s; a_1, \dots, a_s) = \sum_{i,j} \pi_{ij} V_{ij}(t_j, a_j), \tag{5}$$

where π_{ij} is the proportion of families of type ij in the *total* population (of parents), the size of which is normalized at one.[7] Assuming a purely redistributive tax, the revenue constraint is given by:

$$\sum_{i,j} \pi_{ij}(t_j w_i L_{ij} - a_j) = 0. \tag{6}$$

and the Lagrangean expression can be written as follows:

$$\mathscr{L}_1 = \sum_{i,j} \pi_{ij} V_{ij}(t_j, a_j) + \mu \sum_{i,j} \pi_{ij}(t_j w_i L_{ij} - a_j)$$

where μ is the multiplier associated with the revenue constraint.

Differentiating $\mathscr{L}_1$ with respect to t_j and a_j, $j = 1, \ldots, s$, yields after some manipulations these two well-known formulas:[8]

$$\sum_i \pi_{ij}\eta_{ij} = \mu \sum_i \pi_{ij}, \quad j = 1, \ldots, s; \tag{7}$$

and

$$\frac{t_j}{1 - t_j} = -\frac{\operatorname{cov}(\eta_{ij}, w_i L_{ij}) \sum_i \pi_{ij}}{\mu \sum_i \pi_{ij} w_i L_{ij} \tilde{\varepsilon}_{ij}}, \quad j = 1, \ldots, s, \tag{8}$$

where η_{ij} is the marginal utility of income and $\tilde{\varepsilon}_{ij}$ the compensated elasticity of labor (here equal to the uncompensated elasticity).

To interpret these expressions observe that we are solving an optimal linear income tax problem within each of the s different "classes" (a class being characterized by a given family size). These s problems are independent of each other *except* for the fact that there is a *global* budget constraint.[9] Consequently, it is not surprising that the marginal tax rate in each class is determined according to the traditional trade-off between redistributive benefits and efficiency cost of taxation. This is shown by the RHS of (8), with the covariance term measuring redistributive benefits, while the deadweight loss is determined by the compensated elasticity of labor supply.

Turning to (7), this condition says that the average marginal utility of income has to be the same in each of the s classes. In other words, the average marginal utility of income is equalized between different family sizes. This is not surprising. In a first best setting (complete information or when $\tilde{\varepsilon}_{ij} = 0$) the utilitarian government would equalize marginal utilities of income within and between classes. In our second-best setting, where productivities are not observable marginal utilities of income within each class are not equalized. However, since family size is observable, it is clearly desirable to adjust the a_j's so that average marginal utilities of income are the same for any family size.

To get a more precise understanding of the role that family size plays for the optimal tax policy one has to examine how t_j and a_j are affected by the number of children n_j. This relationship will of course depend on the distribution of types and specifically on the distribution of wage conditional on family size. For instance, a positive correlation between wage and family size (the proportion of high wages is larger for large families) will be a factor contributing to a higher tax (a lower level of a_j) on large families.[10] More generally, when the distributions of w and n are not independent, the observable family size can be used as a signal for the unobservable ability and this will clearly affect the structure of the tax policy.

For the remainder of the section we shall concentrate on a special case in which the distribution of wages is independent of family size. In other words, we are abstracting from the signal aspect just mentioned to examine if there is any other reason which would justify some systematic relationship between family size and the parameters of the tax function. Let us thus assume that $\pi_{ij}^j \equiv \pi_{ij} / \sum_i \pi_{ij} = \pi_i$: the distribution of earnings abilities is the same for all family types. For technical reasons it is also convenient to define an "indirect" social welfare function:

$$\widetilde{W}(t_1, \dots, t_s) = \max_{a_1 \dots a_s} \; W(t_1, \dots, t_s; a_1, \dots, a_s)$$

$$\text{s.t.} \quad \sum_{i,j} \pi_{ij}(t_j w_i L_{ij} - a_j) = 0. \tag{9}$$

In words, $\widetilde{W}$ is the maximum level of welfare that can be achieved with marginal tax rates $(t_1, \dots, t_s)$ if the demogrants $(a_1, \dots, a_s)$ are set optimally (i.e., to maximize welfare subject of the budget constraint). We now in a position to prove the following proposition.

Proposition 1. *Assume $\pi_{ij}^j \equiv \pi_{ij} / \sum_i \pi_{ij} = \pi_i$; $j = 1 \dots s$: the distribution of earnings abilities is independent of family size. Consider two different family sizes $n_k > n_l$.*

(i) *Starting from any vector of tax rates with $t_k = t_l$, a welfare improvement can be achieved through a variation in tax rates $dt_k < 0$ and $dt_l > 0$, with $dt_k \sum_i \pi_{ik} = -dt_l \sum_i \pi_{il}$*
(ii) *If we assume in addition that $\widetilde{W}(t_1, \dots, t_s)$ is concave, then (i) implies $t_k^* < t_l^*$: optimal marginal tax rates decrease with family size.*

Proof. To prove (i) we derive the welfare change associated with the variation in tax rates and which using (9) is given by:

$$d\widetilde{W} = \frac{\partial \widetilde{W}}{\partial t_k} dt_k + \frac{\partial \widetilde{W}}{\partial t_l} dt_l \tag{10}$$

Using (5) and (9) we have:

$$\frac{\partial \widetilde{W}}{\partial t_l} = -\sum_i \pi_{il} w_i L_{il} u'(x_{il}) + \mu \sum_i \pi_{il} \left(w_i L_{il} + t_l w_i \frac{\partial L_{il}}{\partial t_l} \right), \tag{11}$$

$$\frac{\partial \widetilde{W}}{\partial t_k} = -\sum_i \pi_{ik} w_i L_{ik} u'(x_{ik}) + \mu \sum_i \pi_{ik} \left(w_i L_{ik} + t_k w_i \frac{\partial L_{ik}}{\partial t_k} \right), \tag{12}$$

where

$$\mu = \sum_i \pi_i u'(x_{ik}) = \sum_i \pi_i u'(x_{il}) \tag{13}$$

Rearranging (11) and (12) yields:

$$\frac{\partial \widetilde{W}}{\partial t_l} = \left(-\operatorname{cov}(w_i L_{il}, u'(x_{il})) + \mu \sum_i \pi_i t_l w_i \frac{\partial L_{il}}{\partial t_l} \right) \sum_i \pi_{il}, \tag{14}$$

$$\frac{\partial \widetilde{W}}{\partial t_k} = \left(-\operatorname{cov}(w_i L_{ik}, u'(x_{ik})) + \mu \sum_i \pi_i t_k w_i \frac{\partial L_{ik}}{\partial t_k} \right) \sum_i \pi_{ik}. \tag{15}$$

Substituting (14) and (15) into (10), using $dt_k \sum_i \pi_{ik} = -dt_l \sum_i \pi_{il}$ and the property that $t_l = t_k$ implies $L_{il} = L_{ik} = L_i$ and $\partial L_{il}/\partial t_l = \partial L_{ik}/\partial t_k$,

$$d\widetilde{W} = \operatorname{cov}(w_i L_i, u'(x_{ik}) - u'(x_{il}))\, dt_l \sum_i \pi_{il} > 0 \tag{16}$$

if at $t_l = t_k$

$$\operatorname{cov}(w_i L_i, u'(x_{ik}) - u'(x_{il})) > 0. \tag{17}$$

To prove that (17) holds, first observe that with $t_l = t_k$, and hence $L_{il} = L_{ik} = L_i$ we obtain from (2):

$$\frac{\partial x_{il}}{\partial w_i} > \frac{\partial x_{ik}}{\partial w_i}.$$

In words, as wage increases, per-capita consumption increases at a faster rate in smaller families, l, than in larger families, k. Condition (13) then implies

$$x_{1l} < x_{1k} \quad \text{and} \quad x_{ml} > x_{mk}. \tag{18}$$

Consequently, $u'(x_{1k}) - u'(x_{1l}) > 0$, while $u'(x_{mk}) - u'(x_{ml}) < 0$; these two inequalities along with the property that $w_i L_i$ increases with wage can then easily be shown to imply (17).[11] This completes the proof of (i). Part (ii) then follows immediately from the concavity assumption. ■

Finally and in the same vein, we would have liked to show that a_j is positively related to n_j. Even though such a result is quite intuitive and is obtained in the numerical examples we conducted, it cannot be proved to always hold.

To discuss the intuition behind these results let us note two properties which can be shown from (7) and (8). First, if t_j were constrained to be constant for all j, a_j would increase with n_j. Second and conversely, if a_j were restricted to be constant, t_j would decrease with n_j. Note, however, that in the first case, a_j can do a lot towards horizontal redistribution but little towards vertical redistribution. In the second case, t_j can achieve both types of redistribution. It is thus not surprising that when combining both instruments, t_j "dominates" a_j and it can even happen that a_j decreases with n_j. To put it another way, consider a tax schedule with uniform t and variable a_j. For the average wage, one can achieve the same outcome with alternative tax schedule characterized by uniform a and a variable t_j. Moreover, with the alternative schedule, the vertical redistribution within each n_j group is better from a utilitarian viewpoint.

4. Non-linear income tax

We now turn to the non-linear income tax problem. To study this case, we assume $m = 2$: there are just two levels of productivity $w_2 > w_1$. The government's objective continues to be a simple utilitarian welfare function. Recall that family size is observable while productivity is not observable. To solve the problem we first determine the allocation which maximizes welfare subject to the resource constraint and the self-selection constraints. A complete solution of the optimal tax problem per-se then requires the design of the implementing income tax functions $T_j(y_{ij})$, $j = 1, \dots, s$.

The individuals' preferences over (r, y), that is in the space of observable variables, are given by (3); they are crucial ingredient in the non linear income tax problem. Specifically, it is useful to define:

$$\alpha_i(y) = \frac{h'\left(\dfrac{y}{w_i}\right)}{w_i}, \tag{19}$$

which represents the individual's marginal rate of substitution (slope of an indifference curve), obtained by differentiating (3). Observe that with the considered structure of preferences α depends only on y and not on r.[12] Further, α depends on the wage, but not on family size. Finally, it is easy to show that α decreases with w:

$$\alpha_2(y) < \alpha_1(y). \tag{20}$$

In other words, for any given level of y, high-wage individuals have flatter indifference curves than low-wage individuals. This corresponds to the traditional "single-crossing property".

For each n_j, the incentive compatibility constraint of high-wage workers is given by:

$$v_{2j}(r_{2j}, y_{2j}) \geq v_{2j}(r_{1j}, y_{1j}).$$

Using (3) this condition can be expressed as:

$$r_{2j} - h\left(\frac{y_{2j}}{w_2}\right) \geq r_{1j} - h\left(\frac{y_{1j}}{w_2}\right). \tag{21}$$

With a utilitarian objective function and given (20), it is easy to show that (21) will be binding within each family size class j, while the incentive constraint of the low wage type is never relevant. Observe that mimicking *between* family size classes is not possible because n is observable.

We are now in a position to state the government's problem. Let λ_j denotes the Lagrange multiplier associated with the incentive compatibility constraints and μ, that associated with the budget constraint. The Lagrangean of the social planner problem can be written as:

$$\mathscr{L}_2 = \sum_{i,j} \pi_{ij}(1+n_j)u\left(\frac{r_{ij} - h\left(\frac{y_{ij}}{w_i}\right)}{1+n_j}\right)$$
$$+ \sum_j \lambda_j \left[r_{2j} - h\left(\frac{y_{2j}}{w_2}\right) - r_{1j} + h\left(\frac{y_{1j}}{w_2}\right)\right]$$
$$+ \mu \sum_{i,j} \pi_{ij}(y_{ij} - r_{ij}).$$

Observe that, like in the linear case, we are again faced with s problems (one for each family size class) which are related only through the budget constraint.[13]

The first-order conditions are:

$$\frac{\partial \mathscr{L}_2}{\partial y_{1j}} : -\pi_{1j}u'(x_{1j})\frac{h'\left(\frac{y_{1j}}{w_1}\right)}{w_1} + \lambda_j \frac{h'\left(\frac{y_{1j}}{w_2}\right)}{w_2} + \mu\pi_{1j} = 0, \tag{22}$$

$$\frac{\partial \mathscr{L}_2}{\partial r_{1j}} : \pi_{1j}u'(x_{1j}) - \lambda_j - \mu\pi_{1j} = 0, \tag{23}$$

$$\frac{\partial \mathscr{L}_2}{\partial y_{2j}} : -\pi_{2j}u'(x_{2j})\frac{h'\left(\frac{y_{2j}}{w_2}\right)}{w_2} - \lambda_j \frac{h'\left(\frac{y_{2j}}{w_2}\right)}{w_2} + \mu\pi_{2j} = 0, \tag{24}$$

$$\frac{\partial \mathscr{L}_2}{\partial r_{2j}} : \pi_{2j}u'(x_{2j}) + \lambda_j - \mu\pi_{2j} = 0. \tag{25}$$

Before proceeding, it is useful to note that the maximization of utility (3) subject to the (after tax) budget constraint (4) yields:

$$1 - T_j'(y_{ij}) = \frac{h'\left(\frac{y_{ij}}{w_i}\right)}{w_i} = \alpha_i(y_{ij}), \tag{26}$$

where $T_j'(y_{ij})$ is the marginal income tax rate. Consequently, we can use the marginal rates of substitution determined by the first-order conditions to characterize the marginal tax rates implied by the implementing tax function.[14] We obtain the following results:

4.1. High-wage individuals

Combining (24) and (25) we obtain,

$$\frac{h'\left(\frac{y_{2j}}{w_2}\right)}{w_2} = 1, \quad \forall j,$$

which from (26) implies

$$T_j'(y_{2j}) = 0, \quad \forall j.$$

In each type of family, there is no distortion at the top; the marginal tax rate is zero for the higher ability parents. Furthermore, given our specification,

$$y_{2k} = y_{2l} \quad \forall k, l. \tag{27}$$

Higher productivity individuals have the same labor supply irrespective of family size.

4.2. Low-wage individuals: sign of marginal tax rate

Dividing (22) by (23), while rearranging and making use of (19) and (20) (for the inequality) yields:

$$\alpha_1(y_{1j}) = \frac{\alpha_2(y_{1j}) + \frac{\mu}{\lambda_j}}{1 + \frac{\mu}{\lambda_j}} < 1, \tag{28}$$

where $\alpha_2(y_{1j})$ is simply the marginal rate of substitution of the mimicking individual. The inequality in (28) comes from μ, $\lambda_j > 0$ and $\alpha_2(y_{1j}) < 1$ which, follows from $y_{ij} < y_{2j}$ and $\alpha_2(y_{2j}) = 1$. In words, whatever their family size all low wage individuals have a marginal rate of substitution which is smaller than one. From (28) and (26) we then obtain:

$$T_j'(y_{1j}) > 0, \quad \forall j. \tag{29}$$

Consequently, we have established that all low wage individuals face a positive marginal tax rate.

Summing up, we have shown that the traditional properties obtained in the two-types optimal income tax model continue to hold here *within every family size class*. Observe that no assumption on the distribution of wages (beyond $s = 2$) were necessary to establish this property. In particular, the fact that the two wage levels are the same for each family size is not necessary. Consequently, the "top" individual in, say, class j would have a zero marginal tax rate even if his wage were in fact lower than that, say, both types in class k. This is a direct implication of the information structure and specifically the assumption that the number of children is observable.

Like in the linear case, we would now like to go beyond this general characterization and examine if there is some systematic relationship between family size and (marginal) taxes. Once again, we shall look at this question for the case where the (conditional) distribution of wages is the same for all family sizes. With the marginal tax rate of the high wage type independent of family size, we are left with the comparison of the marginal tax rate faced by the low wage individual.

4.3. Low-wage individuals: marginal tax rate and family size

The main result is formally stated in the following proposition. Roughly speaking it says that when the distribution of wages is independent of family size, then the marginal tax rate of the low wage individual decreases with family size.

Proposition 2. *Assume* $\pi_{ij}^{j} \equiv \pi_{ij}/\sum_i \pi_{ij} = \pi_i$*: the distribution of earnings abilities is independent of family size. Further assume that* $h''' > 0$*. Consider two different family sizes* $n_k > n_l$*.*

(i) *The optimal utilitarian allocation, constrained by the information structure satisfies:*

$$x_{1k} > x_{1l}; \quad y_{1k} > y_{1l} \tag{30}$$

$$x_{2k} < x_{2l}; \quad y_{2k} = y_{2l}. \tag{31}$$

(ii) *The implementing tax function satisfies:*

$$T_k'(y_{1k}) < T_l'(y_{1l}); \tag{32}$$

Proof. The proof proceeds by combining a certain number of properties in order to show that all cases not satisfying (30) and (31) can be ruled out.

The relevant properties are stated in the following lemmas.

Lemma 1. *Assume that* $n_k > n_l$*. Then*

$$y_{1l} \geq y_{1k} \Rightarrow x_{2l} - x_{1l} \geq x_{2k} - x_{1k}. \tag{33}$$

Proof. Lemma 1 is a direct implication of the incentive constraint. When (21) is binding, we can write:

$$x_{2j} = \frac{r_{2j} - h\left(\dfrac{y_{2j}}{w_2}\right)}{1+n_j} = \frac{r_{1j} - h\left(\dfrac{y_{1j}}{w_2}\right)}{1+n_j}. \tag{34}$$

Therefore:

$$x_{2j} - x_{1j} = \frac{h\left(\dfrac{y_{1j}}{w_1}\right) - h\left(\dfrac{y_{1j}}{w_2}\right)}{1+n_j}. \tag{35}$$

Differentiating the RHS of (35) with respect to y_{1j} gives:

$$\frac{h'\left(\frac{y_{1j}}{w_1}\right)}{w_1} - \frac{h'\left(\frac{y_{1j}}{w_2}\right)}{w_2} > 0. \tag{36}$$

Evaluating (35) for l and k and making use of (36) then implies (33). ■

From now on we assume $\pi_{ij}^{j} \equiv \pi_{ij} / \sum_i \pi_{ij} = \pi_i$ and $h''' > 0$. Turning to the next property, we have:

Lemma 2. *Assume that $n_k > n_\ell$. Then*

$$x_{ik} \gtreqless x_{il} \Leftrightarrow y_{ik} \gtreqless y_{il}. \tag{37}$$

Proof. This relationship follows directly from the condition

$$u'(x_{1j}) \left[\frac{h'\left(\frac{y_{1j}}{w_1}\right)}{w_1} - \frac{h'\left(\frac{y_{1j}}{w_2}\right)}{w_2} \right] = \mu \left[1 - \frac{h'\left(\frac{y_{1j}}{w_2}\right)}{w_2} \right], \tag{38}$$

which in turn is obtained by combining first-order conditions (22) and (23). ■

Next, it is easy to check that $x_{1k} = x_{1l}$ and $y_{1k} = y_{1l}$ for $k \neq l$ are not possible. Consequently, we are then left with two possibilities:

(i)

$$x_{1k} > x_{1l} \quad \text{and} \quad y_{1k} > y_{1l} \tag{39}$$

(ii)

$$x_{1k} < x_{1l} \quad \text{and} \quad y_{1k} < y_{1l}. \tag{40}$$

Finally, we can eliminate case (ii) by showing:

Lemma 3. *Assume that $n_k > n_l$. Then then $x_{1k} < x_{1l}$ and $y_{1k} < y_{1l}$ cannot simultaneously hold.*

Proof. To establish Lemma 3, first note that from Lemma 1, $y_{1l} > y_{1k}$ implies

$$x_{2l} - x_{1l} > x_{2k} - x_{1k}. \tag{41}$$

Furthermore, by adding (23) and (25) we obtain

$$\pi_{1j} u'(x_{1j}) + \pi_{2j} u'(x_{2j}) - \mu(\pi_{1j} + \pi_{2j}) = 0, \quad j = 1, \ldots, s,$$

from which we obtain that

Fig. 1. Pattern of consumption levels for family sizes $n_k > n_l$. Consumption levels of both types and their difference between types decrease with family size

$$x_{1l} > x_{1k} \quad \text{implies} \quad x_{2l} < x_{2k}. \tag{42}$$

But (35) implies $x_{2k} - x_{1k} > x_{2l} - x_{1l}$ which contradicts (41). ■

We can now finalize the proof of Proposition 2. As a consequence of the above lemmas, we must have case (i). Consequently, we have established (30), which from (26) also implies (32). Finally, (31) directly results from (27) along with the incentive constraint. ■

Proposition 2 has a number of interesting implications. First, is shows that low wage families with a larger family size face a smaller marginal tax rate and have higher pre-tax income. Consequently their labor supply is also higher. Second, we obtain a pattern of consumption levels x_{ij} as represented on Fig. 1. Specifically, inequality in per capita consumption (between high and low wage families) decreases with family size. Third, the property that $x_{2k} < x_{2l}$, along with the result that $y_{2k} = y_{2l}$, shows that the tax policy does not fully compensate type k families for their larger size.

In this section, like in the previous one, we have assumed that there is no correlation between family size and productivity. One could of course obtain very different results if such a correlation were introduced. Assume for example, that there is a strong positive correlation between n and w. Then, one can no longer exclude the possibility that the tax rate increases with family size. The effect obtained in the no correlation case continues to be at work. However, family size now also acts as a signal for productivity and this effect calls for a higher tax on large families.

5. Numerical example

Let us now turn to a numerical illustration with two objectives. First, we want to contrast the results obtained in the linear and the non-linear cases. The general expressions cannot give us a good grasp of how these results differ. Second, we want to consider the possibility of non-balanced altruism.

We adopt a setting with three productivity levels ($w_1 = 10$, $w_2 = 20$ and $w_3 = 50$) and three family sizes ($n_1 = 0$, $n_2 = 1$ and $n_3 = 3$). The population is equally shared between all these groups ($\pi_{ij} = 1/9$, $\forall ij$). The objective of the government is utilitarian, that is, the social welfare function is the sum of utilities of parents and children. The social planner observes the number of children (n_j) and the parent's income (y_{ij}) but neither the productivity (w_i), nor the labor supply (L_{ij}). The parent's utility function is:

$$v_{ij} = \frac{(x_{ij})^{1-\varepsilon}}{1-\varepsilon} + \gamma n_j \frac{(c_{ij})^{1-\varepsilon}}{1-\varepsilon}, \quad \text{if } \varepsilon \neq 1,$$

$$v_{ij} = \log(x_{ij}) + \gamma n_j \log(c_{ij}), \quad \text{if } \varepsilon = 1,$$

Table 2. Optimal fiscal parameters in the linear income tax problem

	$n_1 = 0$	$n_2 = 1$	$n_3 = 3$
a_j	3.99	8.73	19.25
d_j	0	4.74	5.08
t_j	0.62	0.56	0.48
$c_{ij} = x_{ij}$	3.9; 6.11; 14.38	4.47; 5.81; 10.7	4.96; 5.78; 8.71
y_{ij}	4.89; 12.77; 38.57	5.24; 13.27; 39.36	5.59; 13.77; 40.15
T_{ij}	−0.97; 3.9; 19.82	−5.8; −1.32; 13.25	−16.53; −12.57; 0.22

where $x_{ij} = y_{ij} - T_{ij} - n_j c_{ij} - h(L_{ij})$ is the parent's consumption, $y_{ij} = w_i L_{ij}$ the gross income, T_{ij} the tax liability, c_{ij} the children's consumption, $h(L_{ij}) = 1/(1 - L_{ij})^\eta$ the monetary disutility of work and γ the factor of altruism. For the time being, we assume $\varepsilon = \eta = \gamma = 1$.

We examine how the tax schedule is affected by the number of children. First, we look at the linear income tax problem. The tax liability is given by:

$$T_{ij} = t_j y_{ij} - a_j$$

where t_j is the linear tax rate and a_j the demogrant. Observe that this demogrant cannot directly be interpreted as a "family allowance". To obtain a measure of the child benefits implied by the tax system, one has to take into account the fact that even childless families may have a demogrant. The allowance per child can then be defined as:

$$d_j = \frac{a_j - a_1}{n_j}$$

where a_1 is the demogrant received by households having no children.

Results are reported in Table 2. The demogrant and the child benefit are increasing with the family size while the tax rate is a decreasing function. Broadly speaking, there is a transfer going from households with high income but no children to poor families with children. So both instruments, namely income taxation and child benefits, contribute to redistribute income horizontally as well as vertically. Another basic finding is that the child cost (c_{ij}) increases with the family income. Therefore, the child benefits more than compensate for the child cost in poor families, but not in rich ones. This is an consequence of our utilitarian objective function along with the rigidity of our tax instruments. The government is utilitarian and families as well. In rich families, children are well-treated and one cannot expect family allowances to fully finance their consumption.

Similar conclusions are drawn in the non-linear case (Table 3): (i) the marginal tax rate (T'_{ij}) is a decreasing function of the family size; (ii) the tax liability (T_{ij}) increases with the productivity level and decreases with the family size; (iii) the higher the household earnings are, the higher the child cost is. We also find the Mirrlees (1971) result according to which there is no distortion at the top ($T'_{ij} = 0$ for high ability households). Compared to the linear case, one observes that now the tax burden is more redistributive. The tax liability for a poor and large family relative to that for a rich and large family is – 23–5

Table 3. Optimal non-linear income tax problem

	$n_1 = 0$	$n_2 = 1$	$n_3 = 3$
$w_1 = 10$			
$c_{ij} = x_{ij}$	5.69	6.27	6.66
y_{ij}	6.51	6.57	6.74
T_{ij}	−2.04	−8.88	−22.95
T'_{ij}	0.18	0.11	0.06
$w_2 = 20$			
$c_{ij} = x_{ij}$	7.07	7.02	7.05
y_{ij}	15.01	15.22	15.36
T_{ij}	3.93	−3.00	−17.14
T'_{ij}	0.20	0.12	0.07
$w_3 = 50$			
$c_{ij} = x_{ij}$	9.65	8.39	7.76
y_{ij}	42.93	42.93	42.93
T_{ij}	26.21	19.07	4.81
T'_{ij}	0	0	0

(compared to – 17–0). When contrasting the two extremes, poor large family and rich small family, one has – 23–26 (compared to – 17–20).

We use this numerical example to explore the case of selfish parents and thus the question of differential altruism. It is possible that parent's altruism is lower than what the social planner would like. For instance, parents could weight their children's welfare less than their own while the government insists on equal weights ($\gamma < 1$). We examine how the tax schedule is affected by this divergence.

There are two effects acting in opposite directions. On the one hand, the government would like to give more to large families to compensate children for the parent's lack of altruism. But, on the other hand, the major beneficiaries are selfish parents, not children. With a logarithmic utility function, both effects cancel. With an isoelastic function and a low ε (we have here assumed $\varepsilon = 1/2$), the second effect overwhelms the first one (see Table 4).[15] More precisely, the demogrant is larger for households without children but smaller for large families, while the tax rate is respectively lower and higher than when parents are as altruistic as the social planner. Family allowances remain increasing and the tax rate decreasing with the family size, but at a slower pace than when there is no difference in altruism. We should keep in mind that the less altruistic parents are, the more they consume relative to their children. Finally, with a high elasticity ($\varepsilon = 3/2$), the profile of tax rates and family allowances for increasing n is relatively independent of the value of γ.

One of the interesting implications of this example is that it underlines the inadequacy of linear income taxation to cope with such a merit-good problem. More effective policies could be introduced in (at least) two ways. The first possibility considers explicitly a consumption *vector* consisting of several goods. This opens the door to possible subsidies for child-specific commodities like in Cigno and Pettini (2001). Alternatively, one could resort to policies targeted towards children. In future research we plan to study this agency issue wherein the government could try to elicit parents' altruism and to provide when needed specific services to the children.

Table 4. Linear income tax problem with altruism differing between parents and the social planner

	$\varepsilon = 1/2$		
	$n_1 = 0$	$n_2 = 1$	$n_3 = 3$
$\gamma = 1$			
a_j	2.98	7.40	17.65
d_j	0	4.42	4.89
t_j	0.5447	0.4750	0.3929
$\gamma = 0.1$			
a_j	6.64	8.38	12.48
d_j	0	1.74	1.95
t_j	0.4841	0.4641	0.4275
	$\varepsilon = 3/2$		
	$n_1 = 0$	$n_2 = 1$	$n_3 = 3$
$\gamma = 1$			
a_j	4.48	9.37	19.96
d_j	0	4.89	5.16
t_j	0.6529	0.6029	0.5368
$\gamma = 0.1$			
a_j	4.01	9.65	20.18
d_j	0	5.64	5.39
t_j	0.6595	0.6005	0.5358

6. Conclusion

In this paper we have studied the optimal income taxation with different family sizes and ability levels. We have taken into account the informational problem to which the social planner is confronted; it cannot observe the sources of income, namely the innate ability or the labor supply.

Under this setting, we have shown that the income linear tax rate should decrease with the family size. With a non-linear tax schedule, we have found the well-known result that there should be no distortion at the top. At lower ability levels, the marginal tax rate should decrease with the number of children and increase with the productivity.

The numerical example shows that fiscal parameters are pretty independent of the parent's altruism. In this model, the government can help children only through their parent's decision. So the major beneficiaries of tax cuts for large families and child benefits are selfish parents, not children.

Endnotes

[1] See Balcer and Sadka (1986), Balestrino (1994), Cremer et al. (1999) and Kaplow (1992).

[2] See Delhausse et al. (1998).

[3] See Carrin (1982), Cigno (1983), Cigno (1986) and Cigno and Pettini (1999) for the case where fertility is endogenous. Balestrino (1998) and Cigno (1996) have studied endogenous as well as exogenous fertility cases.

[4] Balcer and Sadka (1986) have drawn the same conclusion under the condition of strict horizontal equity and utilitarianism.
[5] In that respect, our approach is similar to that of Blomqvist and Horn (1984) who consider a problem of health insurance and assume that the tax parameters could be state-dependent.
[6] Preferences over x and L are quasi-linear. This specification is used for instance by Diamond (1988). It appears to represent a good compromise between simplicity and realism. It implies that there is no income effect in labor supply. In the problem at hand, it also implies that family size has no incidence on labor supply, which is questionable. There is some evidence that labor supply is affected by the size and the structure of the family. However, this occurs for reasons which are not income-related but associated with features not considered here.
[7] This is different from the objective of horizontal equity studied by Balcer and Sadka (1986).
[8] Atkinson and Stiglitz (1980).
[9] Formally, this is very much like a setting in which we would have to design linear income tax functions in s different countries, with the possibility of making transfers between the countries (global budget constraint). The analogy is not perfect, though, because family size also affects preferences in a specific way.
[10] It does not necessarily mean that the marginal tax rate is high, though. Consider an extreme case in which all families with the largest size, n_s, have an identical and large productivity. In that case, since there is no heterogeneity within the class, the marginal tax rate will be zero.
[11] This is most obvious when $m = 2$ (there are only two wage levels). When $m > 2$ a few tedious but straightforward steps are required; the complete argument is available from the authors on request.
[12] All indifference curves of a considered individual are vertically parallel to each other in the (y, r) plane.
[13] Recall that n_j is observable. Consequently, there are no incentive constraints involving individuals with differ in family size.
[14] Subject to the usual caveats regarding the non-differentiability of the implementing tax function in a two group model; see Stiglitz (1987). Where this problem arises we *define*

$$1 - \frac{h'\left(\frac{y_{ij}}{w_i}\right)}{w_i},$$

as the individuals marginal tax rates.
[15] Only results from the linear income tax problem are reported. But same conclusions can be drawn from the non-linear income tax problem.

References

Atkinson AB (1995) On targeting social security: theory and western experience with family benefits. In: van de Walle D, Nead K (eds) *Public Spending and the Poor*, The Johns Hopkins University Press, Baltimore, Md

Atkinson A, Bourguignon F (1989) The design of direct taxation and family benefits (unpublished)

Atkinson A, Stiglitz J (1980) *Lectures on Public Economics*. McGraw-Hill, London

Balcer Y, Sadka E (1986) Equivalence scales, horizontal equity and optimal taxation under utilitarianism. *Journal of Public Economics* 29:79–97

Balestrino A (1998) Does horizontal equity call for subsidies to large families? University of Pisa

Barro R, Becker G (1989) Fertility choice in a model of economic growth. *Econometrica* 57:481–501

Becker G (1981) *A Treatise on the Family*. Harvard University Press, Cambridge, MA

Blomqvist A, Horn H (1984) Public health insurance and optimal income taxation. *Journal of Public Economics* 24:353–371

Carrin G (1982) Optimal family allowances in a simple second-best model. *Public Finance* 37:39–49

Cigno A (1983) On optimal family allowances. *Oxford Economic Papers* 35:13–22

Cigno A (1986) Fertility and the tax-benefit system: a reconsideration of the theory of family taxation. *Economic Journal* 96:1035–1051

Cigno A (1996) Cost of children, parental decisions and family policy. *Labour* 10:461–474

Cigno A, Pettini A (1999) Traitement fiscal des familles quand la fécondité est endogène. *Actualité Economique* 15:239–252

Cigno A, Pettini A (2001) Taxing family size and subsidizing child specific commodities. *Journal of Public Economics* (forthcoming)

Cremer H, Dellis A, Pestieau P (1998) Prestations familiales et taxation optimale. *Economie Publique* 2–3, 145–160

Delhausse B, Dellis A, Pestieau P (1998) Family allowances and poverty in the European Union. *Cahiers de recherche CREPP*

Diamond P (1998) Optimal income taxation. An example with a U-shaped pattern of optimal marginal tax rates. *American Economic Review* 88:83–95

Kaplow L (1992) Optimal distribution and the taxation of the family. *National Bureau of Economic Research Working Paper* 4189

Mirrlees J (1971) An exploration in the theory of optimal income taxation. *Review of Economic Studies* 38:175–208

Mirrlees J (1972) Population policy and the taxation of family size. *Journal of Public Economics* 1:169–198

MISSOC (1998) *Social Protection in the Member States of the EU*. Office des Publications Officielles des Communautés Européennes, Luxembourg

Sinn H-W (1997) The value of children and immigrants in a pay-as-you-go pension system: a proposal for a partial transition to a funded system. *Journal of Population Economics* (forthcoming)