CENTRE FOR HOUSEHOLD, INCOME, LABOUR AND DEMOGRAPHIC ECONOMICS

Working Papers

Agency in family policy: a survey.

A. Cigno

ChilD n. 16/2009

e-mail: de-child@unito.it
Web site: http://www.child-centre.it

Agency in family policy: a survey.*

Alessandro Cigno
Università di Firenze, CESifo, CHILD and IZA

Abstract

Given that young children are under the control of their parents, if the government has an interest in either the welfare or the productivity of the former, it has no option but to act through the latter. Parents are, in the ordinary sense of the word, the government's agents. They are agents also in the sense of Principal-Agent theory if the parental action of concern to the government is private information. This throws doubt on some established optimal-taxation results, and gives rise to some new ones.

Key-words: optimal taxation, optimal family allowances, hidden ability to raise children, hidden educational investments, endogenous and exogenous fertility.

JEL: D13, D82, H24, H31, J13, J24.

*Comments by Alessandro Balestrino, Guy Laroque, Pierre Pestieau and Ray Rees are gratefully acknowledged. Remaining errors are the author's.

1 Introduction

This paper is about the analytical and substantive issues which arise in the fiscal treatment of families, as opposed to singles or childless couples. One has to do with the fact that young children are under the control of their parents. If the government has an interest in either the welfare or the productivity of the former, it then has no option but to act through the latter. Parents are, in the ordinary sense of the word, the government's agents. They are agents also in the technical sense of Principal-Agent theory if the parental action of concern to the government is private information. Another has to do with the fact that the children are domestically produced goods in the sense of Becker (1991). As these goods are not tradeable, a couple's ability to raise children is then as relevant a differentiating characteristic as a worker's ability to produce income. This raises questions of division of labour between couples.[1] It also makes parental decisions separable into a problem of time allocation between domestic and market activities, and a problem of income allocation between expenditure on parents, and expenditure on children. As we will see, separability is of some importance in the design of optimal policy. Yet another issue has to do with the fact that the parental actions and characteristics which affect a child's future income are typically more difficult to observe by a public authority than those which affect the agent's own income. Therefore, the potentiality or actual presence or young children changes the information structure of the policy problem.

The incorporation of asymmetric information in economic models is probably the most important theoretical development of the past forty years. Of particular interest to us here is its incorporation into the theory of optimal taxation at the hands of Mirrlees (1971, 1974), and many others in his wake. As the observable outcome of concern to the government depends not only on an unobservable action undertaken by the agents, but also on other unobservable factors, this theory may be seen as an application of the Principal-Agent model. In one formulation, the unobservable factors are the agent's personal characteristics. Suppose, for example, that the government observes how much a worker earns, but not how much he works. If the government knew the worker's earning ability, it could deduce the amount of time worked from the amount of money earned. As it does not, however, all the government can do is use its statistical information regarding the distribution of this trait in the population of workers to device a tax-subsidy scheme which will in-

[1] It also raises questions of division of labour *within* couples. Given our focus on parent-child relations, however, we shall not be concerned with that.

duce the worker in question to reveal his own earning ability. In most of the existing models, the only hidden characteristic is indeed the agent's earning ability. Variations on the theme can be found in Bovemberg and Jacobs (2005), and Hellwig (2008), who examine the case in which there is still only one hidden characteristic, but this characteristic affects *either* the earning ability of an agent with a given education, *or* the cost to the agent of acquiring that education. Balestrino, Cigno and Pettini (2002, 2003), and Cigno, Luporini. and Pettini (2004), appear to be the first to have extended the analysis to the case where the agents are couples differentiated by two hidden characteristics, namely their earning ability, *and* their ability to raise children. This extension gives rise to important interactions between informational and comparative-advantage considerations in the design of policy.

In another formulation of optimal taxation theory, the unobservable factors are random variables with known probability distributions conditional on the agent's hidden actions, and the policy problem is then to induce the agent to set those actions at the levels deemed optimal by the policy maker. In most of the existing contributions, the only hidden action is again the amount of labour supplied by the agent, and the observable outcome are the agent's earnings. In Cigno, Luporini and Pettini (2003), by contrast, the hidden action is the amount of money or own time that a couple invests in a child, and the observable outcome is the child's (rather than the agent's) lifetime earning and tax-paying capacity. In principle, there is also another hidden action, reproductive activity. Like most of the endogenous-fertility literature, however, the article in question assumes that a couple can deterministically choose how many children to have. As the number of children is observable, the government can then get potential parents to have whatever feasible number of children it deems optimal by imposing what, in the Principal-Agent literature, is known as a "forcing contract" – in plain English, by threatening parents with a large enough penalty if they have a number of children different from the one prescribed.[2] This unsavoury, as well as unrealistic, feature disappears if the number of children also is treated as a random variable with known probability distribution conditional on the agent's unobservable reproductive activity as in Cigno and Luporini (2006, 2009). The same treatment is reserved by Cremer, Gahvari and Pestieau (2006, 2008) to the sum-total of the amounts earned by all the agent's children.[3] Either way, if the conditioning action is private

[2]Of course, the assumption that any type of behaviour can be secured by a large enough penalty is only a convenient analytical simplification. Homicide attracts very severe penalties, but is not entirely deterred by them.

[3]This implicitly assumes that the probability distribution of the number of chil-

information, the government cannot use a forcing contract, and must consequently give would-be parents the incentive to undertake the action in question at the socially optimal level. As the properties of the second-best policy do not change very much compared with the case where a couple can deterministically choose how many children to have, however, we will concentrate on the latter.

2 Comparative advantage and the direction of redistribution

Before tackling asymmetric information, it is useful to set out the welfare implications of assuming that some of the goods entering the utility function of the household members are domestically produced, rather than bought from the market.[4] While obvious once stated, these implications do not appear to be universally appreciated.

Let the utility of household i be given by

$$U^i = U\left(X^i, Y^i\right), \tag{1}$$

where X^i is a domestically produced good, and Y^i a good bought from the market. The utility function $U(.,.)$, the same for all households, is increasing and concave. The quantity of domestically produced good is $X^i = F\left(H^i, Z^i; \theta^i\right)$, where H^i denotes the amount of own time, and Z^i the quantity of a market commodity, used as inputs. The production function $F(.,.)$ is increasing and concave. The positive constant θ^i is a domestic productivity parameter. It seems natural to assume that both H^i and Z^i are essential to production, so that $F\left(0, Z^i; \theta^i\right) \equiv F\left(H^i, 0; \theta^i\right) \equiv 0$.

Household income is given by

$$W^i = \omega^i L^i, \tag{2}$$

where L^i is the labour supply, and ω^i the wage rate, of household i. Each household is endowed with one unit of time, so that

$$H^i + L^i = 1. \tag{3}$$

Suppose that there are only two, equally numerous, categories of households. We can then conduct the analysis in terms of the representative household of each category, labelled $i = 1, 2$. The two can differ in the value of θ, in that of ω, or in both. Having assumed that

dren, and the probability distribution of each child's future earnings, are conditional on the same parental action.

[4]This section draws on Cigno (2001).

one of the goods is domestically produced, we cannot allow for the time of one household to be used in another household, otherwise we might just as well have assumed that the good is produced by firms. If we assume that the domestic product of one household can be consumed by another household, we characterize a Pareto optimum by maximizing $U(X^1, Y^1)$, subject to

$$U\left(X^2, Y^2\right) \geq U^i, \tag{4}$$

where U^i is a parameter,

$$Y^1 + Y^2 + \left(Z^1 + Z^2\right) p = \left(1 - H^1\right) \omega^1 + \left(1 - H^2\right) \omega^2, \tag{5}$$

where p is the price of the commodity used as an input into domestic production, and

$$X^1 + X^2 = F\left(H^1, Z^1; \theta^1\right) + F\left(H^2, Z^2; \theta^2\right). \tag{6}$$

The precise form in which θ^i enters the domestic production function makes a difference to the analytical results, but does not affect the general drift of the argument. For explicitness, we introduce θ^i as a factor that augments the effectiveness of time employed in domestic production (in the same way as ω^i augments the effectiveness of time employed in the production of income),

$$F\left(H^i, Z^i; \theta^i\right) \equiv X\left(\theta^i H^i, Z^i\right). \tag{7}$$

A Pareto-optimal $(H^1, H^2, X^1, X^2, Y^1, Y^2, Z^1, Z^2)$ then satisfies

$$\frac{\theta^i X_H\left(\theta^i H^i, Z^i\right)}{X_Z\left(\theta^i H^i, Z^i\right)} = \frac{\omega^i}{p} \tag{8}$$

for $i = 1, 2$, and

$$\frac{U_X\left(X^1, Y^1\right)}{U_Y\left(X^1, Y^1\right)} = \frac{U_X\left(X^2, Y^2\right)}{U_Y\left(X^2, Y^2\right)}. \tag{9}$$

The first of these conditions tells us that, in each type of household, the MRTS of household time for the commodity used as an input in domestic production must be equated to the relative price. The second says that the MRS of the market good for the domestically produced good must be the same in all households.

The assumption that the domestically produced good can be moved from one household to another suggests that this is something like a soufflé. Some people are better than others at baking it (have a higher θ), but anyone can enjoy it. That is the assumption in Sandmo (1990). If there is a perfectly competitive market for soufflés, the *laissez-faire* equilibrium is then a Pareto optimum. What if we are talking of children, rather than soufflés? Think of X^i as the number of children in a type-i household, or as some measure of their aggregate welfare. Then, X^i cannot be moved from one household to another. That is the assumption in Cigno (2001), and Balestrino, Cigno and Pettini (2003).

Let us then replace (6) with

$$X^i = X\left(\theta^i H^i, Z^i\right), \; i = 1, 2. \tag{10}$$

A Pareto-optimal allocation will again satisfy (8) for $i = 1, 2$. But, instead of (9), it will now satisfy

$$X_Z\left(\theta^1 H^1, Z^1\right) \frac{U_X\left(X^1, Y^1\right)}{U_Y\left(X^1, Y^1\right)} = X_Z\left(\theta^2 H^2, Z^2\right) \frac{U_X\left(X^2, Y^2\right)}{U_Y\left(X^2, Y^2\right)}. \tag{11}$$

The latter reduces to (9) if and only if

$$X_Z\left(\theta^1 H^1, Z^1\right) = X_Z\left(\theta^2 H^2, Z^2\right). \tag{12}$$

The intuition is straightforward. If the domestically produced good can be moved about, the marginal product of the commodity used as an input in its production will be equalized across households, and (12) will always be true. If the domestic product cannot be moved, (12) will be true only for some $\left(\theta^1, \theta^2, \omega^1, \omega^2, p\right)$ configuration.

We shall refer to the allocation that satisfies $(8) - (9)$ as a Pareto optimum, and to that which satisfies (8) and (11) as a constrained Pareto optimum. Since all households buy the market input at the same price p, (12) implies that money must have the same return in every household. The government can then induce a Pareto optimum by lump-sum redistribution.

Let M^i denote a lump-sum money transfer to household i. Suppose the government chooses (M^1, M^2) to maximize the Benthamite social welfare function

$$SWF = U\left(X\left(\theta^1 H^1, Z^1\right), Y^1\right) + U\left(X\left(\theta^2 H^2, Z^2\right), Y^2\right), \tag{13}$$

subject to a household budget constraint for each type of household,

$$Y^i + Z^i p = \left(1 - H^i\right)\omega^i + M^i, \; i = 1, 2, \tag{14}$$

and to the government budget constraint,

$$M^1 + M^2 = 0. \tag{15}$$

The solution satisfies $(8) - (9)$. Of all constrained Pareto optima, the one that is also a Pareto optimum is thus a Benthamite social optimum.

If we carry out the same optimization with $M^1 = M^2 = 0$, we get the *laissez-faire* solution. Let us compare this with the social optimum for various wage and domestic ability configurations.

In the standard income taxation problem, households are differentiated only by their wage rate. As domestic ability is not mentioned, it must be assumed that it is the same for all households. Let us then start by assuming

$$\theta^1 = \theta^2, \ \omega^1 < \omega^2. \tag{16}$$

This implies that type-1 households have a *comparative advantage* in domestic production,

$$\frac{\theta^1}{\theta^2} > \frac{\omega^1}{\omega^2}. \tag{17}$$

The case is illustrated in Figure 1.

The continuous concave-to-the-origin curve through point **SO** is the locus of Pareto optima. In view of (17), this curve is asymmetric around the 45° line. Since type 1 has the comparative advantage, the hump is on the right. The dotted concave-to-the-origin curve through points **LF**° and **SO** is the locus of constrained Pareto optima.[5] The straight line through point **SO**, with slope equal to minus one, is a Benthamite social indifference curve. The point **SO**, belonging to all three curves, is the social optimum. Point **LF**°, on the constrained Pareto frontier, is the *laissez-faire* equilibrium. Since the two household types have the same θ, the one with the higher ω has obviously the higher *laissez-faire* utility. **LF**° will then lie on the left of the 45° line. The socially optimal policy involves redistribution in favour of low-wage households.

The same figure may be used to illustrate also the case where

$$\theta^1 > \theta^2, \ \omega^1 < \omega^2. \tag{18}$$

Since (17) is still true, the shape of the frontiers has not changed, but either type could now have the higher *laissez-faire* utility.[6] If θ^1 is sufficiently higher than θ^2 to more than compensate for the fact that ω^1 is lower than θ^2, the point representing the *laissez-faire* equilibrium will be on the right of the 45° line. The picture is drawn under the assumption that θ^1 is large enough to place this point, labelled **LF'**, not only to the right of the 45° line, but also to the right of **SO**. The socially optimal policy now redistributes in favour of high-wage households! But this is not the only possibility. **LF'** could be on the right of the 45° line, but

[5]Be careful not to mistake this for a second-best frontier.

[6]In trade theory, only comparative advantage matters. In welfare analysis, absolute advantage matters too.

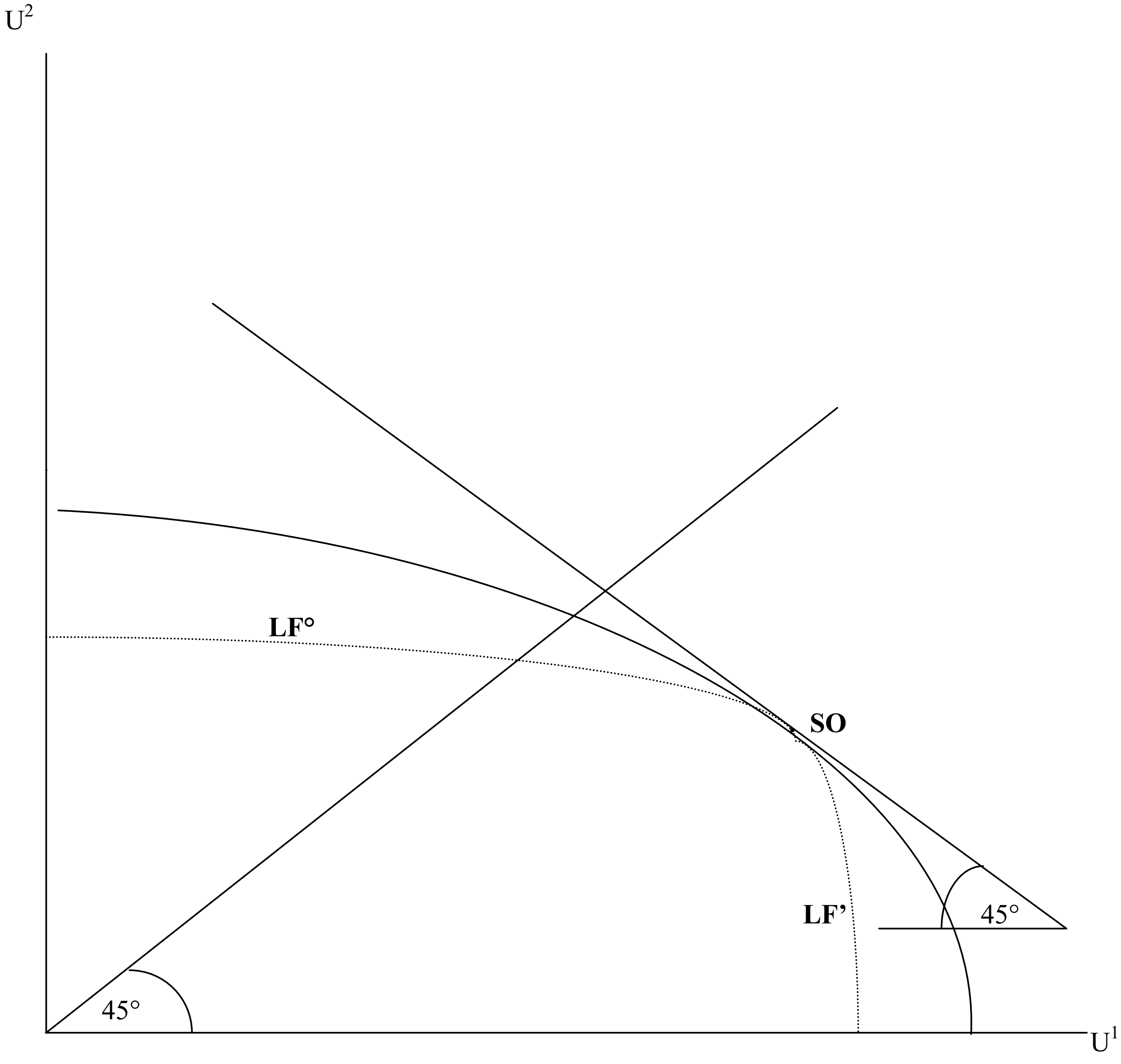

Figure 1

on the left of **SO**, in which case redistribution would be in favour of low-wage households.

Finally, let us look at the case where one category of households is better at both activities,

$$\theta^1 < \theta^2, \ \omega^1 < \omega^2. \tag{19}$$

Now, the low-wage type has clearly the lower *laissez-faire* utility, but either type could have the comparative advantage in domestic production. Figure 2 is drawn under the assumption that the high-wage type has this advantage,

$$\frac{\theta^1}{\theta^2} < \frac{\omega^1}{\omega^2}. \tag{20}$$

If that is the case, the hump in the two frontiers, and the point **SO**, are on the left of the 45° line as pictured. Point **LF** also will be on the left of the 45° line, but it could be on either side of **SO**. The picture is drawn under the assumption that type-2 households are so much better at everything, that **LF** is on the left of **SO**. Redistribution is then in favour of low-wage households. But this is not the only possibility. If **LF** were on the right of **SO**, implying that type 2 is not so much cleverer than type 1, redistribution would be in favour of high-wage households in spite of the fact that they have higher *laissez-faire* utility than low-wage households.

These exercises make the point that optimal redistribution could have the effect of allowing households with a comparative advantage in domestic production to specialize further in that activity as well as, or rather than, the effect of reducing inequality. This conclusion requires some qualification if society dislikes utility inequality, for in that case the social optimum will be a point between the Benthamite social optimum, represented in figures 1 and 2 by **SO**, and the 45° line. If inequality aversion is very strong (for example, in the extreme Rawlsian case where the social optimum is necessarily a point on the 45° line), this may actually change the direction of the redistribution. Otherwise, efficiency considerations will prevail.

Informational asymmetries may prevent the government from carrying out the socially optimal redistribution, and to this we now turn.

3 Inducing parents to reveal their ability to raise children

Let us now introduce asymmetric information about household characteristics. Suppose that the government has statistical information about the distribution of these characteristics in the population of households,

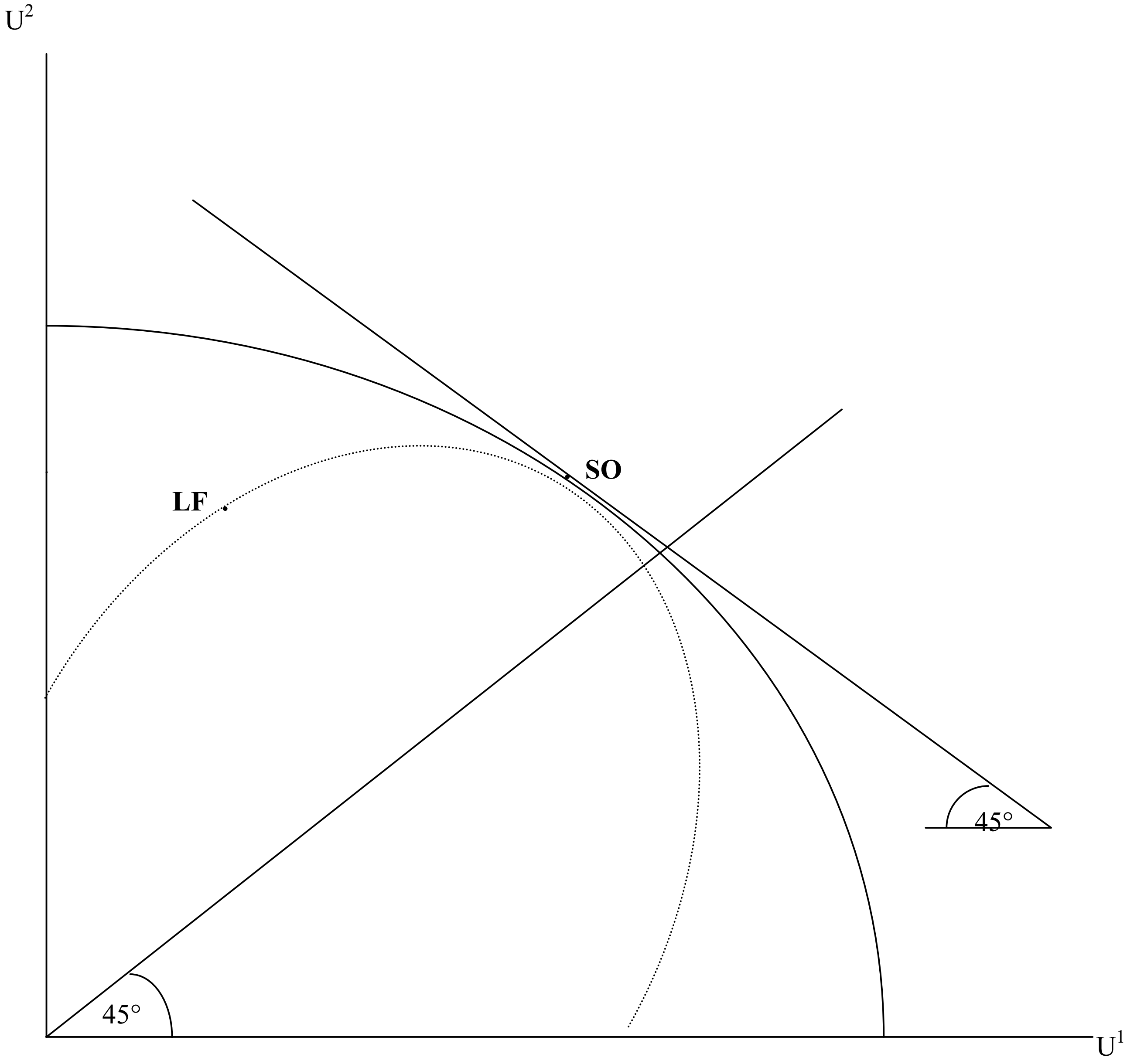

Figure 2

but does not know the characteristics of any particular household. The solution proposed by Mirrlees (1971) is to induce households to reveal their characteristics by offering them a list of alternative fiscal menus, one for each household type (combination of household characteristics). In a popular formulation of this theory, households are differentiated only by their wage rate, and the government observes only their earnings (not their labour supplies). Each menu then specifies the level of household earnings before and after income-tax, and the set of alternative menus implies a non-linear income tax schedule.

According to this approach, the government sets up the list of alternative fiscal treatments so as to make social welfare, defined as the sum of the household utilities, as high as possible within the limits imposed by the government's own budget constraint, and by the self-selection constraint that a household of type i must end up worse-off if it picks the menu intended for type j. If any of these self-selection constraints is binding, the solution to the policy optimization is a second best. As we saw in the last section, if households are only differentiated by their wage rate, the government will want to redistribute from high to low wage households. If that is the case, no household will have an interest in being taken for a high earner. An implication of the model is thus that the marginal rate of income tax implicitly charged to top earners will be zero ("no distortion at the top"). By contrast, the marginal income-tax rate charged to households with wage rates below the top will have to be positive in order to discourage mimicking.

Under these assumptions, mimicking requires only adjusting the labour supply so that the amount earned by the mimicker is equal to that earned by the mimicked. But suppose that fertility is a choice variable, and that the government can see how many children a household has. Then, the mimicker must reproduce not only the income, but also the fertility level of the mimicked. This makes mimicking harder. Another way in which children can make a difference to policy design has to do with the direction of redistribution. As we saw in the last section, if households are differentiated by their ability to raise children as well as by their ability to make money, the government may want to redistribute from low to high wage households. As low-wage households are then potential mimickers, the no-distortion-at-the-top-property need not apply. Furthermore, it may be optimal for the government to use a policy instrument related to the number of children in addition to, or instead of a non-linear income tax. The last two propositions apply whether fertility is endogenous or exogenous. Balestrino, Cigno and Pettini (2002) examine the first case, and allow for differences in ability to raise children. Cremer, Dellis and Pestieau (2003) examine the second, but do not allow for differences in ability to raise children. In essence, the assumption that fertility is endogenous entails a less benign fiscal treatment of larger families than the one that fertility is exogenous. It may also make a difference where the use of indirect taxation is concerned. Distorting prices by differential commodity taxation has in fact a deadweight cost, but may raise welfare all the same if it raises revenue, or relaxes the self-selection constraints. The positive effect may be stronger if fertility is endogenous, than if it is exogenous.

For the simple case where households are differentiated by only one characteristic, and this characteristic is the wage rate, Atkinson and Stiglitz (1976) show that indirect taxation is redundant if the utility function is separable in consumption and leisure, and income is optimally taxed. Kaplow (2006) shows that indirect taxation[7] is undesirable even if the income tax is not optimal.[8] What if the differentiating characteristic is a personal trait affecting either the earning ability of an agent with given education, or the cost of obtaining a given level of education, and education is a choice variable? In the first case, Bovemberg and Jacobs (2005) find that it is optimal to distort labour decisions through an income tax, and expenditure decisions through an education subsidy. In the second, Hellwig (2008) finds that there should not be a distortionary tax on income, but only a distortionary tax on education. The last author finds it "bothersome ... that such substantive conclusions should depend on ... whether a person's 'type' affects the person's productivity or the person's cost of attaining a given education level. In general, an agent's 'type' should be a multi-dimensional characteristic ...". But that is precisely how it is in Cigno (2001), Balestrino, Cigno and Pettini (2002, 2003), and Cigno, Luporini and Pettini (2004), where households are differentiated not only by their ability to make money, but also by their ability to domestically produce goods. In particular, Balestrino, Cigno and Pettini (2003) find that the redundancy-of-indirect-taxation result does not hold if the goods domestically produced in one household are not transferable to another. The analysis that follows draws on this article.

[7] By this we mean taxation which alters the relative prices of traded goods. Uniform taxation of traded goods is equivalent to income taxation.

[8] Cigno and Pettini (2003) find that, if an income tax is not available, but the policy maker can distinguish between goods consumed primarily by adults, and goods consumed primarily by children, it may be optimal to tax the number of children, and subsidize the goods they consume.

3.1 Households

Let household preferences be described by a concave utility function,

$$U = U(N, x, Y,), \tag{21}$$

where Y is again adult (parental) consumption, and N the number of children. The second argument, x, may be alternatively interpreted as a domestically produced commodity consumed by children (Becker, 1991), as the maximized lifetime utility of each child (Becker and Barro, 1988), or as the old-age support that parents expect to receive from each of their children under some self-enforcing family arrangement (Cigno, 1993, 2006). If we define

$$X \equiv Nx,$$

(21) is a generalization of (1).

Whichever its interpretation, x will depend on the amount of child-specific commodities, z, and parental time, h, received by each child,

$$x = x(z, h; \theta), \tag{22}$$

where $x(., .; .)$ is increasing and concave, and θ is again a productivity parameter, this time representing parental ability to rear children. The function $x(.; \theta)$ is taken to be homogeneous of degree one in $(z - z_0, h - h_0)$, with $x(z_0, h_0; k) \equiv 0$, where (z_0, h_0) are the minimum levels of z and h necessary to bring a child into the world, and keep him or her alive. In other words, there are constant returns to scale in domestic production above a certain threshold.

Gross parental income is again given by (2). Since the market input is bought anonymously, the tax or subsidy on this commodity has to be the same for everybody. By contrast, since the number of children in each household is known, it may be optimal for the government to set a different fertility subsidy or tax rate for each household type. The household budget constraint is

$$Y + (1 + t_Z)Z + t_N N = B, \tag{23}$$

where $Z \equiv Nz$, t_Z is an excise tax on child-specific commodities, and t_N a tax on the number of children. B is parental income net of income tax. Defining $H \equiv Nh$, the time budget is again (3). The government observes (B, N, W), but not (h, x, z), and has only statistical information about the distribution of (θ, ω).

The government offers households a list of alternative (B, W, t_N) triplets, one for each household type, from which to choose, and fixes a common t_Z. Parents take their decisions in two stages. First, they find the (H, L, N, X, Y, Z) that maximizes household utility conditionally on (B, W, t_N, t_Z). Second, they effectively declare to be of a certain type by choosing a (B, W, t_N) triplet from those on offer.

A household choosing the (B, W, t_N) that was intended for its type maximizes (21), subject to (3) – (23). The solution satisfies

$$\begin{array}{c} U_Y = \alpha; \quad U_x x_z = \alpha N \left(1 + t_Z\right); \\ U_N = \alpha \left[\left(1 + t_Z\right) z + t_N\right] + \beta h; \quad U_x x_h = \beta N, \end{array} \tag{24}$$

where α and β are the indirect marginal utilities of , respectively, B and h. Solving (24) together with the constraints yields the household demand for (N, X, Y, Z) as a function of (B, W, t_N, t_Z). Substituting back into the utility function, this gives us the indirect utility function, $V(B, W, t_N, t_Z; \theta, \omega)$. We may interpret this as the pay-off to being truthful.

The indirect marginal utility of income before tax is $V_W = -\beta/\omega$. The negative sign of this expression reflects the fact that a rise in W, holding w constant, implies a rise in labour supply. The implicit income tax is $T(W) \equiv W - B$. The marginal income tax rate is

$$T'(W) \equiv \frac{V_W}{V_B} = 1 - \frac{\beta}{\alpha\omega}. \tag{25}$$

The RHS of the first-order condition for N in (24) represents the marginal cost of N. Using (25), this term may be re-written as

$$\pi \equiv z + \tau + \left[w - T'(W)\right] h, \tag{26}$$

where

$$\tau \equiv z t_Z + t_N$$

is the effective tax on the money spent for each child. The cost of an extra child is thus the sum of an actual expenditure $(z + \tau)$, and an opportunity-cost $(\omega - T')\, h$. If parents give their children only the bare necessities of life, an additional child costs

$$\pi_0 \equiv \left(1 + t_Z\right) z_0 + t_N + \left[\omega - T'(W)\right] h_0. \tag{27}$$

We may think of π_0 as the fixed cost, and $(\pi - \pi_0)$ as the marginal cost of a child. The marginal cost of x is

$$\left(\pi - \pi_0\right) N = \left(\left(1 + t_Z\right)\left(z - z_0\right) + \left[\omega - T'(W)\right]\left(h - h_0\right)\right) N. \tag{28}$$

Notice that, while the marginal cost of N is affected by all the policy instruments, the marginal cost of x does not depend on t_N. Therefore, a positive t_N distorts parental choice away from the quantity of children, N, and towards their quality, measured by x.[9]

Let us again suppose that there are only two, equally numerous, household types differentiated by the (unobservable) values of θ and ω, and define an ij mimicker as an i-type household wanting to pass for a j $(i, j = 1, 2)$. This household will set the choice variables observable by the government at the level chosen by the j type. Since the government sees N, as well as W, an ij mimicker has then no choice of time allocation. It must devote (W^j/ω^i) units of time to the labour market, and $[(\omega^i - W^j)/\omega^i]$ to looking after children. As already pointed out, this makes mimicking less attractive than it would be if the mimicker had to worry only about W.

The ij mimicker's optimization has first-order conditions

$$U_Y^{ij} = \alpha^{ij}, \quad U_x^{ij} x_z^{ij} = \alpha^{ij} N^j (1 + t_Z). \tag{29}$$

The indirect utility function $V^{ij} = V(B^j, W^j, t_N^j, t_Z; \theta^i, \omega^i)$, $i \neq j$, is the pay-off to being untruthful. An i-type household will mimic if and only if

$$V(B^j, W^j, t_N^j, t_Z; \theta^i, \omega^i) > V(B^i, W^i, t_N^i, t_Z; \theta^i, \omega^i). \tag{30}$$

3.2 Government

Suppose that the government maximizes a concave function of the utilities of the two household types. The choice of policy instruments is restricted not only by the government budget constraint, but also by the self-selection constraint that neither household type must be better-off mimicking, than revealing its true characteristics. The policy problem is to choose $(B^1, B^2, W^1, W^2, t_N^1, t_N^2, t_Z)$ so as to maximize

$$SWF\left(V\left(B^1, W^1, t_N^1, t_Z; \theta^1, \omega^1\right), V\left(B^2, W^2, t_N^2, t_Z; \theta^2, \omega^2\right)\right), \tag{31}$$

subject to the government budget constraint,

$$\sum_i \left(t_Z Z^i + t_N^i N^i + W^i - B^i\right) = 0, \tag{32}$$

and self-selection constraints,

$$V\left(B^1, W^1, t_N^1, t_Z; \theta^1, \omega^1\right) \geq V\left(B^2, W^2, t_N^2, t_Z; \theta^1, \omega^1\right) \tag{33}$$

and

$$V\left(B^2, W^2, t_N^2, t_Z; \theta^2, \omega^2\right) \geq V\left(B^1, W^1, t_N^1, t_Z; \theta^2, \omega^2\right). \tag{34}$$

[9] The language is Gary Becker's. The specific point appears to have been made for the first time in Cigno (1986).

The Lagrange-multipliers associated with these constraints are, respectively, λ, σ^{12} and σ^{21}.

If neither of the self-selection constraints is binding (*i.e.,* if parents speak the truth), the solution is a first best. What this means is that households find mimicking not worth the trouble, and that the government can thus carry out the desired redistribution using personalized lump-sum transfers. If either of these constraints is binding, however, the government has to distort marginal incentives in order to deter mimicking, and the solution is then a second best. Which, if any, of the self-selection constraints will be binding is determined jointly with the direction of redistribution, and with the optimal choice of policy instruments. In this respect, the presence of domestically produced goods (N and x) makes a difference to the analysis.

In conventional optimal income taxation models, it is customary to impose the "agent monotonicity" condition (Seade, 1982), that $-V_Y/V_B$ is decreasing in the wage rate. When households are differentiated by earning ability only, this implies that the indifference curves are everywhere flatter for high than for low-wage households ("single-crossing") in the (B, Y) plane. Combining agent monotonicity with the assumption that the policy maker maximizes a convex combination of household utilities makes sure that the optimal redistribution is in favour of low-wage households. Since this rules out the possibility that a low-wage household will ever want to be taken for a high-wage household, one of the self-selection constraints can be disregarded. This is not legitimate if households differ also in domestic ability as in the present model, where one has to allow for the possibility that either household type has an interest in mimicking.

3.2.1 The direction of redistribution

As in the last section, we shall suppose that $\omega^1 < \omega^2$, and consider different possibilities regarding θ^1 and θ^2. Consider first the case where

$$\theta^1 > \theta^2,$$

as in Figure 1. In that case, the comparative advantage in raising children rests definitely with low-wage households. If θ^1 is not sufficiently larger than θ^2 to give type 1 higher *laissez-faire* utility than type 2, redistribution will be in favour of low-wage households. Then, $\sigma^{12} = 0$ and $\sigma^{21} > 0$. If the opposite is true, equity and efficiency pull in opposite directions, and we cannot say *a priori* which way the optimal tax system will redistribute. In the second case, the direction of redistribution depends on how strongly the policy maker dislikes inequality (how convex the social indifference curves are). If the social welfare function is

Benthamite as in Section 1, it is likely that the efficiency motive will predominate, and redistribution be in favour of low-wage households even though this means taking from the worse-off and giving to the better-off. Then, (34) will be binding $(\sigma^{21} > 0,\ \sigma^{12} = 0)$. Redistribution is more likely to be in favour of high-wage households if the social welfare function is strictly quasi-concave. If any of the self-selection constraint is binding, that will then be (33) $(\sigma^{12} > 0,\ \sigma^{21} = 0)$. But the countervailing effects of equity and efficiency considerations may require so little redistribution, that mimicking is not worth the trouble for low-wage households to pretend to be otherwise. There is then a chance that neither self-selection constraint will be binding, $(\sigma^{12} = \sigma^{21} = 0)$, and thus that a first best is achieved.

In the case portrayed in Figure 2,

$$\theta^1 < \theta^2,$$

low-wage households have lower *laissez-faire* utility than high-wage households, and may or may not have a comparative advantage in raising children. If they do, redistribution is definitely in their favour, and (34) is binding $(\sigma^{12} = 0,\ \sigma^{21} > 0)$. Otherwise, equity and efficiency pull in opposite directions. Then, redistribution is likely to be in favour of high-wage households $(\sigma^{12} > 0,\ \sigma^{21} = 0)$ if social preferences are Benthamite, in favour of low-wage households $(\sigma^{12} = 0,\ \sigma^{21} > 0)$ if social preferences are sufficiently inequality averse.

3.2.2 Taxing income

Denote by

$$\Theta^i\left(W^i, N^i\right) \equiv t_Z Z^i + t_N^i N^i + W^i - B^i \tag{35}$$

the total tax bill of a type-i household. We can immediately see that the *effective* marginal income tax is not simply $T'\left(W^i\right)$, but

$$\Theta_Y^i = T'\left(W^i\right) + t_Z Z_W + t_N^i N_W \equiv 1 + \frac{V_W}{V_B} + t_Z Z_W + t_N^i N_W. \tag{36}$$

Using the first-order conditions on the policy optimization, and adapting a procedure in Edwards, Keen and Tuomala (1994), it can be shown that the optimal effective marginal income tax rates for the two household types are

$$T'\left(W^1\right) = \frac{\sigma^{21} V_B^{21}}{\lambda}\left(\frac{V_W^{21}}{V_B^{21}} - \frac{V_W^1}{V_B^1}\right) - \left[t_Z \overline{Z}_W^1 + t_N^1 \overline{N}_W^1\right], \tag{37}$$

$$T'\left(W^2\right) = \frac{\sigma^{12} V_B^{12}}{\lambda}\left(\frac{V_W^{12}}{V_B^{12}} - \frac{V_W^2}{V_B^2}\right) - \left[t_Z \overline{Z}_W^2 + t_N^2 \overline{N}_W^2\right], \tag{38}$$

where $\overline{Z}^i_W$ and $\overline{N}^i_W$ are Slutsky terms (partial derivatives of the Hicksian demands, $\overline{Z}^i$ and $\overline{N}^i$).

To interpret these rules, suppose for a moment that $t^1_N = t^2_N = t_Z = 0$, so that the second right-hand-side term in each expression is identically zero. Suppose, also, that type-2 households are interested in mimicking type-1 (as in Figure 2), but not the other way round ($\sigma^{21} > 0$ and $\sigma^{12} = 0$). If the indifference curve of a low-wage household is steeper than that of a high-wage mimicker in the (B, W) plane, the imposition of a positive marginal income tax rate on the former will deter the latter from mimicking. Since type-1 households have no interest in mimicking, however, there is no point in distorting the decisions of type-2 households by imposing a positive marginal rate of income tax on them too. In this particular case, (37) – (38) thus imply $T'(W^1) > 0$ and $T'(W^2) = 0$ as in the Mirrlees-Stiglitz model ("no distortion at the top"). In general, however, σ^{12} could well be positive, and the marginal rate of income tax on high earners could thus be positive.

If, in addition to taxing income, the government taxes (subsidizes) child-specific commodities or the number of children, there is a revenue effect, reflected by the terms in square brackets in (37) – (38). Suppose, for instance, that

$$t_Z\overline{Z}^2_W + t^2_N\overline{N}^2_W < 0,$$

meaning that the revenue from taxing commodities and number of children falls as the labour supply of type-2 households goes up. Even if nobody were interested in mimicking high earners ($\sigma^{12} = 0$), imposing a positive marginal rate of income tax on them would raise tax revenue. There could thus be another reason, in addition to deterring mimicking, for distorting labour decisions. This makes it clear that there are two, quite independent, reasons why the no-distortion-at-the-top proposition need not hold. One is that, as households are differentiated by two characteristics, either self-selection constraint could be binding. The other is that indirect taxation introduces a revenue effect (Nava, Schroyen and Marchand, 1996).

3.2.3 Taxing or subsidizing commodities and the number of children

Using again the first-order conditions on the policy optimization, and noting that $N^{ji} = N^i$, it can be shown that fertility and the excise taxes must satisfy

$$\sum_i \left(t_Z\overline{Z}^i_{t_Z} + t^i_N\overline{Z}^i_{t^i_N}\right) = N^1\frac{\sigma^{21}V^{21}_B}{\lambda}(z^1 - z^{21}) + N^2\frac{\sigma^{12}V^{12}_B}{\lambda}(z^2 - z^{12}) \quad (39)$$

and

$$t_Z \overline{N}^i_{t_Z} + t^i_N \overline{N}^i_{t^i_N} = 0, \quad i = 1, 2. \tag{40}$$

The left-hand-side of (39) is the cost of distorting the demand for child-specific commodities through t_Z and t^i_N . If the effect is negative (positive), we say that the demand for child-specific commodities is "discouraged" ("encouraged"). The right-hand-side represents the corresponding gain. To see the intuition behind the rule, suppose, for instance, that true type-1 households spend less, for each of their children, than 21-mimickers ($\sigma^{21} > 0$, $\sigma^{12} = 0$ and $z^1 < z^{21}$). Since V^{21}_B and λ are positive, it is clear that, in this case, distorting prices in favour of adult-specific commodities would harm mimickers more than genuine low-wage households. Therefore, the relevant self-selection constraint can be relaxed by discouraging the purchase of child-specific commodities. Analogous considerations apply to the other possible cases.

The left-hand-side of (40) is the cost, and the right-hand-side the benefit, of distorting fertility decisions through t_Z and t^i_N. Since the benefit is zero, these taxes must *not* distort fertility decisions (though there may be other reasons, as we shall see in a moment, for distorting fertility using different means). Intuitively, that is because the mimicker has to produce the same number of children as the mimicked, and distorting fertility choices has thus no "screening power". This rule does not imply that $t^i_N = 0$. If (39) provides a second-best rationale for taxing or subsidizing child-specific commodities, the policy prescription is to set t^i_N so that it totally offsets the distortionary effect of t_Z. Thus, (40) implies that t_Z and t^i_N must have opposite signs if Z and N are Hicksian complements, the same sign if Z and N are Hicksian substitutes.

Since t_Z and t^i_N are not the only policy instruments affecting the cost of raising children, we can see that by just looking at (26) that (40) does not imply that fertility will not distorted at a second-best optimum. By taxing income at the margin, the government does in fact reduce the opportunity-cost of child bearing. Even if it so happens that $t_Z = t^i_N = 0$, the post-tax marginal cost of children will still differ from its pre-tax level so long as $T'(Y) \neq 0$. So long as child-specific commodities, or the number of children, are taxed or subsidized, the same is likely to be true even if $T'(Y) = 0$. Although there is no point in distorting fertility decisions to discourage mimicking (because children are visible), there may thus be a point in distorting them for distributional reasons, or in order to balance the effects of other distortions.

The *effective* marginal tax rate on children is

$$\Theta^i_N = \tau^i - T'(W^i)\omega^i h^i = \tau^i + \frac{\beta^i h^i}{\alpha^i} - \omega^i h^i, \tag{41}$$

where the sign of the second equation follows from (25). For a household of type i, a child is a *tax asset* if Θ_N^i is negative, a *tax liability* if Θ_N^i is positive. In first best, children are clearly tax-neutral for every type of household because there is no distortionary taxation (Nerlove, Razin and Sadka, 1993). In second best, however, that will depend on the number of characteristics by which households are differentiated, and on the number of tax instruments available. If households are differentiated by the wage rate only, and there is only an income tax, children are tax-neutral for high earners (the potential mimickers), a tax asset for low-wage households. With taxes or subsidies on child-specific commodities or number of children, anything is possible, because $\tau^i \equiv z^i t_Z + t_N^i$ can have any sign. The same is true if households are differentiated by more than one characteristic, because both τ^i and $T'(W^i)$ can have any sign.

We have thus seen that an optimal policy may include taxes or subsidies on commodities and number of children, in addition to a non-linear income tax. Does this apply also to the case where the utility function is separable in consumption and time use? The well-known Atkinson-Stiglitz theorem says that it does not if households differ only in their wage rate.[10] Balestrino *et al.* (2003) show analytically that it does if households differ also in their domestic ability to produce a non-transferable good. Balestrino *et al.* (2002) report numerical examples of cases where it is optimal to tax or subsidize child-specific commodities and the number of children despite separability of the utility function.

The public debate about tax policy and policies towards the family often seems to assume that children should be a tax asset to their parents, especially in low-income households. Cremer, Dellis and Pestieau (2000) show that this is indeed the case if fertility is exogenous, and households are differentiated only by their wage rate (as well as by number of children). But we have just seen that this is not necessarily true when fertility is endogenous. If parents are differentiated not only by their wage rate, but also by their ability to raise children, an extra child may not reduce the tax bill. Cigno and Pettini (2003) find the same in an indirect-taxation context even without child-rearing ability differentiation.

4 Inducing parents to invest in their children

Let as now interpret x as a child's future earning and thus, for any given income-tax schedule, tax-paying capacity. Parents may, or may not, derive utility from x, but the money-equivalent of this utility is in any case lower than x. There is thus a positive fiscal externality: an increase

[10] See Atkinsons and Stiglitz (1976).

in either N or x relaxes the government budget constraint. Suppose that x depends on an action (educational investment in the broadest sense) undertaken by the child's parents, a, and on random variable (luck). Following Mirrlees (1974), we will treat x itself as a random variable with known density conditional on action a, $f(x, a)$.[11] This means that a couple can choose the probability distribution, rather than the actual value of x. The number of children, N, is deterministically chosen by the parents. Couples are *ex-ante* identical. The policy problem is how to provide parents with the incentive to invest at the level desired by the policy maker, rather than how to induce them to reveal their characteristics as in the last section. The policy instruments are direct and indirect subsidies. Although the income-tax schedule is taken as given, the progressivity of the tax system as a whole, and the direction of the redistribution will be affected by the choice of subsidy schedules.[12]

4.1 Parents

The *ex-post* utility function of the generic couple,

$$U = U(Y + v(x)N), \tag{42}$$

is only a little less general than (21). In the present context, the assumed concavity of $U(.)$ implies risk aversion. The couple's expected utility is

$$E(U) = \int U(Y + v(x)N)f(x, a)dx. \tag{43}$$

The total cost of the children is

$$C = c(N) + z(a)N, \tag{44}$$

where $c(N)$ is the minimum cost (what, in the last section, we called π_0) of raising N children, and $z(a)$ the cost of investing a in each of them. To ensure concavity of $E(U)$, $c(.)$ and $z(.)$ are assumed increasing an convex. The household budget constraint is

$$Y + [m(x) - z(a)]N - c(N) \leq W, \tag{45}$$

where W is again parental income, and $m(x)$ is a per-child government transfer (the negative of what, in the last section, we called t_N), possibly

---

[11] The standard monotone likelihood ratio (MLR) condition, $(\frac{f_a}{f})$ increasing in x, and the convex distribution function (CDF) condition, are assumed to hold. Since a is a continuous variable, the likelihood ratio is increasing in x iff $\left(\frac{f_a}{f}\right)$ is increasing in a (Milgrom, 1981).

[12] The exposition that follows draws on Cigno, Luporini and Pettini (2003).

conditional on x. If a involves the use of parental time, $z(a)$ will include the opportunity-cost of this time, and W is then to be interpreted as full parental income.

Parents choose (a, N) to maximize (43), given (45). The first-order conditions are

$$\int [m(x) + v(x) - z(a) - c'(N)]\, U' f(x, a) dx = 0 \tag{46}$$

and

$$-z'(a)N \int U' f(x, a) dx + \int U f_a(x, a) dx = 0. \tag{47}$$

The former says that the couple will procreate to the point where the expected benefit of an additional child equals the cost (zero), the second that it will invest in each child to the point where the expected marginal benefit equals the expected marginal cost (positive).

4.2 Government

The government maximizes the sum of the expected utilities of all the couples. As the latter are *ex-ante* identical, that is the same as maximizing (43). Assuming that the number of couples, hence the number of future tax payers, is "large", the government faces no uncertainty over its future tax receipts, and the budget constraint may then be written in expected tax revenue terms as

$$n \int [x - m(x)] f(x, a) dx \geq 0. \tag{48}$$

Comparing (48) with (45) makes it clear that, in *laissez faire*, parents would have no reason to take into account the effects of their choice of (a, N) on the government's budget constraint. As a is private information, there is then a moral hazard problem where the amount of parental time and money invested in each child is concerned. As N is observable, by contrast, the government can use a forcing contract to get a couple to deliver the desired number of children.

The government will then choose N and $m\,(.)$ to maximize (43), subject to its own budget constraint (48), and to the incentive-compatibility constraint (47). For every possible realization of x, the corresponding value of m will be determined by the first-order condition,

$$U' f - \lambda f + \mu \left(-n z' U'' f + U' f_a\right) = 0, \tag{49}$$

where λ is the Lagrange-multiplier associated with (48), and μ the one associated with (47).

As in the hidden-characteristics model, a first best is again defined as a solution to the policy optimization problem subject only to the government budget constraint. There, a first best would be achieved if the government could observe the parental type. Here, it would be achieved if the government could observe a. In first best, (49) reduces to $U' = \lambda$, implying that

$$m(x) + v(x) = \text{constant}. \tag{50}$$

This says that the government would fully compensate the couple for having more than the privately optimal number of children, and fully insure it against the risk of getting low-earning children. This full-insurance property is typical of first-best policies. In standard Principal-Agent models, however, it follows from the assumption that the principal maker is less risk-averse than the agent. Here, by contrast, the principal is as risk-averse as the agents, and the result derives from the fact that, as there are many agents, the government does not face any risk over future tax revenues.

As a is not observable, however, the government must satisfy the incentive-compatibility constraint. If (47) is binding, the government will have to depart from the full-insurance principle, and the solution will be a second best. It is then convenient to re-write (49) as

$$\frac{\lambda}{U'} = 1 + \mu \left(nrw' + \phi \right), \tag{51}$$

where $r \equiv -U"/U'$ is the Arrow-Pratt measure of absolute risk aversion, assumed constant, and $\phi \equiv (\frac{f_a}{f})$ is a close relative of the likelihood ratio. Using standard arguments, it can be shown that (51) implies

$$\frac{dm}{dx} = \frac{\phi'}{\frac{\lambda}{\mu} \frac{N}{U'} r} - v'. \tag{52}$$

Since $\phi'$and v' are positive by assumption,and λ and μ must be positive at an optimum, (52) can have any sign. As v' is decreasing in x, however, the sign is more likely to be negative if x turns out to be small, positive if it turns out to be large. Therefore, the second-best subsidy schedule may well turn out to be U-shaped as in Figure 3 – decreasing in x at low realizations of this variable, where the insurance principle is more likely to prevail, increasing at high ones, where the incentive principle is more likely to predominate.

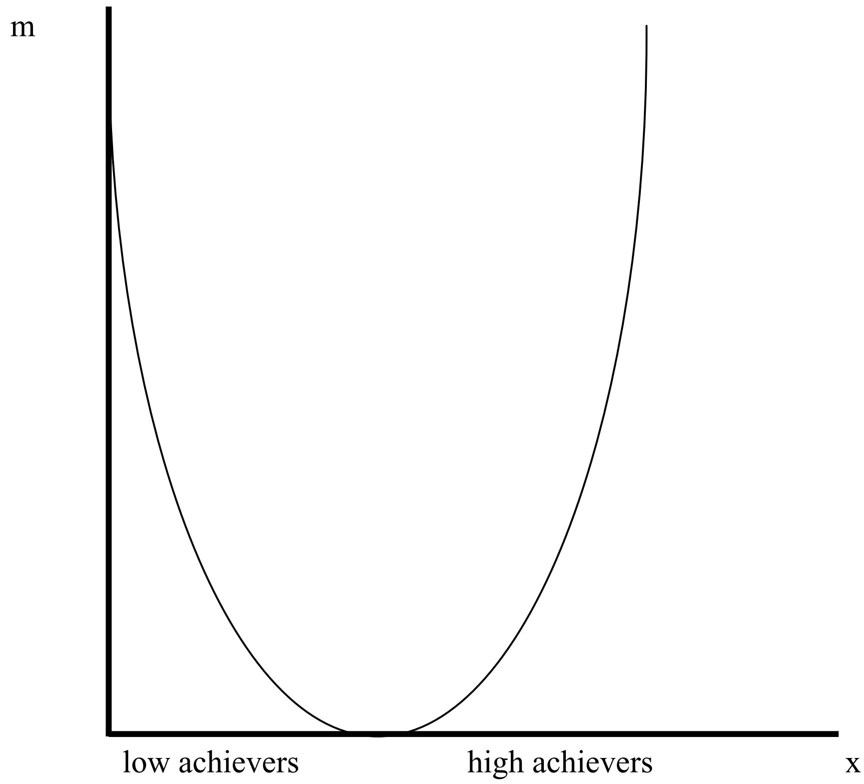

Figure 3

With regard to fertility, the government will offer parents the following forcing contract:

$$\begin{array}{ll} g_1 = m(x)N^* & \text{if } N = N^* \\ g_2 < g_1 & \text{if } N \neq n^* \end{array}, \tag{53}$$

where N^* is the optimal (first or second best) value of N, and g_2 low enough (even zero or negative) to deter anyone from having a different number of children.

4.3 Price subsidies

Let us now see if the government can get any closer to a first best by using indirect taxes or subsidies as in the hidden-characteristics model of the last section. Suppose that a involves the purchase of a certain commodity, and the use of fixed household resources. The assumption that z' increases with a may then be justified by the increasing scarcity of these resources. For example, if a is measured as number of children's books, increasing a implies that the parent must divert more time from work or leisure to reading the books to, or with, the child. If a is measured as use of an external facility, such as a school or a gym, an increase in a implies that more parental time, and more of the family car, must be diverted from other uses to bringing and fetching the child.

Let ξ denote a subsidy (the negative of what, in the last section, we called t_Z) on the commodity in question. The household budget constraint then becomes

$$Y = W + [m(x) - z(a) + \xi a]N - c(N). \tag{54}$$

For any given N, the first-order condition on the agent's choice of a is now

$$-\left[z'(a) - \xi\right] N \int U' f(x, a)dx + \int U f_a(x, a)dx = 0. \tag{55}$$

At an optimum, $(z' - \xi)$ must be positive (otherwise, the agent's utility could be raised by reducing a).

So far, we have implicitly assumed zero cost of delivering the direct subsidy. As m was the only policy instrument, that made no qualitative difference to the results. The moment we introduce a second policy instrument, however, it becomes important to compare administration costs. Let k_ξ denote the cost of administering the price subsidy, and k_m that of delivering the direct subsidy. Given (54), the government now maximizes (43), subject to the government budget constraint,

$$N \int [x - (1 + k_\xi)\, \xi a - (1 + k_m)\, m(x)] f(x, a)dx \geq 0, \tag{56}$$

and to the incentive-compatibility constraint (55). Using (59), the first-order condition on the government's choice of ξ may be written as

$$a(k_\xi - k_m) = \frac{\mu}{\lambda} \int U' f(x,a) dx. \tag{57}$$

If $k_\xi \leq k_m$, the policy optimization has a corner solution, with only the price subsidy in use. As the incentive-compatibility constraint will not be binding, the first-order condition on the choice of ξ is then

$$\int U' f(x,a) dx = \lambda \left(1 + k_\xi\right). \tag{58}$$

The government must raise the price subsidy to the point where the expected marginal benefit equals the marginal cost. The intuitive explanation is that, while ξ is certain, m is uncertain (because it is conditional on the realization of x). Since parents are averse to risk, the incentive effect of a subsidy given through the price system is then larger than that of an uncertain payment with the same expected value. Doing without the direct subsidy has the disadvantage that it is not possible to insure parents against the risk of getting a child with low x. If delivering m costs as much or more than administering ξ, however, providing this kind of insurance is not a cost-effective way of rising social welfare.

If $k_\xi > k_m$, we get an interior solution, with both policy instruments in use. As this is a possibility, the redundancy-of-indirect taxation result is once again refuted. The first-order condition on the government's choice of $m\left(.\right)$ is now

$$\left(1 + k_m\right) \frac{\lambda}{U'} = 1 + \mu \left[(v' - \xi) N r + \phi \right]. \tag{59}$$

The slope of the direct payment schedule can again take any sign, but the indirect subsidy makes it more likely that this will be negative at low realizations of x, positive at high ones. In other words, the possibility of providing the investment incentive indirectly through the price system (e.g., subsidizing educational services) is an argument for directly subsidizing very unsuccessful (e.g., educationally subnormal), as well or instead of highly successful, children.

5 Conclusion

Extending the optimal taxation approach to would-be parents and families with children yields a rich crop of new results, and throws doubt on some established ones. Among the latter, are the propositions that indirect taxation is redundant in the presence of an income tax, and that top wage earners should not have their labour decisions distorted by a positive marginal rate of income tax. The new results depend crucially on the assumptions one makes with regard to fertility determination. If the number of children is taken to be exogenous, an extra birth should reduce the net tax bill paid by the parents. That is not necessarily true, however, if fertility is taken to be endogenous. As the differentiating characteristic is then the couple's ability to raise children, rather the number of children, the more able parents should have more children, and get larger per-child subsidies, than the less able parents. If a child's chances in life are positively affected by some unobservable parental action (e.g., time and money spent on the child), the subsidy should be made contingent on some predictor of the child's future earning and tax-paying capacity. If the child's educational record is a good predictor of this capacity, we may interpret the subsidy as a scholarship.[13] Otherwise, it may be better to rely on the child's actual work and tax-paying record up to a certain date. For this to be reliable, however, the date should be set as late as possible (say, when the child is in middle life, and the parents are on the point of retirement), and it then comes natural to interpret the subsidy as a pension entitlement for the parents.[14]

6 References

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

Balestrino, A., Cigno, A. and A. Pettini (2002), Endogenous fertility and the design of family taxation, *International Tax and Public Finance* 9, 175-193

———, Cigno, A. and A. Pettini (2003), Doing wonders with an egg: optimal re-distribution when households differ in market and non-market abilities, *Journal of Public Economic Theory* 5, 479-498

Becker, G. (1991), *A Treatise on the Family* (2nd enlarged edition), Harvard University Press, Cambridge Mass.

——— and R. J. Barro (1988), A Reformulation of the economic theory of fertility, *Quarterly Journal of Economics* 103, 1-25

Bovemberg, A. L. and B. Jacobs (2005), Redistribution and education subsidies are Siamese twins, *Journal of Public Economics* 89, 2005-2035

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

[13] For a model of university scholarships along these lines, see Cigno and Luporini (2009a).

[14] Not for the children, because it is the parents who make the hidden eductional investment which the government wants to encourage. For an actual policy proposal based on this idea, see Cigno (2009).

——— (1993), Intergenerational transfers without altruism: family, market and state, *European Journal of Political Economy* 9, 505-518

——— (2001), Comparative advantage, observability, and the optimal tax treatment of families with children, *International Tax and Public Finance* 8, 455-470

——— (2006), A constitutional theory of the family, *Journal of Population Economics* 19, 259-283

——— (2009), How to avoid a pension crisis. A question of intelligent mechanism design, *CESifo WP* 2590

——— and A. Luporini (2006), Optimal policy towards families with different amounts of social capital, in the presence of asymmetric information and stochastic fertility, *CESifo WP* 1664

——— and A. Luporini (2009a), Scholarships or student loans? Subsidizing higher education in the presence of moral hazard, *Journal of Public Economic Theory* 11, 55-87

——— and A. Luporini (2009b), Optimal family policy in the presence of moral hazard, when the quantity and quality of children are stochastic, *IZA* DP 4179

———, Luporini, A. and A. Pettini (2003), Transfers to families with children as a principal-agent problem, *Journal of Public Economics* 87, 1165-1177

———, Luporini, A. and A. Pettini (2004), Hidden information problems in the design of family allowances, *Journal of Population Economics* 17, 645-655

——— and A. Pettini (2003), Taxing family size and subsidizing child-specific commodities? *Journal of Public Economics* 87, 75-90

Cremer, H., Dellis, A. and P. Pestieau (2003), Family size and optimal income taxation, *Journal of Population Economics* 16, 37-54

———, Gahvari, F. and P. Pestieau (2006), Pensions with endogenous and stochastic fertility, *Journal of Public Economics* 90, 2303-2321

———, Gahvari, F. and P. Pestieau (2008), Pensions with heterogeneous agents and endogenous fertility, *Journal of Population Economics* 21, 961-981

Edwards, J., Keen, M. and M. Tuomala (1994), Income tax, commodity taxes and public good provision: a brief guide, FinanzArchiv, 51, 472-487

Hellwig, M. (2008), A generalization of the Atkinson-Stiglitz (1976) theorem on the undesirability of nonuniform excise taxation, *Max Planck Institute for Research on Collective Goods Preprints* 2008/47

Milgrom, P., 1981, Good news and bad news: representation theorems and applications, *Bell Journal of Economics* 12, 380-391

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

——————— (1974), Notes on welfare economics, information and uncertainty, in Balch M.S., McFadden D. L. and S. Y. Wu (eds.) *Essays on Economic Behavior and Uncertainty*, North-Holland, Amsterdam.

Nava, M., Schroyen, F. and M. Marchand (1996), Optimal fiscal and public expenditure policy in a two-class economy, *Journal of Public Economics* 61, 119-137

Nerlove, M., Razin, A. and E. Sadka (1993), Children as capital goods or a base for income redistribution policies, *Public Finance* 48, 78-84

Sandmo, A. (1990), Tax distortions and household production, *Oxford Economic Papers* 42, 78-90

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