The Redistributive Role of Child Benefits Revisited

Tomer Blumkin[*] Yoram Margalioth[•] Efraim Sadka[•]

January, 2014

Abstract

In this paper we reexamine the commonly invoked argument that due to the existence of a negative correlation between earning ability and family size, the latter can be used as a 'tagging' device, justifying subsidizing children (via provision of child allowances) to enhance egalitarian objectives. Employing a benchmark setting where the quality-quantity paradigm holds, we show that the case for subsidizing children is far from being a forgone conclusion. We demonstrate that the desirability of subsidizing children crucially hinges on whether benefits are means-tested or being accorded on a universal basis.

JEL Classification: D6, H2, H5

Key Words: child allowance, re-distribution, means-testing, universal, tagging, optimal taxation

[*] Department of Economics, Ben-Gurion University, Beer-Sheba 84105, Israel, CesIfo, IZA. E-mail: tomerblu@bgumail.bgu.ac.il.

[•] The Buchman Faculty of Law, Tel-Aviv University, Tel-Aviv 69978, Israel. E-mail: margalio@post.tau.ac.il

[•] The Eitan Berglas School of Economics, Tel Aviv University, Tel-Aviv 69978, Israel, CesIfo, IZA. E-mail: sadka@post.tau.ac.il

1. Introduction

Family size is a key component in the determination of income tax liability in all OECD countries [see, e.g., Bradshaw and Finch (2002)]. Two major decisions affect family size: (i) marriage/cohabitation and (ii) fertility. In this paper we focus on the latter aspect, namely, the optimal fiscal treatment of children. In practice, the existence of children generally reduces the household's tax liability, taking a variety of forms, including: income splitting amongst (a standardized number of) family members (as in France); exemptions or standard deductions (as in the US); specific childcare deductions; tax credits; and the provision of child allowances, which could be either universal or means-tested.[1] In most countries the policy implemented is a mixture of some or all of the above measures.

There are different rationales for the commonly observed preferential tax treatment of children, such as enhancing fertility, encouraging labor market participation of women and promoting income re-distribution. In this paper we focus on the latter without discounting the importance of the other rationales.

The optimal income tax literature [starting with the seminal work of Mirrlees (1971)] emphasizes the screening problem that the government is facing due to its inability to observe the earning capacities of the individuals. According to the quality/quantity paradigm [see the pioneering studies of Becker (1960) and Becker and Lewis (1973)], low-ability families may choose to 'specialize' in quantity, that is, to raise more children relative to higher-ability households.[2] Child-related subsidies can, therefore, be used to enhance re-distribution: family size can be employed as an indicator [a 'tagging' device, à la Akerlof (1978)] for the earning capacity of the household.

A relatively recent strand in the optimal income tax literature [see Cigno (2011) for a comprehensive recent review] examines the potential supplementary re-distributive role of extending the tax base to account for the number of children in the household and child-related consumption (such as, education and daycare). This new strand of the literature emphasizes a key distinction from the standard optimal tax

[1] Nearly all developed countries provide universal child allowances; namely, child allowances that do not depend on household's income (but may well vary with the number of children) with the notable exception of the US, where the child allowances system (which is embedded in the EITC program) is (partly) means-tested. In this paper we draw attention to the importance of distinguishing between a universal and a means-tested system, in analyzing re-distributive policy implications of child benefits.
[2] For evidence of the existence of a quality-quantity trade-off see, e.g. Hanushek (1992).

setting, which derives from the unique characteristics of children: a crucial part of the process of rearing children may be viewed as consumption of a non-transferable domestically produced good (e.g., parental attention and affection), the production of which requires expertise (ability to nurture) that is different from the ability to earn. The literature considers parental capacity as a second source of heterogeneity across households, in addition to the variation in earning ability, which is the only source of variation in the standard Mirrleesian optimal tax framework [Cigno (2001), Balestrino et al. (2002)]. This bears new re-distributive implications, affecting both policy goals and system design. In particular, the new literature shows that the direction of re-distribution is not necessarily in favor of the low earning-ability individuals, because the latter may enjoy some marked advantage in child-rearing, which may, all-in-all, compensate (in utility terms) for their low earning capacity. It further shows that the tax system design can employ observed family attributes to enhance target efficiency ('tagging'). The properties of the optimal integrated tax-transfer system (which allows the tax liability to depend on income level, on family size as well as on expenditure on child-related goods) are generally shown to depend on both comparative- and absolute-advantage (in domestic vis-à-vis market production) considerations.

The new literature challenges also some of the key results of the optimal income tax literature, such as, the desirability of a zero marginal tax rate levied on top-earners, when the skill distribution is bounded [Balestrino et al. (2002)] and the redundancy of commodity taxation [Balestrino et al. 2003)]. Most surprisingly, counter to conventional wisdom and common practice, Balestrino et al. (2002) demonstrate that it *may* be optimal to tax children at the margin; namely, to make the net tax bill an *increasing* function of the number of children.

In this paper, we attempt to provide a fairly general characterization of the optimal income tax cum child benefit system. Among other things, we revisit the surprising result mentioned above and provide an explanation for the mechanism at work. Unlike the aforementioned literature, we assume, a la Mirrlees, that households differ *only* in their earning capacities. Notice, however, that unlike the standard optimal tax framework which rules out household production (such as rearing children), in our setting the variation in earning capacities implies that households also differ in their ability to raise children (the opportunity cost of raising children which is assumed to be time-consuming would be lower for households with a lower earning capacity). We do so in order to allow for closed-form solutions, which would facilitate the interpretation of the results and would elicit the forces at play. We extend the two-type framework used by Cigno (1986) and the subsequent literature by considering a continuum of households. This allows us to derive meaningful policy implications by relating the properties of the optimal tax-and-transfer system to actual skill distribution. We impose only few restrictions on the preferences of the households: we assume, following Diamond (1998) and Salanie (2003), that the households' utility is quasi-linear with respect to parental (as opposed to child related) consumption; and further assume that the preferences are 'calibrated' according to the quality-quantity paradigm [Becker (1973)] in the benchmark (no-tax) case, implying the existence of a negative correlation between family size and earning ability. We assume that the government is free to use an integrated system of income taxes and fertility-related subsidies. As this general system allows for the possibility of making the level of child benefits dependent on the household's level of income, we will henceforth refer to it as a means-tested system. We also examine the special case of a *separable* system comprised of an income tax component, which does not depend on the household's number of children, and a *universal* child-benefit component, which does not depend on the household's level of income. The importance of examining this special case derives from its policy relevance and is twofold: (i) first, as discussed earlier on, universal child benefits systems are fairly common amongst OECD countries; (ii) second, our analysis allows the evaluation of adding child benefits piece-meal to an existing fiscal system (rather than considering the case where the entire tax-benefits system can be designed from scratch).

Our contribution to the literature is in showing that the desirability of subsidizing/taxing children at the margin crucially hinges on whether child benefits are means-tested or universally provided. Notably, we show that it is *unambiguously* optimal on re-distributive grounds to tax children at the margin; that is, the total tax liability should rise with the number of children for a given level of income, when child benefits are means-tested. In contrast, when child benefits are universally provided, it *may* be optimal to subsidize children at the margin. We provide empirically plausible parametric assumptions regarding the underlying skill distribution and the labor supply elasticity (labor-leisure preferences) under which subsidies are warranted.

The structure of the remainder of the paper will be as follows. In the following section we introduce the analytical framework. In section 3 we formulate the government problem and derive the properties of the general (means-tested) income tax cum child benefit system. The universal case is discussed in section 4. Section 5 concludes.

2. The Model

Consider an economy with a continuum of households. The number of households is normalized to unity, with no loss in generality. We assume that the production technology of all market goods employs labor only, and exhibits constant returns to scale and perfect substitution across the various skill levels. Following the standard assumption in the optimal tax literature, we assume that households differ in a single attribute: their earning ability/skill level (equaling the wage rate, assuming a competitive labor market). We let w denote the wage rate and assume that w is distributed over some, possibly unbounded, support $[\underline{w}, \overline{w}]$, with a cumulative distribution function $F(w)$ and corresponding density $f \equiv F'$. We follow Mirrlees (1971) by assuming that abilities (wage rates) are unobserved by the government, thus constraining the latter to second-best re-distributive policies.[3]

All households share the same preferences, represented by the following utility function:

$$(1) \quad V(c,l,n,e,\alpha) = c + g(l,n,e,\alpha);$$

where c denotes (parental) consumption, n denotes the number of children, e denotes the (per-child) amount of child-specific market goods/services provided by the parents (education, day-care etc.), α denotes the amount of time (per-child) dedicated by parents to child rearing/nurturing activities (parental attention) and l denotes (parental) leisure.[4] We assume that g is jointly concave, strictly increasing in all its arguments, and satisfies INADA conditions to ensure interior solutions.

Several remarks are in order. Note first that our setting captures the fundamental quantity-quality trade-off [a la Becker (1960)] faced by the household, whether to increase the quantity (number of children, n) or invest in their quality (say, human capital/education, e, and/or parental attention, α).[5] We make several

[3] As differences in earning ability are assumed to be the single source of heterogeneity in the economy, we refrain from introducing horizontal equity considerations into the analysis.

[4] Notice that e is measured per-capita, for simplicity; namely, there are no economies of scale embodied in the consumption of children.

[5] The variable e is henceforth interpreted as the level of parental provision of child related consumption goods, such as education (Becker, 1991), but may well take additional plausible interpretations, such as

simplifying assumptions to render our model tractable. We invoke a quasi-linear specification, which is fairly common in the optimal tax literature [see, e.g., Diamond (1998) and Salanie (2003)] and rules out income effects.[6] It is worth noting that Becker (1960) conjectured that the elasticity of family size (quantity/number of children) with respect to income would be rather small, which is consistent with some of the empirical evidence [see, e.g., Hotz, Klerman and Willis (1997), and more recently Cohen, Dehejia and Romanov (2013)]. We follow the standard approach in the endogenous fertility literature and assume that the household can deterministically choose the number of children.[7] Finally note that in our setting there is no difference between the household and the individual, as we adopt a unified approach to household decision making (making no distinction between primary and secondary earners).

Each household is faced with the following budget constraint:

$$\begin{aligned} & c + n \cdot e = z(y, n); \\ & y = w \cdot (1 - l - n \cdot \alpha), \end{aligned} \tag{2}$$

where y and z denote gross and net income levels, respectively. Several remarks are in order. First notice, that we normalize each household's time endowment as well as the price levels of both c and e to unity, with no loss in generality. Notice further that high-ability households find it more costly to raise children, due to the larger opportunity cost they incur (forgoing time in the labor market).[8] Finally, note that we consider a general non-linear tax schedule, which depends both on the number of children and on the level of gross income (both of which are assumed to be observable by the government). This tax schedule is implicitly defined by the difference between the gross and net income levels, $t(y,n) \equiv y - z(y,n)$. Note that $t(y,$

---

the maximized lifetime utility of each child (Becker and Barro, 1988). An alternative interpretation of the utility form given in equation (1) is that $g(l, n, e, \alpha) = s[l, n, x(e,\alpha)]$, where x denotes the domestic production function of 'quality' (child related consumption) employing parental time as well as market goods such as private tutoring and day care services as inputs [see Cigno (2001)].

6 By continuity considerations, provided that the utility function is separable between parental consumption, c, and its other arguments $(l, n, e$ and $\alpha)$, accommodating moderate income effects will not change the qualitative nature of our results.

7 We acknowledge that parents have less than perfect control over the number of children (by proper choice of intercourse frequency, the use of contraceptives, etc.). For a recent paper that allows households to control only the probability distribution of the number of children see Cigno and Luporini (2011). For models assuming exogenous fertility see, for instance, Cremer, Dellis and Pestieau (2003).

8 Higher-skilled households will partially mitigate this by replacing their own time inputs with relatively cheaper outsourced day-care/tutoring services (via choosing to increase e and decrease α). We will further discuss this point below.

n) denotes an integrated income tax and child benefit system. From an economic point of view, this system, referred to as a means-tested system, cannot be decomposed into separate income tax and means-tested child benefit components, except in the special case where $t(y, n)$ takes the additively-separable form: $t(y,n) = a(y) + b(n)$. The latter is referred to as a universal system, with $a(y)$ denoting an income tax component and $b(n)$ denoting a non means-tested (universal) child benefit system.

The typical household seeks to maximize the utility function in equation (1), subject to the budget constraint in (2). Substituting from the budget constraint in (2) into the utility function in equation (1) to eliminate c and l, we obtain the indirect utility function $U(w)$ given by:

(3) $U(w) = \max_{n,y,e,\alpha}\{[z(y,n) - n \cdot e)] + g[(1 - y/w - n \cdot \alpha), n, e, \alpha]\}$,

which can be re-written as:

(3') $U(w) = \max_{n,y}[z(n, y) + h(n, y/w)]$,

where $h(n, y/w) = \max_{e,\alpha}[-n \cdot e + g[(1 - y/w - n \cdot \alpha), n, e, \alpha]]$.

The first-order-conditions for the typical w-household's optimal choice with respect to y and n are given by:

(4) $z_n(y,n) + h_1(n, y/w) = 0$,

(5) $z_y(y,n) + h_2(n, y/w)/w = 0$,

where z_n and $1 - z_y$ denote, respectively, the marginal subsidy provided to an additional child, and, the marginal tax rate levied on labor income; and, h_k denotes the partial derivative with respect to the k^{th} argument.[9]

The first-order conditions given in (4) and (5) define implicitly the optimal choice of the w-household given by $n(w)$ and $y(w)$. We will henceforth make two additional assumptions with respect to the optimal household choice in the benchmark setting with no taxes in place; namely, when $z \equiv y$, hence, $z_y = 1$ and $z_n = 0$. First we assume that, labor supply is upward sloping. Formally, $\partial[y(w)/w]/\partial w > 0$. This assumption is fairly standard and is satisfied, for instance, in the quasi-linear specification examined by Diamond (1998) and Salanie (2003), amongst others. We

[9] We will henceforth assume that the second order conditions are always satisfied, thus employ first-order conditions only to characterize the individual incentive constraints when formulating the government problem. This latter assumption will ensure no 'bunching' in the optimal solution of the government problem [see Ebert (1992), for a rigorous treatment of 'bunching' in the context of optimal non-linear labor income tax in the continuum case].

further impose a 'calibrating' assumption, requiring that family size will be negatively correlated with earning ability. Formally, $\partial n(w)/\partial w < 0$. That is, in line with the quantity-quality paradigm, poor (low-ability) families will 'specialize' in quantity and hence choose to have a larger number of children.[10]

Combining the two assumptions implies a negative relationship between (the optimally set) family size and labor supply. By virtue of (4) it then follows, assuming no taxation in place ($z_n = 0$), that $-h_{11}/h_{12} < 0$. By virtue of the household's second-order conditions $h_{11} < 0$, hence, $h_{12} < 0$.

Clearly, the two assumptions, which are stated as comparative statics properties of the optimal household choice in the no-tax regime, impose restrictions on the form of utility function given in (1). It is straightforward to provide specific functional forms that satisfy the two properties. In appendix A, for instance, we show that the properties are satisfied, when g takes an additively-separable form; namely, $g(l,n,e,\alpha) = \psi(l) + v(n) + u(e) + \Phi(\alpha)$, where ψ, v, u and Φ are strictly increasing and strictly concave; and, in addition, the following property is satisfied, $-\alpha \cdot \Phi''(\alpha)/\Phi'(\alpha) > 1$; that is, the degree of concavity of Φ, as measured by the familiar coefficient of relative-risk aversion, is sufficiently large. The latter assumption ensures that the opportunity cost of raising children, given by $\alpha(w) \cdot w$, will rise in w, ensuring that patterns of specialization will remain in line with the quality-quantity paradigm [$n'(w)<0$, $e'(w)>0$]. Notice that high-skill households faced with high-opportunity costs associated with the time dedicated to nurturing activities will respond by downward adjusting the latter (resorting instead to outsourcing services such as tutoring and day-care). The concavity assumption implies that the adjustment effect would be relatively small.[11]

We next turn to characterize the properties of the integrated income tax cum child benefit system.

---

[10] Notice that to obtain our qualitative results all we need is a negative correlation between quantity (number of children) and ability (of parents). Our qualitative results would remain unchanged even if quality were fixed at some exogenous level. In our setting quality is endogenously determined by the households and quantity and quality are negatively related in the benchmark (no-tax) case, in line with the quality-quantity paradigm.

[11] In a previous version of the paper we have examined the simpler case where α is a fixed parameter (rather than being endogenously determined, as we assume). In this case the negative correlation between the skill level and family size is satisfied with the additively separable functional form, without imposing additional assumptions [see Moav (2005) for a similar setting].

3. The General (Means-Tested) System

The government seeks to maximize an egalitarian social welfare function given by:

(6) $$W = \int_{\underline{w}}^{\overline{w}} G[U(w)]dF(w);$$

where G is strictly increasing and strictly concave,[12] by choosing the tax schedule, $t(y,n)$, subject to a revenue constraint:

(7) $$\int_{\underline{w}}^{\overline{w}} t[y(w), n(w)]dF(w) = R;$$

where $y(w)$ and $n(w)$ are the optimal individual choices of the gross income level and number of children, respectively, given by the first-order-conditions in (4) and (5); and R denotes the (pre-determined) level of government revenue needs. Notice that we start by analyzing the most general (means-tested) setting in which taxes/benefits may vary across income levels as well as family size. Below, we also consider a universal system (as is often the case in many countries) in which the tax function takes an additively separable form: $t(y,n) = a(y) + b(n)$.[13]

Following Salanie (2003), the government optimization problem can be formulated as an optimal control problem where the government is choosing the functions $U(w), n(w)$ and $y(w)$, so as to maximize the social welfare function in equation (6), subject to the revenue constraint [re-formulated, employing equation (3') and the fact that $t(y,n) \equiv y - z(y,n)$]:

(8)

$$\int_{\underline{w}}^{\overline{w}} \left[y(w) - U(w) + h[n(w), y(w)/w]\right]dF(w) = R,$$

[12] In the formulation of the welfare function in (6), we take $U(w)$ as the argument; namely, the utility driven by the parent. This utility includes an altruistic component derived from providing consumption to the offspring. One could also include the utility derived by the offspring per-se in the welfare calculus in addition to that of the altruistic parent. This type of double counting would create a positive externality, justifying the subsidization of children. However, as this paper focuses on the re-distributive motive for taxing/subsidizing children, we set aside this alternative motive, without discounting its importance, by 'laundering out' the child utility component.

[13] It is implicitly assumed that the government cannot observe the household's expenditure on education, so the latter cannot be subsidized or taxed. Our results will not change if we allow for such taxation, as long as we plausibly assume that e is a form of anonymous transaction and hence cannot be observed on the individual level. It implies that only linear taxation of e can be used. In such a case, the price of e would be given by $1+t$, where t denotes the unit tax on e. As all households will still be faced with the same (after-tax) prices of e, such linear taxation of e will not affect the redistributive role of non-linear and potentially means-tested child benefits system. For incorporating taxation of child-specific commodities in an optimal tax setting with endogenous fertility, see Cigno (2011).

and, the incentive compatibility constraint:

(9) $U'(w) = -h_2[n(w), y(w)/w]\cdot y(w)/w^2$, for all w.

We next turn to solve the optimization program employing *Pontryagin's* maximum principle. We choose $n(w)$ and $y(w)$ as the two control variables and $U(w)$ as the state variable. Formulating the *Hamiltonian* then yields:

(10) $H = \left[G(U) + \lambda\cdot\left[y - U + h(n, y/w) - R\right]\right]\cdot f - \mu\cdot h_2(n, y/w)\cdot y/w^2,$

where $\mu(w)$ denotes the co-state multiplier and λ is the multiplier associated with the government revenue constraint.

Formulating the necessary first-order conditions for the *Hamiltonian* in (10), employing the *transversality* conditions, one can prove the following proposition (see appendix B for details):

Proposition 1: In the optimal integrated tax/benefit system, total tax liability rises with the number of children (for a given level of pre-tax income).[14]

We obtain a fairly strong result. In a system of child allowance, many may advocate reducing the allowance for each additional child on the grounds of economies of scale in child rearing.[15] Formally, in our setting this would imply that the allowance per additional child; namely, z_n, would decline with n, that is $z_{nn} < 0$. Proposition 1 suggests that z_n itself (not z_{nn}) should be negative. Moreover, suppose that statutorily, the tax/benefit system is separated into an income tax component, $a(y)$, and a means-tested per-child allowance, $k(y,n)$. That is, $t(y,n) = a(y) - k(y,n)\cdot n$. The standard argument of economies of scale in child rearing calls for the average child allowance, k, to decline with the number of children, n. The proposition is in fact stronger, as it calls for total child allowance, kn, to decline with n. This implies that k must decline at a faster rate than the rise in n. That is, the elasticity of the per-child allowance, k, with respect to the number of children, n, is higher than one (in absolute value). We emphasize that we obtain this result even though there are no economies of scale in child rearing in our setting.

The rationale for this result is as follows. In the absence of taxes, low-skill households are faced with a lower opportunity (time) cost of raising children relative to high-skill ones. Hence, they choose to 'specialize' in quantity (number of children),

[14] With a bounded skill distribution, the standard efficiency at the top property continues to hold; namely, the marginal tax on children is zero for the top-earning household.

[15] This is essentially the rationale underlying the common use of equivalence scales.

whereas high-skill households choose to 'specialize' in quality (e.g., education). In a second best setting, (observed) family size may be employed as an indicator of the (unobserved) earning capacity of the household (a 'tagging' device).[16] The negative correlation between family size and ability provides the rationale behind the conventional wisdom calling for subsidizing children on equity grounds. However, in a system in which child benefits can be made means-tested, the government employs a more refined concept of correlation between ability and family size for 'tagging' purposes; namely, the correlation between these two variables, which is conditional on income. To see this, note that for a given level of income, a high-skill household has more leisure than a low-skill one, as it has to work less in order to obtain the same level of income. Hence, conditional on income, a high-skill household has a comparative advantage in raising children over the low-skill household. Thus, conditional on income, the correlation between family size and ability is positive (and not negative as conventional wisdom suggests). In light of the positive correlation between family size and ability (conditional on income), taxing (rather than subsidizing) children at the margin would be socially desirable.[17] [18]

Notice that our prediction differs from the classic result of the redundancy of commodity taxation [Atkinson and Stiglitz (1976)] for the case of homogenous preferences and separable utility. This derives from the fact that, even if one allows for the utility to be separable between leisure and the set of consumption goods (as in the specific functional form examined in appendix A), households are faced with different costs of raising children (due to the fact that raising children is time-consuming and households differ in their earning skills).[19]

4. The Universal Case

In section 3 we have demonstrated, counter to conventional wisdom, that taxing children at the margin would be socially desirable for re-distributive purposes, when

[16] Note that conditioning transfers on family size serves as a *second-best* 'tagging' device because fertility is an endogenous variable in our setting, which responds to financial incentives offered by the government [for recent empirical attempts to estimate the effect of financial incentives on fertility, see Cohen, Dehejia and Romanov (2007) and Laroque and Salanie (2012)].

[17] It is important to emphasize that in equilibrium, high ability households will choose to spend more hours in the labor market and raise a lower number of children, relative to low-ability households. However, our argument suggests that if they mimic the low ability households (an out-of-equilibrium strategy which will not be incentive compatible by construction of our optimal policy rule), then by choosing the same level of income, they will find it relatively cheaper to raise children.

[18] Cigno (2001) and Balestrino et al. (2002) demonstrate the 'tagging' role played by child benefits.

[19] This observation was first made by Cigno (1986).

child benefits are allowed to be means-tested. However, in many countries (in fact, in most developed countries,) benefits are offered on a universal basis and are not subject to means testing. That is, the net income/benefit schedule essentially takes an additively separable form: $z(y,n) = a(y) + b(n)$. Therefore it is of interest and policy relevance to see under what conditions, a universal system can justify subsidizing children at the margin. We attempt to address the following question: starting from any given income tax system, under what conditions will a universal system of child allowances with marginal subsidies be desirable?[20]

To address this issue we must first re-formulate the government optimization program. In this case $y(w)$ is no longer a control variable, but is rather implicitly defined by the first-order condition of the household's utility maximization problem (as a function of family size, n):

(5') $$a_y(y) - h_2(n, y/w)/w = 0.$$

The government then chooses $U(w)$ and $n(w)$ so as to maximize the social welfare function given by (6), subject to the revenue constraint:

(8') $$\int_{\underline{w}}^{\overline{w}} \left[y[n(w)] - U(w) + h\left[n(w), y[n(w)]/w\right]\right] dF(w) = R,$$

and the incentive-compatibility constraint:

(9') $$U'(w) = -h_2\left[n(w), y[n(w)]/w\right] \cdot y[n(w)]/w^2, \quad \text{for all } w.$$

where $y[n(w)]$ is implicitly defined by the first-order condition in (5').

The *Hamiltonian* in this case becomes:

(10')

$$H = \left[G(U) + \lambda \cdot \left[y(n) - U + h[n, y(n)/w] - R\right]\right] \cdot f - \mu \cdot h_2[n, y(n)/w] \cdot y(n)/w^2,$$

where $\mu(w)$ denotes the co-state multiplier and λ is the multiplier associated with the government revenue constraint.

Assuming that the second-order conditions for the government program are satisfied, starting from a system where the marginal child subsidy is set to zero, a necessary and a sufficient condition for the desirability of subsidizing children at the margin is that social welfare will rise by introducing a small marginal child subsidy (thereby increasing the number of children). Formulating the necessary first-order

[20] We will examine, in particular, the desirability of providing a marginal child subsidy, when the labor income tax is set at the optimum.

conditions for the *Hamiltonian* in (10'), employing the *transversality* conditions, the household first order condition [given in (5')] and following some re-arrangements (see appendix C for details) yield the following necessary and sufficient condition for the desirability of subsidizing children at the margin; namely, $b_n > 0$ (some of the arguments of the functions are omitted to abbreviate notation):

(11) $$\left.\frac{\partial H}{\partial n}\right|_{b_n=0} > 0 \Leftrightarrow \left[1-\frac{D(w)}{D(\underline{w})}\right]\cdot\left[\frac{1-F(w)}{f(w)\cdot w}\right]\cdot\left[a_{yy}\cdot y + a_y\right] > (1-a_y),$$

Where the term $D(w) \equiv \frac{1}{1-F(w)}\int_{w}^{\bar{w}} G'[U(t)]dF(t)$ measures the average social marginal utility of income over the interval $[w,\bar{w}]$.

The two first terms in brackets on the left-hand-side of the condition given in (11) are standard in the optimal tax literature and are unambiguously positive [note that $D(w)$ is decreasing with respect to w by virtue of the concavity of G]. The sign of the third term in brackets on the left-hand-side of (11) is however ambiguous. Therefore, the sign of the whole expression on left-hand side of condition (11) is ambiguous too. One can show (see appendix D) that when the marginal child subsidy is set to zero (b_n =0), the third term in brackets [hence, the left-hand side of condition (11)] has the opposite sign of $n'(w)$, which reflects the correlation between earning ability and family size. Thus, when the correlation is negative, the whole expression on the left-hand side is positive, hence, works in the direction of providing a marginal subsidy, and, vice-versa. This left-hand side expression captures, therefore, a 'tagging' component: the existence of a negative correlation implies that by subsidizing children at the margin the government is supporting the low-skill households, thereby enhancing re-distribution. The term on the right-hand-side is the marginal income tax rate, which is exogenously given in our formulation (we examine below also the case where the marginal income tax is set at the optimum). It is plausibly assumed that this term is positive, as our model focuses on the intensive margin of individual labor supply choice; hence, it works in the direction of levying a marginal tax on children. Thus, one cannot a-priori determine the sign of b_n. Naturally, and as is also evident from condition (11), determining whether providing a marginal child subsidy would be socially desirable or not depends on the properties of the income tax schedule.

To gain some intuition, we consider several special cases. Consider first the simple case in which the marginal tax rate is zero for all levels of income (that is, either there is no tax in place, or, a lump-sum tax is being levied). In such a case, $1-a_y=0$ and $a_{yy}=0$. It follows then that $n'(w) < 0$ and the term on the left-hand side of condition (11) is positive. Because the term on the right-hand side of condition (11) vanishes, it follows that providing a marginal child subsidy would be unambiguously socially desirable. The rationale for the clear-cut result obtained for this special case is as follows. In the absence of taxes, low-skill households will have a comparative advantage in raising children, and will hence choose to raise more children than high-skill ones [namely, $n'(w) < 0$]. The negative correlation between earning ability and family size in this case can be employed by the government for re-distributive purposes. Subsidizing children at the margin allows the government to target benefits to low-ability (poor) households, thereby to enhance re-distribution.

We turn next to the case where a flat income tax is in place; namely, $1-a_y>0$ and $a_{yy}=0$. As can be observed from condition (11), both the left-hand side term and the right hand side term are unambiguously positive. Thus, one cannot determine a-priori whether a marginal child subsidy would be desirable. Similar to the case where no tax is in place, the positive sign of the term on the left-hand side derives from the fact that with a flat tax in place, low-skill families still choose to 'specialize' in quantity (namely, $n'(w) < 0$); hence, the government can still employ the ensuing negative correlation between ability and family size for re-distributive purposes by subsidizing children at the margin. However, unlike the case where the marginal income tax rate is zero, the desirability of a marginal subsidy is not forgone conclusion, as the sign of the term on the right-hand side is also positive. This term, which is equal to the marginal income tax rate, reflects the cost associated with a fiscal crowding out effect due to the interaction between the income tax and the child benefit instruments. A child subsidy will induce households to give birth to more children and hence to spend less hours in the labor market. This will reduce the government revenues collected from the income tax system and hence, indirectly, the level of re-distribution. Obviously, when the marginal income tax rate is zero, that is $1-a_y=0$, this term disappears (there is no crowding out effect). In general, this term will work in the direction of levying a tax on children. Thus, although the negative correlation between ability and family size is maintained under a flat (linear) income tax system, one cannot determine a-priori whether a marginal subsidy is desirable or not.

In the two cases examined above the marginal income tax rate is constant across different levels of income. Hence, the term on the left-hand side of (11), which captures the welfare gain from 'tagging', was unambiguously positive. Clearly, this need not be the case with a non-linear income tax system in place. To see this, consider the case where the marginal income tax rate rises with respect to income, that is $a_{yy} < 0$. When the marginal income tax rate rises sufficiently rapidly (that is, a_{yy} is sufficiently negative), then the third term in brackets on the left-hand-side of condition (11), and with it the entire expression on the left-hand-side of this condition, become negative. In such a case, the expression on the left-hand side of (11) will work, all-in-all, in the direction of levying a marginal tax on children. The rationale for this result is as follows. In general, we expect high-ability households to choose a higher level of gross labor income than that chosen by low-ability households. Thus, high-ability households face a higher marginal income tax rate than that faced by low-ability households. When the marginal tax rate rises sufficiently rapidly the difference between the gross levels of income chosen by a low- and a high-skill household will be rather small. In such a case we will be in a scenario similar to the one observed under a means-tested system, where the marginal subsidy was conditional on income, hence the two households were confined to <u>the same</u> gross level of income (that is, a high-skill mimicking household has to choose the same gross level of income as that chosen by his low-skill mimicked counterpart). The patterns of comparative advantage of child-rearing will reverse (as in the means-tested case), and high-ability households will choose to raise more children than low-ability ones (namely, $n'(w) > 0$). The ensuing positive correlation between ability and family size implies that a marginal child tax (rather than a subsidy) would be desirable.

Naturally, in the case where the marginal income tax rate diminishes with respect to income (namely, $a_{yy} > 0$), the net-of-tax wage rate (hence, the gross level of income) of high-ability households is higher than that of low-ability households, with the difference becoming even larger than in the flat-tax case. Hence, the negative correlation between earning ability and family size holding under a flat-tax regime becomes yet stronger in this case. The term on the left-hand side of condition (11) is definitely positive, and hence calls for subsidizing children at the margin as a 'tagging' device. If this effect is stronger than the crowding out effect reflected by the positive term on the right-hand side of condition (11), then a marginal child subsidy is desirable.[21]

To sum up, we have demonstrated that the desirability of subsidizing children at the margin under a universal child allowance system is far from being forgone conclusion and is highly sensitive to the properties of the income tax schedule. Notably, a necessary condition for the desirability of subsidizing children (at the margin) is that the (effective) marginal tax rate does not rise too rapidly (marginal-rate progressivity would be bounded).

We turn next to examine whether the condition in (11) for the desirability of a marginal child subsidy can hold under reasonable parametric assumptions. Saez (2002) approximates the US tax system by a flat tax schedule with a constant marginal tax rate of 40 percent. That is, we set $a_y = 0.6$ and $a_{yy} = 0$. Following Diamond (1998), we assume a single peaked density of skills (a property satisfied by commonly used distributions like the log-normal distribution), which is approximated by a Pareto distribution above the modal skill level. Thus, the term $\frac{1-F(w)}{f(w)\cdot w}$ initially decreases up to the modal skill level and is then constant. It follows that the term $\frac{1-F(w)}{f(w)\cdot w}$ is bounded from below, where the lower bound is given by one over the coefficient of the Pareto distribution. Following Finberg and Poterba (1993), we assume a Pareto coefficient in the range 0.5 to 1.5, which implies that that the lower-bound of the term $\frac{1-F(w)}{f(w)\cdot w}$ varies in the range of 2/3 to 2. Assuming a *Rawlsian* social welfare function implies that $D(w)$=0. Substituting the parametric values into the condition in (11) implies that a marginal subsidy is desirable over the entire range of productivities (wage rates).

Obviously, with a higher marginal tax rate than the one used above, the desirability of providing a marginal subsidy across-the-board (or even providing a subsidy at all) may fail to hold. However, with a flat tax in place and a *Rawlsian*

[21] One may argue that, in practice, most income tax systems exhibit marginal-tax rate progressivity; namely, the statutory marginal tax rate is rising with (gross) income. However, when the bulk of welfare (transfer) programs are means-tested, the effective marginal tax rate at low levels of gross income is relatively high. Therefore, the integrated tax-transfer system exhibits marginal-tax regressivity at the lower end of the income distribution; namely, the effective marginal tax rate is decreasing with (gross) income. Clearly, the marginal tax rates derived in our context are the effective rather than the statutory ones.

government, our parametric assumptions about the skill distribution imply that in general there would be a cutoff level of income, below which a marginal child subsidy would be provided and above which a marginal child tax would be imposed [this follows from the fact that the term $\frac{1-F(w)}{f(w)\cdot w}$ on the left-hand side of (11) is (weakly) decreasing with respect w]. Thus, while in general the desirability of providing a marginal subsidy is ambiguous, it is more likely that such a subsidy would be given to low-skill households.

So far we have examined the desirability of subsidizing/taxing children at the margin, taking the income tax system as given. We turn now to address the same question, while assuming, instead, that an optimal income tax system is in place. Assuming that the marginal child subsidy is set to zero ($b_n = 0$), one can show (see appendix E for details) that the formula for the optimal income tax rule is given by:

$$(12) \qquad \frac{1-a_y}{a_y} = \left[1 - \frac{D(w)}{D(\underline{w})}\right] \cdot \frac{1-F(w)}{f(w)\cdot w} \cdot \left[1 + \frac{1}{\varepsilon_L}\right],$$

where ε_L denotes the labor supply elasticity, given by:

$$\varepsilon_L = -\frac{w \cdot a_y}{y \cdot (n' \cdot h_{12} + h_{22} / w)}.$$

The formula given in (12) is similar to the standard optimal tax formulae available in the literature [see, e.g., Salanie (2003)]. However, there is a subtle difference. In our setting the household's shift in labor supply in response to a change in the after-tax wage rate is accompanied by adjustments in family size and the allocation of time to nurturing activities (as opposed to the single labor-leisure margin which appears in the standard setting). The elasticity given in the expression in (12) takes these adjustments into account. Most estimates of the intensive margin elasticity (hours of work conditional on participating in the labor market) are fairly small [see, e.g., the survey by Blundell, MaCurdy and Meghir (2007)]. However, the notion of elasticity used in our setting captures also traditional 'extensive margin' considerations related to the family composition. Parents will respond to higher anticipated net wage rates by choosing to give birth to a smaller number of children (focusing on their career, instead) and replacing their own (parental-attention) time with paid day-care services. These, in turn, enable the parents to put yet more efforts in their jobs. Thus, the commonly measured intensive margin elasticity underestimates the notion of elasticity used in our setting.

Dividing both sides of the condition given in (11) by a_y, then substituting for the term $(1-a_y)/a_y$ from (12) into (11) and re-arranging, yields the following necessary and sufficient condition for the desirability of subsidizing children at the margin:

$$(13) \qquad \frac{a_{yy} \cdot y}{a_y} > \frac{1}{\varepsilon_L}.$$

We turn next to examine whether the condition in (13) can hold under reasonable parametric assumptions. We assume, as before, a *Rawlsian* government. We further maintain the assumption about a single-peaked skill distribution approximated by a *Pareto* distribution above the modal skill level. Finally, as is common in the literature, we assume constant labor supply elasticity. By virtue of our parametric assumptions, it follows from condition (12) that the optimal marginal tax rate is constant for income levels chosen by households at the higher end of the skill distribution ($a_{yy} = 0$). Substituting into (13) implies that for these households, unambiguously, children should be taxed at the margin!

Turning next to households at the lower end of the distribution, it follows from (12), by virtue of the single-peaked skill-distribution, that the optimal marginal tax rate is declining with income ($a_{yy} > 0$). Thus, both the right-hand side and the left-hand side terms in (13) are positive. The ambiguous result is again due to the already alluded to trade-off between the 'tagging' component on the left-hand side of (13) and the fiscal crowding out effect (on the right-hand side) which is captured by the elasticity component (a higher elasticity, other things equal, implies a higher level of distortion associated with the income tax system; hence, a lower optimal marginal income tax rate). With rapidly declining marginal tax rates and relatively high labor supply elasticity, the 'tagging' component (on the left hand side) will prevail and will all in all, call for subsidizing children at the margin.[22]

With more general welfare functions, the optimal marginal tax rate may well rise with income over some range. Indeed, the literature [see Diamond (1998) and

[22] Notice, that over the income-range where the tax rate is optimally set to be flat, the marginal child subsidy is unambiguously negative (that is, a marginal tax). This is in contrast to the general case, where, with an arbitrary flat-rate system in place, we have shown the result to be ambiguous.

Salanie (2003), amongst others] suggests that the optimal marginal tax schedule is likely to be U-shaped. It then follows from the conditions in (25) and (26) that over the range in which the marginal tax rate increases with income, children should be, unambiguously, taxed at the margin.

5. Conclusion

The economic literature, starting with the seminal contributions of Becker (1960) and Becker and Lewis (1973), viewed household family planning as an economic decision, where the household chooses the number of children to raise and the bundle of goods to consume, so as to maximize its utility. Thus, the size of the household is optimally determined by comparing the costs and benefits associated with raising children.

The literature has emphasized a fundamental trade-off between the quantity (of children) and their quality (e.g., parental investment in education and commodities consumed by the children). Under plausible assumptions (supported by empirical evidence), comparative advantage considerations would induce low-skill (poor) households to specialize in quantity, whereas high-skill (wealthy) households would choose to specialize in quality. The existence of a negative correlation between skill level and family size suggests that in a second-best setting, where the government is unable to observe skill levels directly, (observed) family size could serve as a screening ('tagging') device for re-distributive purposes by an egalitarian government, calling for subsidizing children. Recent literature has called this into question.

In this paper we adopt a benchmark setting in which the quantity-quality paradigm holds; namely, there exists a negative correlation between family size and earning ability. We show, however, that the desirability of subsidizing children in this case is far from being forgone conclusion and hinges on whether child benefits are provided on a universal or a means-tested basis. In fact, we demonstrate that when means-testing is allowed, it is always optimal to tax children at the margin (namely, setting the total child benefits to decline with the number of children), rather than subsidizing them. Under a universal (non means-tested) child-allowance system, subsidizing children may be warranted when the tax-and-transfer system is marginal-tax regressive (namely, when the effective-marginal tax of the integrated tax-and-transfer system decreases with the gross level of income). This is likely to occur at the lower end of the skill distribution, under plausible parametric assumptions.

Our results were derived under the plausible assumption that the quantity-quality paradigm holds, implying a negative correlation between observed family size and unobserved innate ability. Alternatively, assuming that the patterns of specialization are reversed; namely, that there is a positive correlation between family size and ability, will result in reversing our predictions, suggesting that under means-testing, children should be subsidized at the margin (and not taxed as one would predict in light of the positive correlation between family size and ability); whereas, under a universal system taxing children may be desirable. The surprising desirability of taxing children is, therefore, not a byproduct of our (calibrating) parametric assumptions, but is rather derived from the difference between the means-tested system and the universal one, yielding opposite policy implications.

We employ the standard setting used in the optimal taxation literature, where differences in innate earning abilities form a single source of heterogeneity. Clearly, households may differ in other attributes such as preferences for children and nurturing capacity that are likely to affect both the quantity and the quality margins (the choice of family size, time dedicated to nurturing activities and investment in education) and accordingly, the policy recommendations. Rather than attempting to draw direct policy implications, our main message is that one should be cautious in applying simple correlations to derive clear-cut policy recommendations. Calling for taxing children may appear perverse, but what our study suggests is that providing generous child allowances may prove to be counterproductive in promoting re-distributive goals; particularly, when benefits are based on means-testing.

Finally we emphasize that our focus was on re-distributive issues. Thus, even when our analysis suggests unequivocally that children should be taxed at the margin, one has to take into account other important considerations (notably, fertility related externalities that call for subsidizing children at the margin) in order to draw policy recommendations.

Appendix A: The Additively-Separable Case

We show that in the absence of taxes ($z_n = 0$ and $z_y = 1$) when the utility in (1) is taking an additively-separable form, $g(l,n,e,\alpha) = \psi(l) + v(n) + u(e) + \Phi(\alpha)$; and, in addition, satisfies the property that, $-\alpha \cdot \Phi''(\alpha) / \Phi'(\alpha) > 1$, then $\partial[y(w)/w]/\partial w > 0$ and $\partial n(w)/\partial w < 0$.

We first turn to show that $\partial n(w)/\partial w < 0$. The *w*-household is seeking to maximize the utility in (1) subject to the budget constraint in (2). Formulating the household first-order conditions (assuming no taxes in place) yields:

(A1) $1 - \psi'(1 - y/w - n \cdot \alpha)/w = 0$

(A2) $v'(n) - \alpha \cdot \psi'(1 - y/w - n \cdot \alpha) - e = 0$,

(A3) $u'(e) - n = 0$,

(A4) $\Phi'(\alpha) - n \cdot \psi'(1 - y/w - n \cdot \alpha) = 0$.

Substituting from (A1) into (A2) and (A4) yields:

(A2') $v'(n) - \alpha \cdot w - e(n) = 0$,

(A4') $\Phi'(\alpha) - n \cdot w = 0$,

where *e*(*n*) is given by the implicit solution to (A3).

The system of two equations [(A2') and (A4')] implicitly defines the optimal solution for the number of children and the level of parental time invested per child, as a function of the wage rate [*n*(*w*) and $\alpha(w)$]. Fully differentiating the two first-order conditions in (A2') and (A4') with respect to *w* yields:

(A5) $v''(n) \cdot n'(w) - [\alpha + w \cdot \alpha'(w)] - e'(n) \cdot n'(w) = 0$,

(A6) $\Phi''(\alpha) \cdot \alpha'(w) - [n + w \cdot n'(w)] = 0$.

Applying *Cramer's Rule*, one then obtains:

(A7) $$n'(w) = \frac{\alpha \cdot \Phi''(\alpha) + n \cdot w}{[v''(n) - e'(n)] \cdot \Phi''(\alpha) - w^2},$$

where $[v''(n) - e'(n)] \cdot \Phi''(\alpha) - w^2 > 0$, by the household's optimization second order conditions. Substituting for the term $n \cdot w$ from (A4') into the numerator on the right-hand side of (A7) implies that $\partial n(w)/\partial w < 0$ if-and-only-if $\Phi'(\alpha) + \alpha \cdot \Phi''(\alpha) < 0 \Leftrightarrow -\alpha \cdot \Phi''(\alpha)/\Phi'(\alpha) > 1$. The latter follows from our assumption. This completes the first part of the proof.

We turn next to show that $\partial[y(w)/w]/\partial w > 0$. Fully differentiating the condition in (A1) with respect to w yields:

(A8) $\quad 1+\psi''(1-y/w-n\cdot\alpha)\cdot[\partial(y/w)/\partial w+\partial(n\cdot\alpha)/\partial w]=0$.

By the strict concavity of ψ it suffices to show that $\partial(n\cdot\alpha)/\partial w<0$ in order to establish that $\partial[y(w)/w]/\partial w > 0$.

Employing (A4') and (A6) and re-arranging, one obtains:

(A9) $$\partial(n\cdot\alpha)/\partial w = n'(w)\cdot\alpha+\alpha'(w)\cdot n = \frac{n^2+n'(w)\cdot[\Phi'(\alpha)+\alpha\cdot\Phi''(\alpha)]}{\Phi''(\alpha)} < 0.$$

The sign of the inequality follows from the concavity of Φ, the fact that $n'(w)<0$ by the first part of the proof and our assumption that $-\alpha\cdot\Phi''(\alpha)/\Phi'(\alpha)>1$. This concludes the proof.

Appendix B: Proof of Proposition 1

Assuming that the second order conditions for the government optimization are satisfied, to determine whether taxing or subsidizing children at the margin would be optimal it suffices to examine the sign of the derivative of the *Hamiltonian* [given in (10)] with respect to *n*, starting from a system where the marginal child tax/subsidy is set to zero and the income tax is set at the optimum. A positive sign would imply that social welfare will rise by introducing a small marginal child subsidy (thereby increasing the number of children), whereas, a negative sign would suggest taxing children at the margin would be desirable.

Formulating the first-order condition for the *Hamiltonian* in (10) with respect to *U* yields:

(B1) $$\frac{\partial H}{\partial U} = G'(U)\cdot f - \lambda\cdot f = -\mu'.$$

The *transversality* conditions are given by:

(B2) $\quad \mu(\underline{w})=\mu(\overline{w})=0$, ($\lim_{\overline{w}\to\infty}\mu(\overline{w})=0$, when the distribution of skills is unbounded).

Integrating condition (B1), employing the *transversality* condition, $\mu(\overline{w})=0$, yields:

(B3) $$\mu(w)=\int_{w}^{\overline{w}}\left[G'[U(t)]-\lambda\right]dF(t).$$

Employing the second *transversality* condition, $\mu(\underline{w})=0$, yields:

(B4) $\lambda = \int_{\underline{w}}^{\bar{w}} G'[U(t)]dF(t)$.

Now define the function *D* by:

(B5) $D(w) = \frac{1}{1-F(w)} \int_{w}^{\bar{w}} G'[U(t)]dF(t)$.

In words, the function *D* measures the average social marginal utility of income over the interval $[w, \bar{w}]$. Moreover, employing (B3)-(B5) yields:

(B6) $\mu(w) = [1-F(w)]\cdot[D(w) - D(\underline{w})] < 0$

(B7) $\lambda = D(\underline{w})$,

where the negative sign of the expression on the right-hand side of (B6) follows from the fact that *D*(*w*) [defined in (B5)] is decreasing by virtue of the concavity of *G*.

Setting the marginal child tax/subsidy to zero ($z_n = 0$) and differentiating the *Hamiltonian* with respect to *n* yields

(B8) $$\left.\frac{\partial H}{\partial n}\right|_{z_n = 0} = \lambda \cdot f \cdot h_1(n, y/w) - \mu \cdot h_{12}(n, y/w) \cdot \frac{y}{w^2} < 0.$$

The negative sign follows from: (i) the fact that $h_1(n, y/w) = 0$, by virtue of the household's first-order condition in (4), (ii) the fact that $h_{12}(n, y/w) < 0$ by virtue of the assumptions on the utility function ensuring a negative relationship between family size and labor supply, and, (iii) the fact that $\mu(w) < 0$, by virtue of (B6).

Appendix C: Derivation of Equation (11)

In this appendix we derive the necessary and sufficient condition for the social desirability of providing a marginal child subsidy given in equation (11). As in the means-tested case (analyzed in appendix B), assuming that the second order conditions for the government optimization are satisfied, to determine whether taxing or subsidizing children at the margin would be optimal it suffices to examine the sign of the derivative of the *Hamiltonian* [given in (10')] with respect to *n*, starting from a system where the marginal child tax/subsidy is set to zero and the income tax is set at the optimum.

Formulating the necessary first-order condition for the *Hamiltonian* in (10’) with respect to the state variable *U*, employing the *transversality* conditions and re-arranging (repeating the same steps as in appendix B, which are therefore omitted), one obtains:

(C1) $\mu(w) = [1 - F(w)] \cdot [D(w) - D(\underline{w})] < 0$

(C2) $\lambda = D(\underline{w})$,

where $D(w) = \frac{1}{1 - F(w)} \int_{w}^{\bar{w}} G'[U(t)] dF(t)$.

Notice that the conditions in (C1) and (C2) replicate the conditions in (B6) and (B7). The negative sign of the expression on the right-hand side of (C1) follows from the fact that $D(w)$ is decreasing by virtue of the concavity of G.

Setting the marginal child tax/subsidy to zero ($z_n = 0$) and differentiating the *Hamiltonian* with respect to n yields:

$(C3)$

$$\left.\frac{\partial H}{\partial n}\right|_{z_n = 0} = \lambda \cdot f \cdot [y' + h_1 + h_2 \cdot y'/w] - \mu \cdot [h_{12} \cdot y/w^2 + h_{22} \cdot y' \cdot y/w^3 + h_2 \cdot y'/w^2],$$

where we omit some of the arguments to abbreviate notation.

The household's first-order conditions with respect to y and n (replicated for convenience), assuming that the marginal child tax/subsidy is set to zero, are given, respectively, by:

(C4) $a_y + h_2 / w = 0,$

(C5) $h_1 = 0$.

Fully differentiating the condition in (C4) with respect to n yields:

(C6) $a_{yy} \cdot y' + h_{12} / w + h_{22} \cdot y'/w^2 = 0$.

Re-arranging yields:

(C7) $y' = -\frac{h_{12} / w}{a_{yy} + h_{22} / w^2} < 0,$

where the negative sign follows from the fact that $h_{12} < 0$, by our earlier assumptions regarding the household utility function, and the household second-order condition [differentiation of the first-order condition in (C4) with respect to y implies that $a_{yy} + h_{22} / w^2 < 0$]. That is, an increase in the number of children (say in response to offering a marginal subsidy) results in a reduction in the labor supply, and consequently the gross level of income, as expected.

Substituting for λ and μ from (C1) and (C2) into (C3), employing the conditions in (C1)-(C4) and following some re-arrangements yield the following necessary and sufficient condition for the desirability of providing a marginal child subsidy:

$$\text{(C8)} \qquad \left.\frac{\partial H}{\partial n}\right|_{b_n=0} > 0 \Leftrightarrow \left[(1-a_y) + \left[1-\frac{D(w)}{D(\underline{w})}\right]\cdot\frac{1-F(w)}{f(w)\cdot w}\cdot\left[-a_{yy}\cdot y - a_y\right]\right]\cdot y' > 0.$$

By virtue of (C7), the condition in (C8) holds if-and-only if the following condition [equivalent to the one given in equation (11)] is satisfied:

$$\text{(C9)} \qquad \left.\frac{\partial H}{\partial n}\right|_{b_n=0} > 0 \Leftrightarrow \left[1-\frac{D(w)}{D(\underline{w})}\right]\cdot\left[\frac{1-F(w)}{f(w)\cdot w}\right]\cdot\left[a_{yy}\cdot y + a_y\right] > (1-a_y).$$

Appendix D: The Correlation between Family Size and Earning Ability and the Properties of the Income Tax Schedule

In this appendix we state and prove the following claim:

Claim: $n'(w) > = < 0$ if and only if $a_{yy}\cdot y + a_y < = > 0$

Proof: We reproduce, for convenience, the *w*-household's first-order conditions, given in equations (4)-(5), assuming a universal system, namely, $z(y,n) = a(y) + b(n)$ and setting the marginal child subsidy to zero ($b_n = 0$):

(D1) $\quad K(y,n,w) \equiv a_y + h_2 / w = 0,$

(D2) $\quad H(y,n,w) \equiv h_1 = 0.$

The systems of two equations [(D1) and (D2)] provide an implicit solution for *n*(*w*) and *y*(*w*), the optimal choices of the *w*-household.

Fully differentiating the two conditions in (D1) and (D2) with respect to *w* yields:

(D3) $\quad \partial K/\partial y\cdot y'(w) + \partial K/\partial n\cdot n'(w) + \partial K/\partial w = 0,$

(D4) $\quad \partial H/\partial y\cdot y'(w) + \partial H/\partial n\cdot n'(w) + \partial H/\partial w = 0.$

Employing *Cramer's Rule* then yields:

$$\text{(D5)} \qquad n'(w) = \frac{-\partial K/\partial y\cdot\partial H/\partial w + \partial H/\partial y\cdot\partial K/\partial w}{\partial K/\partial y\cdot\partial H/\partial n - \partial H/\partial y\cdot\partial K/\partial n}$$

By the second-order conditions of the household's optimization it follows that $\partial K/\partial y\cdot\partial H/\partial n - \partial H/\partial y\cdot\partial K/\partial n$>0. Thus, it follows:

(D6) $\quad Sign[n'(w)] = Sign[-\partial K/\partial y\cdot\partial H/\partial w + \partial H/\partial y\cdot\partial K/\partial w].$

Differentiating the household's first-order conditions in (D1) and (D2) yields:

(D7) $\partial H / \partial w = -h_{12} \cdot y / w^2$,

(D8) $\partial H / \partial y = h_{12} / w$,

(D9) $\partial K / \partial y = a_{yy} + h_{22} / w^2$,

(D10) $\partial K / \partial w = -h_{22} \cdot y / w^3 - h_2 / w^2$.

Substituting from (D7)-(D10) into the expression on the right-hand side of (D6) and re-arranging yields:

(D11) $$-\partial K / \partial y \cdot \partial H / \partial w + \partial H / \partial y \cdot \partial K / \partial w = \frac{h_{12}}{w^2} \cdot [a_{yy} \cdot y - h_2 / w].$$

The result follows by substituting a_y for the term $-h_2 / w$ by virtue of the first-order condition in (C4) and by noting that $h_{12} < 0$ by our assumption regarding the utility function.

Appendix E: Derivation of Equation (12)

In this appendix we derive the formula for the optimal marginal income tax given in equation (12). The *Hamiltonian* for the government program is given by:

(E1) $$H = \left[G(U) + \lambda \cdot \left[y - U + h[n(y), y / w] - R\right]\right] \cdot f - \mu \cdot h_2[n(y), y / w] \cdot y / w^2,$$

where $n(y)$ is implicitly given by the household's first-order condition in (5).

Formulating the necessary first-order conditions for the *Hamiltonian* in (E1), omitting the arguments to abbreviate notation, yields:

(E2) $$\frac{\partial H}{\partial y} = \lambda \cdot f \cdot \left[1 + h_1 \cdot n' + h_2 / w\right] - \mu \cdot [(h_{12} \cdot n' + h_{22} / w) \cdot y / w^2 + h_2 / w^2] = 0,$$

(E3) $$\frac{\partial H}{\partial U} = G'(U) \cdot f - \lambda \cdot f = -\mu'.$$

The *transversality* conditions are given by:

(E4) $\mu(\underline{w}) = \mu(\overline{w}) = 0$, ($\lim_{\overline{w} \to \infty} \mu(\overline{w}) = 0$, when the distribution of skills is unbounded).

Integrating condition (E3), employing the *transversality* conditions and repeating the steps made in appendix B (which are hence omitted) yield:

(E5) $\mu(w) = [1 - F(w)] \cdot [D(w) - D(\underline{w})] < 0$, (E6) $\lambda = D(\underline{w})$,

where $D(w) = \dfrac{1}{1-F(w)} \int_{w}^{\bar{w}} G'[U(t)]dF(t)$.

The negative sign of the expression on the right-hand side of (E5) follows from the fact that $D(w)$ is decreasing by virtue of the concavity of G.

Substituting from (E5) and (E6) into (E2), employing the household's first-order conditions in (4)-(5) yields, after some re-arrangements, the following expression [which is identical to equation (12)]:

(E7) $$\frac{1-a_y}{a_y} = \left[1 - \frac{D(w)}{D(\underline{w})}\right] \cdot \frac{1-F(w)}{f(w)\cdot w} \cdot \left[1 + \frac{1}{\varepsilon_L}\right],$$

where,

(E8) $$\varepsilon_L = -\frac{w \cdot a_y}{y \cdot (n' \cdot h_{12} + h_{22} / w)}.$$

It remains to show that ε_L denotes the labor supply elasticity.

To see this, let $w_{net} \equiv w \cdot a_y$ denote the after-tax wage rate and let $m \equiv y / w$ denote labor supply. Reproducing the household first order conditions in (4)-(5) obtains:

(E9) $w_{net} + h_2(n,m) = 0,$

(E10) $h_1(n,m) = 0.$

Fully differentiating the system of two equations given by (E9) and (E10) with respect to w_{net} yields:

(E11) $$1 + h_{12}(n,m) \cdot \frac{\partial n}{\partial w_{net}} + h_{22}(n,m) \cdot \frac{\partial m}{\partial w_{net}} = 0,$$

(E12) $$h_{11}(n,m) \cdot \frac{\partial n}{\partial w_{net}} + h_{12}(n,m) \cdot \frac{\partial m}{\partial w_{net}} = 0.$$

Employing *Cramer's Rule* then yields:

(E13) $$\frac{\partial m}{\partial w_{net}} = \frac{-h_{11}}{h_{11} \cdot h_{22} - h_{12}^{\ 2}}.$$

The labor supply elasticity is given by:

(E14) $$\varepsilon_L = \frac{\partial m}{\partial w_{net}} \cdot \frac{w_{net}}{m} = -\frac{h_{11}}{h_{11} \cdot h_{22} - h_{12}^{\ 2}} \cdot \frac{w^2 \cdot a_y}{y}.$$

Differentiating the first order condition in (5) with respect to y and re-arranging yields:

(E15) $n'(y) = \frac{-h_{12}}{w \cdot h_{11}}$.

Substituting for *n*' from (E15) into (E8) and re-arranging yields the expression in (E14). This concludes the derivation of (12).

References

Akerlof, G. (1978). “The Economics of Tagging as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning,” American Economic Review, 68, 8-19.

Atkinson, A. and Stiglitz, J. (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.

______ . (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. (1960). “An Economic Analysis of Fertility,” in Demographic and Economic Change in Developed Countries, National Bureau of Economic Research Series, Number 11, Princeton, NJ: Princeton University Press, 209–231.

______ . (1991). “A Treatise on the Family” (2nd enlarged edition), Harvard University Press, Cambridge, Mass.

Becker, G. and Barro, R. (1988). “A Re-formulation of the Economic Theory of Fertility,” Quarterly Journal of Economics, 103, 1-25.

Becker, G. and Lewis, G. (1973) "On the Interaction between the Quantity and Quality of Children," Journal of Political Economy, 81, 279–288.

Blundell, R., MaCurdy, T. and Meghir, C. (2007) “Labor Supply Models: Unobserved Heterogeneity, Nonparticipation and Dynamics," in: J.J. Heckman & E.E. Leamer (ed.), Handbook of Econometrics, edition 1, volume 6, chapter 69 Elsevier.

Bradshaw, J. and Finch, N. (2002). “A Comparison of Child Benefit Packages in 22 Countries,” Department for Work and Pensions Research Report No.174, Corporate Document Services: Leeds.

Cigno, A. (1986). “Fertility and the Tax-Benefit System: A Reconsideration of the Theory of Family Taxation,” Economic Journal, 96, 1035-1051.

______ . (2001). "Comparative Advantage, Observability, and the Optimal Tax Treatment of Families with Children," International Tax and Public Finance, 8, 455-470.

______ . (2011). “Agency in Family Policy: A Survey,” CESifo Economic Studies, 57, 305-331.

Cigno, A. and Luporini A. (2011). “Optimal Family Policy in the Presence of Moral Hazard when the Quantity and Quality of Children are Stochastic,” CESifo Economic Studies 57, 349-364.

Cohen, A., Dehejia, R. and Romanov, D. (2013). “Financial Incentives and Fertility”, Review of Economics and Statistics, 95, 1-20

Cremer, H., Dellis, A. and Pestieau, P. (2003). "Family Size and Optimal Income Taxation," Journal of Population Economics 16, 37-54.

Diamond, P. (1998). “Optimal Income Taxation: An Example with a U-shaped Pattern of Optimal Marginal Tax Rates,” American Economic Review, 88, 83-95.

Ebert, U. (1992). “A Reexamination of the Optimal Nonlinear Income Tax,” Journal of Public Economics 49, 47–73.

Hanushek, E. (1992) "The Trade-Off between Child Quantity and Quality," Journal of Political Economy, 100, 84-117.

Hotz, V.J., Klerman, J.A. and Willis, R.J. (1997). "The Economics of Fertility in Developed Countries," in M.R. Rosenzweig and O. Stark (editors), Handbook of Population and Family Economics, Elsevier.

Laroque, G. and Salanie, B. (2012). “Identifying the Response of Fertility to Financial Incentives,” forthcoming in the Journal of Applied Econometrics.

Mirrlees, J. (1971) “An Exploration in the Theory of Optimum Income Taxation,” Review of Economic Studies, 38, 175-208.

Moav, O. (2005) "Cheap Children and the Persistence of Poverty," Economic Journal, 115, 88-110.

Saez, E. (2002) "Optimal Income Transfer Programs: Intensive Versus Extensive Labor Supply Responses", Quarterly Journal of Economics, 117, 1039-1073.

Salanie, B. (2003). “The Economics of Taxation”, The MIT Press, Cambridge, Mass.