Blumkin, Tomer; Margalioth, Yoram; Sadka, Efraim

Working Paper
Taxing children: the re-distributive role of child benefits - revisited

CESifo Working Paper, No. 2970

Provided in Cooperation with:
Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Blumkin, Tomer; Margalioth, Yoram; Sadka, Efraim (2010) : Taxing children: the re-distributive role of child benefits - revisited, CESifo Working Paper, No. 2970, Center for Economic Studies and Ifo Institute (CESifo), Munich

This Version is available at:
http://hdl.handle.net/10419/30727

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
Taxing Children: The Re-distributive Role of Child Benefits - Revisited

Tomer Blumkin
Yoram Margalioth
Efraim Sadka

CESifo Working Paper No. 2970
Category 1: Public Finance
February 2010

An electronic version of the paper may be downloaded
• from the SSRN website: www.SSRN.com
• from the RePEc website: www.RePEc.org
• from the CESifo website: www.CESifo-group.org/wp
Taxing Children: The Re-distributive Role of Child Benefits - Revisited

Abstract

In this paper, we challenge the conventional wisdom that due to the negative correlation between family size and earning ability, family size can be used as a 'tagging' device, and calls for subsidizing children (via child allowances) to enhance egalitarian objectives. We show that the case for subsidizing children crucially hinges on child allowances being provided on a universal basis. Notably, when child benefits are means-tested, taxing children at the margin (namely, setting the total benefits to decline with the number of children) is socially optimal.

JEL-Code: D60, H20, H50.

Keywords: child allowance, re-distribution, means-testing, universal, tagging, optimal taxation.

Tomer Blumkin
Department of Economics
Ben-Gurion University
Beer-Sheba 84105, Israel
tomerblu@bgu.ac.il

Yoram Margalioth
The Buchman Faculty of Law
Tel-Aviv University
Tel-Aviv 69978, Israel
margalio@post.tau.ac.il

Efraim Sadka
The Eitan Berglas School of Economics
Tel Aviv University
Tel-Aviv 69978, Israel
sadka@post.tau.ac.il

February, 2010
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 optimal fiscal treatment of children [for the fiscal treatment of the former see, for instance, two recent papers by Cremer et al (2009) and Kleven et al (2009)]. In practice, the existence of children generally reduces the household's tax liability. This may take 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.

The economic rationale underlying the preferential tax treatment of children is based on the following four key arguments. First, the existence of children raises the question of horizontal equity, which in the family size context implies that the tax liability of a household, which is determined based on its ability to pay (say, measured by the level of income), should also account for the (standardized) number of family members. According to this view, children are not a form of consumption good of their parents, but rather part of the tax-paying unit. As such, they reduce the

\(^1\) 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. Thus, for example, in 2009, for the income range of 0-5,970 USD, a household with no children is entitled to a wage subsidy of 7.65%, whereas, a household with one child is eligible for a wage subsidy of 34% (within the same income range households with 2 and 3 children are entitled to a wage subsidy of 40% and 45%, respectively). In this income range, as well as in the income phase-out ranges that are also structured with different rates for different numbers of children, child allowances are means-tested. However, parts of the program are universal. For example, the difference between the allowance of households with 3 children and that of households with 2 children (with the same level of income) is constant for all levels of income above 12,570 USD; and in the overlapping income ranges, in which the child allowance is fixed (the plateaus), the program is obviously universal, as the allowances depend only on the number of children and not on income.
disposable income per-capita, hence the ability-to-pay of the household [for incorporation of horizontal equity considerations into the design of optimal tax-transfer systems see, for instance, Balcer and Sadka (1982) and (1986)]. A second argument draws on demographic considerations (primarily, those related to the looming pension crisis in many countries). The sharp increase in dependency ratios, stemming both from the drop in fertility rates and the corresponding increase in life expectancy, is casting a shadow on the financial sustainability of many national pension systems. The economic rationale for providing child-related subsidies in this case is essentially Pigouvian: subsidies are aimed at internalizing fiscal externalities.\(^2\)

The third argument warrants the provision of child-related subsidies as a means to motivate women to participate in the labor market. This is achieved by the provision of subsidized child care services (as in Sweden), deduction of the costs of daycare,\(^3\) and provision of child tax credits that are limited to mothers’ income tax liability (as in Israel and the UK). The fourth argument, which is the focus of the current paper, justifies the use of child-related subsidies on re-distributive grounds. Family size is used as an efficient indicator ['tagging' device, a la Akerlof (1978)] for the earning capacity of the household. 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

---

2 France, Sweden and Quebec, are notable examples of countries that have implemented policies with the explicit goal of enhancing fertility [see, e.g., Laroque and Salanie (2008)].

3 The cost of daycare is a business related expense of parents with young children who need someone to look after their children when they are at work. It is a business and personal (that is, consumption) mixed cost, as daycare provides value beyond the mere safekeeping of the children. According to basic income tax principles, the business part of it should be deducted in computing the taxable income of the parent (as is done, for example, in Canada and in Germany). Most countries do not allow an outright deduction due to the difficulty of separating the business and consumption elements, but instead reach some sort of a compromise such as the exclusion of employer provided child care (or daycare expenses reimbursed by the employer) from taxable income (as in the US, the Netherlands and Japan), or providing a credit in lieu of deduction (as in France).
to higher-ability households.\(^4\) In such a case, in a second-best setting [a la Mirrlees, (1971)], where earning abilities are un-observed by the government; subsidizing larger families can promote a re-distributive goal.

Indeed, a relatively recent strand in the optimal income tax literature 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). For a comprehensive recent review of the literature, see Cigno (2009). This literature challenges 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 [see Phelps (1973) and Sadka (1976)] when the skill distribution is bounded, the redundancy of commodity taxation [see Atkinson and Stiglitz (1976) and Mirrlees (1976)], as well as, the conventional wisdom in tax policy design that the existence of children merits a reduction in tax liability (that is, children being a tax asset for their parents). The literature emphasizes a key distinction from the standard optimal tax 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 fundamentally different from the ability to earn (the single source of variation across households in the standard optimal tax setting). The introduction of a second source of heterogeneity (alongside variation in earning ability) bears new re-distributive implications, affecting both policy goals and system design. In particular, it is shown 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,

\(^4\) For evidence of the existence of a quality-quantity trade-off see, e.g. Hanushek (1992).
compensate (in utility terms) for their low earning capacity.\footnote{In our setting, we will maintain the standard assumption in the optimal tax literature that individuals only differ in their earning capacity.} Moreover, 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, family size as well as expenditure on child-related goods) are generally shown to depend on both comparative- and absolute-advantage (in domestic vis-à-vis market production) considerations.

In this paper, we address the key policy issue of the optimal tax treatment of children. Employing a continuum version of the two-household framework used by Cigno (1986) and (2001), and Cigno and Pettini (2003), we derive the properties of the general optimal income tax cum child benefit system set by an egalitarian government. 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. The special case of an integrated system comprised of an income tax component, which does not depend on the household's number of children, and a child-benefit component, which does not depend on the household's level of income, will be henceforth referred to as a universal system.

We start by examining the properties of the optimal general system. We show that, counter to conventional wisdom, it is desirable 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). The mechanism at work is associated with the nature of the quality-quantity trade-off faced by the household. 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), whereas high-skill households choose to ‘specialize’ in the quality of children (e.g.,
Therefore, (observed) family size may be employed as an indicator for the (unobserved) earning capacity of the household (a ‘tagging’ device). This 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 can employ a more refined notion of correlation between ability and family size; namely, the correlation between these two variables which is conditional on income. 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, thereby calling for taxing (rather than subsidizing) children at the margin.

Clearly, this somewhat surprising result hinges on the ability of the government to set child benefits that are means-tested. If the government is restricted to a universal system the conditional correlation between ability and family size can no longer be of use. The relevant correlation then becomes the unconditional one. With taxes in place, the latter correlation cannot be unambiguously signed; hence, one cannot determine unequivocally whether children should be taxed (or subsidized) at the margin. Nonetheless, we are able to provide some plausible numerical examples, in which subsidizing children at the margin is socially desirable, in sharp contrast to the general (means-tested) case.

Naturally, a universal system can never do better than a general (means-tested) one. In fact, we are able demonstrate the strict dominance of the means-tested system when the skill distribution is discrete (with any arbitrary finite number of skill levels).
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. In section 4, we compare the general (means-tested) system with the restricted (universal) one. The universal case is discussed in section 5. Section 6 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 employs labor only, and exhibits constant returns to scale and perfect substitution across the various skill levels. Households differ in 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 $[\underline{w}, \bar{w}]$, with a cumulative distribution function $F(w)$ and corresponding densities $f = F'$. 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.\(^6\)

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

$$V(c, l, n, e) = c + h(l) + [v(n) + u(e)];$$

\(^6\) Differences in earning ability are assumed to be the single source of heterogeneity in the economy. We thus refrain from introducing horizontal equity considerations into the analysis.
where \( c \) denotes consumption, \( n \) denotes the number of children, \( e \) denotes the education level per child and \( l \) denotes leisure.\(^7\) We assume that \( v, u \) and \( h \) are strictly concave and strictly increasing, and further assume INADA conditions so that interior solutions are guaranteed throughout.

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 number of children (quantity) or invest in their human capital/education (quality).\(^8,9\) The quasi-linear specification rules out income effects, and is assumed for tractability purposes [see Diamond (1998) and Salanie (2003) for application in the optimal tax literature]. 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 (2007)]. Note, finally, that we follow the standard approach in the endogenous fertility literature and assume that the household can deterministically choose the number of children [for models assuming exogenous fertility see, for instance, Cremer, Dellis and Pestieau (2003)].

Each household is faced with the following budget constraint:

\[
\begin{align*}
c + n \cdot e &= z(y, n); \\
y &= w \cdot (1 - l - n \cdot \alpha),
\end{align*}
\]

\(^7\) Notice that \( e \) is measured per-capita, for simplicity; that is, there are no economies of scale embodied in the consumption of children. In reality, some economies of scale are likely to exist and are often addressed by reference to equivalents scales. Ignoring economies of scale does not affect the qualitative nature of our key arguments. In fact, assuming economies of scale could even strengthen our argument.\(^8\) See, for example, Moav (2005), for a similar setting.\(^9\) The variable \( e \) is interpreted as the level of parental investment in their children’s education, but it may well take alternative interpretations to encompass any commodity consumed by the children (Becker, 1991), the maximized lifetime utility of each child (Becker and Barro, 1988) or old age support expected by parents from each of their children (Cigno, 1993).
where $y$ and $z$ denote gross and net income levels, respectively, and the parameter $\alpha$ measures the fraction of time parents need to allocate to nurturing activities (raising their children).\textsuperscript{10} 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 wealthier households find it more costly to raise children, due to the larger opportunity cost they incur (forgoing time in the labor market). 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. This tax schedule is implicitly defined by the difference between the gross and net income levels, $t(y,n) = y-z(y,n)$. Note that $t(y,n)$ 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 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:

\begin{equation}
U(w) = \max_{n,y,e} \left[ z(y,n) - n \cdot e \right] + h(1 - y/w - n \cdot \alpha) + [v(n) + u(e)]
\end{equation}

The first-order-conditions for the typical $w$-household’s optimal choice are given by:

\textsuperscript{10}We simplify by implicitly assuming that the (time) cost of raising a child cannot be replaced by day-care services. Our results would remain valid if we allowed for replacement of parents’ time by paid child-care services, as long as parents maintained some role in raising their children, which is obviously the case in reality.
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.  

It is straightforward to verify (see appendix A for details) that in the absence of any form of government intervention; namely, when \( z = y \), hence, \( z_y = 1 \) and \( z_n = 0 \), the model yields the plausible result suggested by the quantity-quality paradigm: poor families will ‘specialize’ in quantity and hence choose to have a larger number of (less educated) children. The opposite will hold true for wealthy families: they will ‘specialize’ in quality (educating their offspring). This key observation will later play a crucial role in the design of the welfare system. We next turn to characterize the properties of the integrated income tax cum child benefit system.

3. The General (Means-Tested) System

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

\[
W = \int_{\mathbb{R}} G[U(w)]dF(w);
\]

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

\[11\] 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; notice that in the two-type case bunching (that is, pooling) will never be part of the optimal solution as shown by Stiglitz (1982)].
\[ \int_{w} t[y(w), n(w)]dF(w) = R; \]

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)-(6); 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 Mirrlees (1971) and (1976) and Salanie (2003), we reformulate the government optimization problem (see appendix B for details) as choosing the functions \( U(w), n(w) \) and \( y(w) \), so as to maximize the social welfare function in equation (7), subject to the revenue constraint:

\[ \int_{w} [y(w) - U(w) - n(w) \cdot e[n(w)] + h[1 - y(w)/w - \alpha \cdot n(w)] + \nu[n(w)] + u[e[n(w)]]]dF(w) = R, \]

and the incentive compatibility constraint:

\[ U'(w) = h'[1 - y(w)/w - n(w) \cdot \alpha] \cdot y(w)/w^2, \quad \text{for all } w, \]

where \( e[n(w)] \) is implicitly defined by the first-order condition in (6).

---

\[ ^12 \] In the formulation of the welfare function in (10), 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 [a type of altruism a la Barro (1974) rather than joy-of-giving as in Andreoni (1990)]. 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. For incorporating taxation of child-specific commodities in an optimal tax setting with endogenous fertility, see Cigno (2009).
It is useful to point out that we do not directly derive the integrated net-income function $z(y, n)$. We rather derive the optimal functions $y(w), n(w), e(w)$ and $U(w)$; and then calculate $z[y(w), n(w)]$, the net income, employing condition (3):

$$z[y(w), n(w)] = U(w) + n(w) \cdot e(w) - h[1 - y(w) / w - \alpha \cdot n(w)] - v[n(w)] - u[e(w)].$$

Note that in this way, we can only define the net income function, $z$, and the tax function, $\tau = y - z$, at bundles $[y(w), n(w), e(w)$ and $U(w)]$ that are actually observed (chosen by individuals) in the optimal solution. Thus, $z(y, n)$ is not well defined elsewhere. Therefore, strictly speaking, one cannot directly derive the marginal income tax rate, $1 - z_y$, and the marginal child benefit, $z_n$. Instead, as is common in the literature, we define these rates through the relevant individual marginal rates of substitution. Indeed, $z_n$ is defined as the marginal rate of substitution of net income ($c + ne$) for children ($n$) and $z_y$ is defined as the marginal rate of substitution of net income for gross income ($y$). Formally, using the individual first-order conditions in (4) and (5), yields:

$$z_n = -v' + \alpha \cdot h' + e$$

and

$$z_y = h' / w.$$

We next turn to solve the optimization program as an optimal control problem 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:

$$H = [G(U) + \lambda \cdot [y - U - n \cdot p \cdot e(n) + h(1 - y / w - n \cdot \alpha) + v(n) + u[e(n)] - R] \cdot f + \mu \cdot [h' (1 - y / w - n \cdot \alpha) \cdot y / w^2],$$
where $\mu(w)$ denotes the co-state multiplier, $\lambda$ is the multiplier associated with the government revenue constraint and $e(n)$ is given by the implicit solution to the first-order condition in (6).

The necessary first-order conditions are:

\begin{equation}
\frac{\partial H}{\partial n} = \left[ -\lambda \cdot p \cdot e - \lambda \cdot n \cdot p \cdot e'(n) - \lambda \cdot h'(1 - y / w - n \cdot \alpha) \cdot \alpha + \lambda \cdot v'(n) + \lambda \cdot u'(e) \cdot e'(n) \right] \cdot f \\
- \mu \cdot h''(1 - y / w - n \cdot \alpha) \cdot \frac{\alpha \cdot y}{w^2} = 0,
\end{equation}

\begin{equation}
\frac{\partial H}{\partial y} = \left[ \lambda - \lambda \cdot h'(1 - y / w - n \cdot \alpha) \cdot 1 / w \right] \cdot f \\
+ \mu \cdot [-h''(1 - y / w - n \cdot \alpha) \cdot y / w^3 + h'(1 - y / w - n \cdot \alpha) / w^2] = 0,
\end{equation}

\begin{equation}
\frac{\partial H}{\partial U} = G'(U) \cdot f - \lambda \cdot f = -\mu'.
\end{equation}

The transversality conditions are given by:

\begin{equation}
\mu(w) = \mu(\bar{w}) = 0, \quad \lim_{w \to \infty} \mu(\bar{w}) = 0, \text{ when the distribution of skills is unbounded}).
\end{equation}

Integrating condition (17), employing the transversality condition, $\mu(\bar{w}) = 0$, yields:

\begin{equation}
\mu(w) = \int_{\bar{w}}^{w} G'[U(t)] - \lambda \right] dF(t).
\end{equation}

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

\begin{equation}
\dot{\lambda} = \int_{\bar{w}}^{w} G'[U(t)] dF(t).
\end{equation}

Now define the function $D$ by:

\begin{equation}
D(w) = \frac{1}{1 - F(w)} \int_{\bar{w}}^{w} G'[U(t)] dF(t).
\end{equation}
In words, the function $D$ measures the average social marginal utility of income over the interval $[w, \bar{w}]$. By virtue of the concavity of $G$, it follows that $D(w)$ is decreasing. Moreover, employing (19) and (20) yields:

(22) \[ \mu(w) = [1 - F(w)] \cdot [D(w) - D(w)], \]

(23) \[ \lambda = D(\bar{w}). \]

Substituting from (22) and (23) into (15), employing the first-order conditions in (4) and (6), after some re-arrangements yields:

(24) \[ z_n[y(w), n(w)] = \left[ 1 - \frac{D(w)}{D(\bar{w})} \right] \cdot \frac{1 - F(w)}{f(w)} \cdot h''[1 - y(w)/w - n(w) \cdot \alpha] \cdot \frac{\alpha \cdot y(w)}{w^2}. \]

Strikingly, because $h'' < 0$, the optimal condition in (24) suggests that $z_n < 0$. That is, children should be taxed at the margin. Formally, 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).

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.\(^{14}\) Formally, in our setting this would imply that the allowance per additional child; namely, $z_n$, would decline with $n$, that is $z_{mn} < 0$.

Proposition 1 suggests that $z_n$ itself (not $z_{mn}$) 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, $kn$, \footnote{This is essentially the rationale underlying the common use of equivalence scales.}

\[ z_n[y(w), n(w)] = \left[ 1 - \frac{D(w)}{D(\bar{w})} \right] \cdot \frac{1 - F(w)}{f(w)} \cdot h''[1 - y(w)/w - n(w) \cdot \alpha] \cdot \frac{\alpha \cdot y(w)}{w^2}. \]
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), 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).\(^\text{15}\) 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

---

\(^\text{15}\) 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 (2008)].
between family size and ability (conditional on income), taxing (rather than subsidizing) children at the margin would be socially desirable.\footnote{16}

We turn next to graphically demonstrate our surprising result of the desirability of taxing children. For this purpose, recall that we used the first-order condition given in (6) to eliminate \(e\), the level of parental investment in per-child education, from the government optimization program. This procedure essentially defines a restricted indirect utility function (where the individual utility maximization is conducted with respect to \(e\) only). Define this function by:

\[
J(w, z, y, n) = \max \left[ z - n \cdot e + v(n) + u(e) + h(1 - y / w - n \cdot \alpha) \right].
\]

Fixing the gross level of income and employing the envelope theorem, we derive the marginal rate of substitution between net income and the number of children (conditional on the gross level of income, \(y\)):

\[
MRS_{z,n}(w, z, y, n) = \frac{-\partial J(w, z, y, n)}{\partial n} = \frac{\partial J(w, z, y, n)}{\partial z} = e - v'(n) + \alpha \cdot h'(1 - y / w - n \cdot \alpha).
\]

Differentiating the \(MRS\) condition in (26) with respect to \(n\) (along the indifference curve) yields, after re-arrangement:

\[
\frac{\partial}{\partial n} \left[ \frac{-\partial J / \partial n}{\partial J / \partial z} \right] / \partial n = \text{const} =
\]

By differentiation of the indirect utility in (25), it follows that \(\partial J / \partial z = 1\) (hence, \(\partial^2 J / \partial z^2 = 0\)) and \(\partial^2 J / \partial n \partial z = 0\). Substitution into (27) then yields:

\footnote{16}{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.}
\[ \frac{\partial}{\partial n} \left( \frac{\partial J/\partial n}{\partial J/\partial z} \right) \bigg|_{J=\text{const}} = -\frac{\partial^2 J}{\partial J/\partial z}, \]

where the inequality follows by virtue of the second order conditions for the \( w \)-household optimization \((\partial^2 J/\partial n^2 < 0)\).

We thus conclude that the indifference curve of a \( w \)-household in the \( n-z \) space (for a given level of gross income, \( y \)) is U-shaped. Moreover, the slope of the indifference curve is decreasing with respect to \( w \). This implies a single crossing property.\(^{17}\) We let the bundle \((z_i, y_i, n_i) = [z(w_i), y(w_i), n(w_i)]\), denote the choice of a \( w_i \)-household in the optimal solution for the government problem. Let \( J(w_i, z, y_i, n) \) describe the indifference curve of a \( w_i \)-household in the \( n-z \) space (where the level of gross income is fixed at \( y_i \)), which passes through the bundle \((n_i, z_i)\). We now show that the bundle \((n_i, z_i)\) must lie on the declining portion of the indifference curve \( J(w_i, z, y_i, n) \).

Suppose to the contrary that \((n_i, z_i)\), as depicted in figure 1 below, lies on the rising portion of the indifference curve. Let \( J(w_2, z, y_i, n) \), as depicted in figure 1, describe the indifference curve of some \( w_2 \)-household that passes through \((n_i, z_i)\), where \( w_2 > w_1 \). For concreteness, suppose that the intersection of the two curves occurs on the declining portion of the indifference curve of the \( w_2 \)-household.\(^{18}\)

\(^{17}\) To see this, note that the optimal level of education is given by: \( u'(e) - n = 0 \), which implies that \( e \) is independent of \( w \) conditional on \( n \). Differentiation with respect to \( w \) then yields: \( \partial \text{MRS}_{z,n} / \partial w = \alpha \cdot h''(1 - y/w - n \cdot \alpha) \cdot y/w^2 < 0 \).

\(^{18}\) The fact that the intersection occurs on the declining portion of the indifference curve of the high-skill household is non-essential. A similar argument would apply for the other case and is hence omitted.
Consider a downward shift along the indifference curve of the \( w_1 \)-household [to the bundle \((n'_1, z'_1)\)]. By construction, this would maintain the same level of utility for this household. Moreover, this will not violate the incentive-compatibility constraint of the higher-ability household (the \( w_2 \)-household), as the new bundle lies below the indifference curve of this household. At the same time, the net income, \( z \), of the \( w_1 \)-household would fall. Recalling that the gross income is kept constant at the level of \( y_1 \), this would imply that the tax liability of the \( w_1 \)-household would rise. Thus, we were able to show that the government can increase its revenues without reducing the utility of any household. This yields the desired contradiction.\(^{19}\)

\(^{19}\) Notice that we have demonstrated that subsidizing children at the margin would be socially undesirable. The result that it would be actually optimal to tax children at the margin, derives from the fact that when the marginal tax/subsidy is set to zero, a small increase in the marginal tax will have no effect on government revenues (to the first order) but will mitigate the incentive compatibility constraint of the higher ability household. This will allow the (egalitarian) government to enhance redistribution.
Turning back to the optimal condition for the marginal tax in (24), it is straightforward to verify that when $\alpha = 0$, the marginal tax (or subsidy) on children is set to zero, namely, we replicate the classic result of the redundancy of commodity taxation [Atkinson and Stiglitz (1976)]. In this particular case, high-ability and low-ability households faced with the same budget will choose the same consumption allocation (by virtue of the separability of the utility function with respect to leisure), as they are faced with the same costs of raising children (the cost of educating them in our context, which does not depend on earning ability). It is also easy to verify that the standard zero tax at the top [Phelps (1973) and Sadka (1976)], when the distribution of skills is bounded from above, and at the bottom [Seade (1977)] apply.

Finally, for the sake of completeness, we can derive a formula for the optimal income tax rule in our case, which is similar to the one commonly available in the literature [e.g., Salanie (2003)]:

\[
\frac{1 - z_y}{z_y} = \left[1 - \frac{D(w)}{D(w^*)}\right] \cdot \frac{1 - F(w)}{f(w) \cdot w} \cdot \left[1 + \frac{1}{\varepsilon_L}\right],
\]

where $\varepsilon_L$ denotes the labor supply elasticity, given by:

\[
\varepsilon_L = -\frac{1}{h' \cdot (1 - y \cdot w - n \cdot \alpha)} \cdot \frac{w^2 \cdot z_y}{y}.
\]

4. **Means-Tested versus Universal Systems**

Naturally, a general (means-tested) tax/benefit system cannot do worse than a universal one, which is confined to a separable form (between an income tax and child benefit components). An interesting question is whether a universal system (which is fairly prevalent in many OECD countries) can nevertheless suffice in certain conditions to attain the social optimum. We are able to show that this is never the case
when the distribution of skill levels is discrete. Formally, we state and prove the following proposition:

**Proposition 2:** When the distribution of skills is discrete, any universal system can be replaced by a means-tested system that attains a higher level of social welfare.

**Proof:** See appendix C.

5. **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 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 would introducing a universal system of child allowances with marginal subsidies be desirable?

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:

\[
(5') \quad a_y(y) - 1/w \cdot h'(1 - y/w - n \cdot \alpha) = 0.
\]

The government then chooses \( U(w) \) and \( n(w) \) so as to maximize the social welfare function given by (7), subject to the revenue constraint:

\[
(9') \quad \int \left[ y[n(w)] - U(w) - n(w) \cdot e[n(w)] + h[1 - y[n(w)]/w - \alpha \cdot n(w)] + v[n(w)] + u[e[n(w)]] \right] dF(w) = R,
\]
and the incentive-compatibility constraint:

\[ U'(w) = h'[1 - y[n(w)]/w - n(w) \cdot \alpha] \cdot y[n(w)]/w^2, \] for all \( w, \)

where \( y[n(w)] \) and \( e[n(w)] \) are implicitly defined by the first-order conditions in (5') and (6), respectively.

The Hamiltonian in this case becomes:

\[ (14') \]

\[ H = [G(U) + \lambda \cdot (y(n) - U - n \cdot p \cdot e(n) + h[1 - y(n)/w - n \cdot \alpha] + v(n) + u(e(n))] - R] \cdot f \\
+ \mu \cdot \left[ h'[1 - y(n)/w - n \cdot \alpha] \cdot y(n)/w^2 \right] \]

where \( \mu(w) \) denotes the co-state multiplier and \( \lambda \) is the multiplier associated with the government revenue constraint. The first-order conditions are given by:

\[ (31) \]

\[ \frac{\partial H}{\partial n} = \left[ \lambda \cdot y'(n) - \lambda \cdot p \cdot e - \lambda \cdot n \cdot p \cdot e'(n) + \lambda \cdot h'(1 - y/w - n \cdot \alpha) \cdot \alpha \cdot y(n)/w^2 \right] + \lambda \cdot v'(n) + \lambda \cdot u'(e) \cdot e'(n) \cdot f + \mu \cdot \left[ -h''(1 - y/w - n \cdot \alpha) \cdot \frac{\alpha \cdot y(n)}{w^2} \right] \\
- h''(1 - y/w - n \cdot \alpha) \cdot \frac{y(n)}{w^2} \cdot y'(n) + h'(1 - y/w - n \cdot \alpha) \cdot \frac{y'(n)}{w^2} = 0, \]

\[ (32) \]

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

The transversality conditions are given, as before, by:

\[ (33) \]

\[ \mu(w) = \mu(\bar{w}) = 0 \quad (\lim_{w \to \infty} \mu(w) = 0, \text{ when the distribution of skills is unbounded}). \]

Substituting the individual first-order conditions (4)-(6) into (31) and rearranging yields:

\[ b_n(a) \cdot \lambda \cdot f = \lambda \cdot f \left[ 1 - a_s \cdot (y(n)) \right] \cdot y'(n) + \mu \left[ -h''[1 - y(n)/w - n \cdot \alpha] \cdot \frac{\alpha \cdot y(n)}{w^2} \right] \\
- h''[1 - y(n)/w - n \cdot \alpha] \cdot \frac{y(n)}{w^2} \cdot y'(n) + h'[1 - y(n)/w - n \cdot \alpha] \cdot \frac{y'(n)}{w^2} \]
Following some algebraic manipulations (which replicate those carried out in the general case, and are hence omitted) yields the following expression for the optimal marginal child subsidy (where some of the arguments of the functions are omitted to abbreviate notation):

\[
(35) \quad b_n = (1-a_y) \cdot y' + \left[ 1 - \frac{D(w)}{D(w)} \right] \cdot \frac{1-F(w)}{f(w) \cdot w} \cdot \left[ h'' \cdot \frac{\alpha \cdot y}{w} + h'' \cdot \frac{y}{w^2} \cdot y' - a_y \cdot y' \right].
\]

Notice, that as in the general (means-tested) case, one cannot directly derive the marginal child benefit, \( b_n \). Instead, we define, as before, the marginal child benefit as the marginal rate of substitution of net income (\( c + ne \)) for children (\( n \)). Formally, using the individual first-order condition in (4), yields:

\[
(12') \quad b_n = -y' + \alpha \cdot h' + e.
\]

It follows by full differentiation of the first-order condition in (5’) with respect to \( n \) that:

\[
(36) \quad y'(n) = \frac{-\alpha \cdot h''}{h'' / w + a_{yy} \cdot w}.
\]

Employing then the second-order-condition [differentiation of the first-order condition in (5’) with respect to \( y \)], it follows that:

\[
(37) \quad a_{yy} + h'' / w^2 < 0.
\]

Thus, by virtue of (36), \( y'(n) < 0 \). 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 \( y'(n) \) from equation (36) into equation (35) and re-arranging yields:

\[
(38) \quad b_n = \left[ (1-a_y) + \left[ 1 - \frac{D(w)}{D(w)} \right] \cdot \frac{1-F(w)}{f(w) \cdot w} \cdot \left[ a_{yy} \cdot y - a_y \right] \right] \cdot y'.
\]
Assuming that the second-order conditions for the government program are satisfied, a necessary and sufficient condition for the desirability of subsidizing children at the margin, $b_n > 0$, is:

\[
(39) \quad \frac{\partial H}{\partial n} \bigg|_{b_n = 0} > 0 \iff \left(1 - a_y \right) + \left[1 - \frac{D(w)}{D(w)} \right] \cdot \frac{1 - F(w)}{f(w)} \cdot \frac{w}{w} \left[-a_{ys} \cdot y - a_y\right] \cdot y' > 0.
\]

Namely, starting from a system where the marginal child subsidy is set to zero, social welfare will rise by introducing a small marginal child subsidy (thereby increasing the number of children).

By virtue of the fact that $y' < 0$, the condition in (39) holds if-and-only-if:

\[
(40) \quad \left[1 - \frac{D(w)}{D(w)} \right] \cdot \left[1 - \frac{F(w)}{f(w)} \cdot \frac{w}{w}\right] \cdot \left[a_{ys} \cdot y + a_y\right] > (1 - a_y).
\]

The two first terms in brackets on the left-hand-side of equation (40) are positive [recall that $D(w)$ is decreasing with respect to $w$]; hence, they work in the direction of providing a marginal child subsidy. The sign of the third term in brackets on the left-hand-side of (40) is however ambiguous. Therefore, the sign of the left-hand side of condition (40) 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 (40)] has the opposite sign of $n'(w)$, which reflects the correlation between earning ability and family size. The term on the right-hand-side is the marginal income tax rate, which is exogenously given in our formulation. 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 (40), 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 (40) is positive. Because the term on the right-hand side of condition (40) 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 (40), 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 \(1 - \alpha_y = 0\), 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 (40), 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 (40), 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 (40) 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 will rise sufficiently rapidly, the net-of-tax
wage rate of high-ability households may fall below that of low-ability households. When the net-of-tax wage rate of high-ability households will be sufficiently smaller than that of low-ability ones, the patterns of comparative advantage of child-rearing will reverse, 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 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 becomes yet stronger. The term on the left-hand side of condition (40) 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 (40), then a marginal child subsidy is desirable.

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.

We resort next to numerical simulations to examine whether the condition in (40) for the desirability of a marginal child subsidy can hold under reasonable parametric assumptions. Saez (2002) approximates the US tax system by a linear 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, 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 (40) implies that a marginal subsidy is indeed desirable over the entire range of productivities (wage rates).

Notice that, as pointed out by Saez (2001), the correct Pareto coefficient should be the one associated with (empirically unobserved) productivity distribution rather than that associated with observed income distribution. As suggested by Saez (2001), for the case of the Pareto distribution, assuming an iso-elastic labor supply, the coefficient associated with the productivity distribution ($a^p)$ is related to that associated with income distribution ($a^i$) by the following formula: $a^i \cdot (1+\varepsilon_L) = a^p$, where $\varepsilon_L$ denotes the (uncompensated) labor supply elasticity. Thus, a more conservative estimate of the lower bound of the term $\frac{1-F(w)}{f(w) \cdot w}$ would be slightly lower than the values in the range specified above (in light of the fact that labor supply is fairly inelastic according to empirical findings). This would suggest that at least for the lower end of the productivity distribution, a marginal subsidy is fairly plausible. In light of our parametric assumption, there would be a cutoff level of productivity, below

---

20 Most estimates of the intensive margin elasticity (hours of work conditional on participating in the labor market) are small. See, for example, the survey by Blundell and MaCurdy (1999).
which a marginal child subsidy would be provided and above which a marginal child
tax would be imposed.

To see this, notice first that with a flat income tax schedule, the expression in
equation (36) reduces to:

$$y'(n) = -\alpha \cdot w.$$  

(41)

Then, assuming that the marginal tax rate is $0 < t < 1$ and a *Rawlsian* social welfare
function, the optimal marginal child subsidy/tax in (38) is given by:

$$b_n = -\alpha \cdot w \cdot t + \alpha \cdot \left(1 - \frac{F(w)}{f(w)}\right) \cdot (1 - t).$$

(42)

Thus,

$$b_n > 0 \iff -\alpha \cdot w \cdot t + \alpha \cdot \left(1 - \frac{F(w)}{f(w)}\right) \cdot (1 - t) > 0 \iff \frac{1 - F(w)}{f(w) \cdot w} > t/(1 - t).$$

(43)

Note that the term $\frac{1 - F(w)}{f(w) \cdot w}$ is sufficiently high for sufficiently low $w$, and is decreasing
over the range of skills up to the modal skill level and is then constant. Thus $b_n$ is
positive (a marginal child subsidy) at the lower end of the wage distribution and may be
negative (a marginal child tax) at the higher end of the wage distribution.

Another interesting observation follows from the expression for the optimal
marginal subsidy/tax in equation (42). As the expression on the right-hand side of this
equation is decreasing in $w$ for the lower end of the wage distribution [recall that we
assume, following Diamond (1998), a single-peaked density of skills], then as low-
ability (poor) households tend to raise more children (due to their comparative
advantage), it follows that the marginal subsidy increases with respect to the number of
children, for households with a sufficiently large number of children. Strikingly, such a
system of subsidies has been in place in Israel since 1975 but allowances were mostly
flattened (subject to some grandfathering clauses) in 2003. Indeed, Israel has a fairly
generous universal (non-means-tested) child allowance system. A notable feature of the program until June 2003 was that the size of the allowance per child increased substantially with the birth order of the child, with the first two children up to the age of 18 receiving minimal benefits, and each child from the fourth on receiving a large benefit.\(^{21}\)

6. 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. Conventional wisdom therefore suggests that in a second-best setting, the (observed) family size could be used as a screening (‘tagging’)

\(^{21}\) Such a structure went against standard arguments of increasing returns to scale in child rearing. In 2001, the allowances for children from the fifth up were close to doubled as a result of a strong ultra-orthodox lobby in the Israeli parliament. The ultra-orthodox (and Muslims) in Israel traditionally give birth to a larger number of children, rendering them the primary beneficiaries of such a non-linear pattern of subsidies. This controversial legislation sparked off a major public outcry that resulted in a backlash in 2003, equalizing the per child allowance for all children. What we show in the current setting is that there could be in fact a re-distributive argument supporting this patently counter-intuitive pattern.
device for re-distributive purposes by an egalitarian government, and call for subsidizing children.

In this paper, we challenge this conventional wisdom. Specifically, we show that subsidizing children may indeed be warranted under special circumstances, provided that child allowances are universal (non means-tested). However, when means-testing is allowed, it is optimal to tax children at the margin (namely, setting the total child benefits to decline with the number of children), rather than to subsidize them. Thus, deciding whether to subsidize or to tax children crucially hinges on whether child benefits are provided on a universal or a means-tested basis.
Appendix A: Demonstration of the Quantity-Quality Trade-off

In this appendix we show that in the absence of government intervention, poor (low-skill) families will choose to 'specialize' in quantity; whereas, wealthy (high-skill) households will choose to 'specialize' in quality.

Substituting for $h'()$ from (5) into (4) and (6) and setting $z_n = 0$ and $z_y = 1$ yields:

(A1) $v'(n) - \alpha \cdot w - e = 0$,

(A2) $u'(e) - n = 0$.

The system of two equations [(A1) and (A2)] implicitly define the optimal solution for the number of children and the level of education as a function of the wage rate [$n(w)$ and $e(w)$]. Fully differentiating the two first-order conditions in (A1) and (A2) with respect to $w$ yields:

(A3) $v''(n) \cdot n'(w) - \alpha - e'(w) = 0$,

(A4) $u''(e) \cdot e'(w) - n'(w) = 0$.

Applying Cramer’s Rule, one then obtains:

(A5) $n'(w) = \alpha \cdot u''(e) / \Delta; \quad e'(w) = \alpha / \Delta$,

where $\Delta \equiv v''(n) \cdot u''(e) - 1 > 0$, by the second order conditions.

It therefore follows that $n'(w) < 0$ and $e'(w) > 0$ (the former follows from the strict concavity of $u$).
Appendix B: Reformulation of the Government Optimization Problem

In this appendix we derive the expressions for the revenue constraint [equation (9)] and the incentive constraint [equation (10)] that appear in the re-formulated government problem in the main body of the text.

Following the standard approach in the optimal tax literature, the government is essentially offering the individuals with a tax schedule given by the triplet of functions, $z(w), y(w)$ and $n(w)$, that is, bundles comprised of the net-income, gross income and the number of children, from which each household is self-selecting the optimal bundle. The total tax revenues collected given the optimal choices taken by the individuals satisfy the government (pre-determined) revenue needs.

We first turn to derive the incentive constraint given by equation (10). We let $J[w, z, y, n]$ denote the maximal level of utility derived by a household with ability level $w$, net income level $z$, gross income level $y$ and number of children $n$. Formally,

\[
J[w, z, y, n] = \max_{w, z, y, n} \left[ z - n \cdot e + v(n) + u(e) + h(1 - y / (w - n \cdot \alpha)) \right].
\]

By definition of the indirect utility function in (3) it follows that:

\[
U(w) = \max_{z, y, n} J[w, z, y, n].
\]

Denoting by $z(w), y(w)$ and $n(w)$, the optimal choices of the $w$-household, it follows that:

\[
U(w) \equiv J[w, z(w), y(w), n(w)].
\]

Fully differentiating the identity in (B3) with respect to $w$ yields:

\[
\frac{dU(w)}{dw} = \frac{\partial J}{\partial w} \cdot z'(w) + \frac{\partial J}{\partial y} \cdot y'(w) + \frac{\partial J}{\partial n} \cdot n'(w).
\]

The $w$-household's incentive constraint is defined by the following condition:

\[
w = \arg \max_{w'} J[w, z(w'), y(w'), n(w')].
\]
In words, the \( w \)-household has no incentives to mimic other types; hence it will choose to reveal its true type (rather than to pretend to be some other type, \( w' \)). Re-formulating the condition in (B5) as a first-order condition yields:

\[
(B6) \quad \frac{\partial J}{\partial z} \cdot z'(w) + \frac{\partial J}{\partial y} \cdot y'(w) + \frac{\partial J}{\partial n} \cdot n'(w) = 0.
\]

Substituting from (B6) into (B4) then yields:

\[
(B7) \quad U'(w) = \frac{\partial J[w, z(w), y(w), n(w)]}{\partial w}.
\]

Thus, the condition in (B6) holds if-and-only-if the condition in (B7) holds. Employing (B1) yields then the following expression, which is identical to equation (10):

\[
(B8) \quad U'(w) = h'(1 - y(w)/w - n(w) \cdot \alpha) \cdot y(w)/w^2.
\]

We turn next to the government revenue constraint, given by equation (9). For convenience, we modify the tax schedule offered by the government [given by the triplet of functions, \( z(w), y(w) \) and \( n(w) \)] by replacing the function \( z(w) \), the net income level, with the function \( U(w) \), the indirect utility level in the individual optimization.\(^{22}\) Recalling that the tax function is implicitly defined by the difference between the gross income and the net income, \( t(y, n) = y - z(y, n) \), the revenue constraint in (8) can be re-written as follows:

\[
(B9) \quad \int_{w}^{\bar{w}} [y(w) - z[y(w), n(w)]]dF(w) = R.
\]

\(^{22}\) Notice, that, given the optimal choice of the level of education as a function of the number of children [determined by the first-order-condition in (6)], the triplet \(< z(w), y(w), n(w) >\) uniquely determines the level of utility, \( U(w) \), by virtue of condition (3). Thus, our transformation is with no loss in generality.
Recalling that $z(w)$, $y(w)$, $n(w)$ and $e(w)$, denote the optimal choices of the $w$-household; substitution for $z[y(w), n(w)]$ from condition (3) into (B9) and rearrangement yields the following expression, which is identical to equation (9):

(B10)

$$\int \left[ y(w) - U(w) - n(w) \cdot e[n(w)] + h[1 - y(w) \cdot w - \alpha \cdot n(w)] + v[n(w)] + u[e[n(w)]] \right] dF(w) = R.$$  

This completes the derivation.
Appendix C: A Means-Tested System Strictly Dominates the Universal one

In this appendix we demonstrate the dominance of a means-tested child benefit system over a universal one. We restrict attention to a setting with a discrete distribution of skill levels.

Consider an economy with a finite number of skill levels denoted by: \( w_i; i = 1, 2, \ldots, N \), where \( 0 < w_1 < w_2 < \ldots < w_{N-1} < w_N \). We normalize the number of individuals of each type to unity with no loss in generality. We let \( J[w_i, z(y, n), y, n] \) denote the maximal level of utility derived by a household with ability level \( w_i \), gross income level \( y \), net income level \( z(y, n) \) and number of children \( n \). Formally,

\[
(C1) \quad J[w_i, z(y, n), y, n] = \max \left[ z(y, n) - n \cdot e + v(n) + u(e) + h(1 - y/w_i - n \cdot \alpha) \right].
\]

We assume that the government is implementing a universal system; namely:

\[
(C2) \quad z(y, n) = a(y) + b(n).
\]

We turn to show that any universal system of the form given in condition (C2) can be replaced by a means-tested system that maintains the government revenue requirement and attains a higher level of welfare.

Denote by \( (y_i, n_i) \) the optimal choice of an individual with skill-level \( w_i \), faced with the universal system given in condition (C2). Consider the following (means-tested) alternative tax-transfer schedule:

\[
(C3) \quad \hat{z}(y, n) = \begin{cases} 0 & y = y_i, n \neq n_i, \\ z(y, n) & \text{otherwise} \end{cases}
\]

Notice, that indeed the schedule given in (C3) is means-tested. To see this consider some arbitrary number of children, \( n \neq n_i \). Then, the marginal child subsidy is given by \( b'(n) \) for levels of income \( y \neq y_i \) and zero when the income level is \( y_i \).
It is straightforward to verify that the optimal choices of the individuals under the modified system given by (C3) will coincide with choices taken under the system in (C2).

By revealed preference considerations, it follows that:

\[(C4) \quad J[w_i, z(y_i, n_i), y_i, n_i] \geq \max_n J[w_i, z(y_i, n), y_i, n_i] \geq J[w_i, z(y_i, n), y_i, n],\]

for all \(i > 1\).

In fact, the last inequality in condition (C4) is strict as long as \(\alpha > 0\). To see this, notice that by the individual first-order conditions:

\[(C5) \quad v'(n_i) + b'(n_i) - \alpha \cdot h'(1 - y_i / w_i - n_i \cdot \alpha) - e_i = 0,\]

\[(C6) \quad u'(e_i) - n_i = 0.\]

Hence by the concavity of \(h\),

\[(C7) \quad v'(n_i) + b'(n_i) - \alpha \cdot h'(1 - y_i / w_i - n_i \cdot \alpha) - e_i > 0,\]

for all \(i > 1\).

Thus, conditional on the level of income, \(y_i\), the optimal number of children chosen by an individual of type \(i > 1\) differs from \(n_i\). Obviously, when \(\alpha = 0\), all types will chose the same number of children (conditional on income). In this case, child benefits are redundant.

Let \(\delta_i = \max_n J[w_i, z(y_i, n), y_i, n] - J[w_i, z(y_i, n_i), y_i, n_i]\). As we have just shown, \(\delta_i > 0\) for all \(i > 1\). Further, let \(\delta = \min_{i=1}^{N} (\delta_i) > 0\).

Consider next the following tax-transfer schedule:

\[(C8) \quad z^*(y, n) = \begin{cases} 0 & \text{if } y = y_i, n \neq n_i \\ z(y, n) + (N - 1) \cdot \delta / N & \text{if } y = y_i, n = n_i \\ z(y, n) - \delta / N & \text{otherwise} \end{cases} \]
The tax-transfer schedule in (C8) is a modification of the schedule given in (C3). This schedule is obtained by levying a lump sum tax by the amount $\delta$ on all income levels different than $y_1$, and then distribute the entailed extra tax revenues in a lump sum fashion. Notice that by construction of $\delta$, the optimal choices of the individuals will still coincide with those obtained under the schedule given in (C3) and (C2). Moreover, the government revenue constraint will still be satisfied. However, as we redistribute resources from individuals of type $i>1$ towards the least well-off individual ($i=1$), we obtain a welfare gain, due the strictly concave welfare function [given in (10)].

The idea underlying the construction was that by forcing individuals of type $i>1$ who mimic the least well-off type ($i=1$), by choosing her level of income, also to choose the same number of children, we create a slack in the binding incentive constraints. We can then employ this slack to redistribute more towards the least well-off individual.
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: \( n'(w) \geq 0 \) if and only if \( a_{yy} y + a_y \leq 0 \)

Proof: We reproduce, for convenience, the \( w \)-household’s first-order conditions, given in equations (4)-(6), assuming a universal system, namely, \( z(y, n) = a(y) + b(n) \) and setting the marginal child subsidy to zero \( (b_n = 0) \):

\[
\begin{align*}
(D1) \quad & v'(n) - \alpha \cdot h'(1 - y/w - n \cdot \alpha) - e = 0, \\
(D2) \quad & a_y \cdot w - h'(1 - y/w - n \cdot \alpha) = 0, \\
(D3) \quad & u'(e) - n = 0.
\end{align*}
\]

We let \( g \) denote the inverse of \( u' \) (which is well-defined by the strict concavity of \( u \)). Hence, \( g(n) = e \). Substituting into (D1) and (D2) then yields:

\[
\begin{align*}
(D1') \quad & H(n, y, w) \equiv v'(n) - \alpha \cdot h'(1 - y/w - n \cdot \alpha) - g(n) = 0, \\
(D2') \quad & K(n, y, w) \equiv a_y \cdot w - h'(1 - y/w - n \cdot \alpha) = 0.
\end{align*}
\]

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:

\[
\begin{align*}
(D4) \quad & \partial H / \partial n \cdot n'(w) + \partial H / \partial y \cdot y'(w) + \partial H / \partial w = 0, \\
(D5) \quad & \partial K / \partial n \cdot n'(w) + \partial K / \partial y \cdot y'(w) + \partial K / \partial w = 0.
\end{align*}
\]

Employing Cramer’s Rule then yields:

\[
(D6) \quad n'(w) = \frac{-\partial H / \partial w \cdot \partial K / \partial y + \partial K / \partial w \cdot \partial H / \partial y}{\partial H / \partial n \cdot \partial K / \partial y - \partial K / \partial n \cdot \partial H / \partial y}
\]
By the second-order conditions of the household's optimization it follows that
\[ \frac{\partial H}{\partial n} \cdot \frac{\partial K}{\partial y} - \frac{\partial K}{\partial n} \cdot \frac{\partial H}{\partial y} > 0. \]
Thus, it follows:

(D7) \[ \text{Sign}[n'(w)] = \text{Sign}[-\frac{\partial H}{\partial w} \cdot \frac{\partial K}{\partial y} + \frac{\partial K}{\partial w} \cdot \frac{\partial H}{\partial y}]. \]

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

(D8) \[ \frac{\partial H}{\partial w} = -\alpha \cdot h''(1 - y/w - n \cdot \alpha) \cdot y/w^2, \]

(D9) \[ \frac{\partial H}{\partial y} = \alpha \cdot h''(1 - y/w - n \cdot \alpha)/w, \]

(D10) \[ \frac{\partial K}{\partial y} = a_{ys} \cdot w + h''(1 - y/w - n \cdot \alpha)/w, \]

(D11) \[ \frac{\partial K}{\partial w} = a_y - h''(1 - y/w - n \cdot \alpha) \cdot y/w^2. \]

Substituting from (D8)-(D11) into (D7) and re-arranging then yields:

(D12) \[ -\frac{\partial H}{\partial w} \cdot \frac{\partial K}{\partial y} + \frac{\partial K}{\partial w} \cdot \frac{\partial H}{\partial y} = \alpha \cdot h''(1 - y/w - n \cdot \alpha)/w \cdot [a_{ys} \cdot y + a_y]. \]

The result follows by virtue of the concavity of \( h \).
References


2906 Xavier Vives, Asset Auctions, Information, and Liquidity, January 2010

2907 Edwin van der Werf, Unilateral Climate Policy, Asymmetric Backstop Adoption, and Carbon Leakage in a Two-Region Hotelling Model, January 2010

2908 Margarita Katsimi and Vassilis Sarantides, Do Elections Affect the Composition of Fiscal Policy?, January 2010

2909 Rolf Golombek, Mads Greaker and Michael Hoel, Climate Policy without Commitment, January 2010

2910 Sascha O. Becker and Ludger Woessmann, The Effect of Protestantism on Education before the Industrialization: Evidence from 1816 Prussia, January 2010

2911 Michael Berlemann, Marco Oestmann and Marcel Thum, Demographic Change and Bank Profitability. Empirical Evidence from German Savings Banks, January 2010

2912 Øystein Foros, Hans Jarle Kind and Greg Shaffer, Mergers and Partial Ownership, January 2010

2913 Sean Holly, M. Hashem Pesaran and Takashi Yamagata, Spatial and Temporal Diffusion of House Prices in the UK, January 2010

2914 Christian Keuschnigg and Evelyn Ribi, Profit Taxation and Finance Constraints, January 2010

2915 Hendrik Vrijburg and Ruud A. de Mooij, Enhanced Cooperation in an Asymmetric Model of Tax Competition, January 2010

2916 Volker Meier and Martin Werding, Ageing and the Welfare State: Securing Sustainability, January 2010

2917 Thushyanthan Baskaran and Zohal Hessami, Globalization, Redistribution, and the Composition of Public Education Expenditures, January 2010

2918 Angel de la Fuente, Testing, not Modelling, the Impact of Cohesion Support: A Theoretical Framework and some Preliminary Results for the Spanish Regions, January 2010

2919 Bruno S. Frey and Paolo Pamini, World Heritage: Where Are We? An Empirical Analysis, January 2010
2920 Susanne Ek and Bertil Holmlund, Family Job Search, Wage Bargaining, and Optimal Unemployment Insurance, January 2010

2921 Mariagiovanna Baccara, Allan Collard-Wexler, Leonardo Felli and Leeat Yariv, Gender and Racial Biases: Evidence from Child Adoption, January 2010

2922 Kurt R. Brekke, Roberto Cellini, Luigi Siciliani and Odd Rune Straume, Competition and Quality in Regulated Markets with Sluggish Demand, January 2010

2923 Stefan Bauernschchuster, Oliver Falck and Niels Große, Can Competition Spoil Reciprocity? – A Laboratory Experiment, January 2010

2924 Jerome L. Stein, A Critique of the Literature on the US Financial Debt Crisis, January 2010

2925 Erkki Koskela and Jan König, Profit Sharing, Wage Formation and Flexible Outsourcing under Labor Market Imperfection, January 2010

2926 Gabriella Legrenzi and Costas Milas, Spend-and-Tax Adjustments and the Sustainability of the Government’s Intertemporal Budget Constraint, January 2010

2927 Piero Gottardi, Jean Marc Tallon and Paolo Ghirardato, Flexible Contracts, January 2010

2928 Gebhard Kirchgässner and Jürgen Wolters, The Role of Monetary Aggregates in the Policy Analysis of the Swiss National Bank, January 2010

2929 J. Trent Alexander, Michael Davern and Betsey Stevenson, Inaccurate Age and Sex Data in the Census PUMS Files: Evidence and Implications, January 2010

2930 Stefan Krasa and Mattias K. Polborn, Competition between Specialized Candidates, January 2010

2931 Yin-Wong Cheung and Xingwang Qian, Capital Flight: China’s Experience, January 2010

2932 Thomas Hemmelgarn and Gaetan Nicodeme, The 2008 Financial Crisis and Taxation Policy, January 2010

2933 Marco Faravelli, Oliver Kirchkamp and Helmut Rainer, Social Welfare versus Inequality Concerns in an Incomplete Contract Experiment, January 2010

2934 Mohamed El Hedi Aoruri and Christophe Rault, Oil Prices and Stock Markets: What Drives what in the Gulf Corporation Council Countries?, January 2010

2935 Wolfgang Lechthaler, Christian Merkl and Dennis J. Snower, Monetary Persistence and the Labor Market: A New Perspective, January 2010

2936 Klaus Abberger and Wolfgang Nierhaus, Markov-Switching and the Ifo Business Climate: The Ifo Business Cycle Traffic Lights, January 2010
2937 Mark Armstrong and Steffen Huck, Behavioral Economics as Applied to Firms: A Primer, February 2010

2938 Guglielmo Maria Caporale and Alessandro Girardi, Price Formation on the EuroMTS Platform, February 2010

2939 Hans Gersbach, Democratic Provision of Divisible Public Goods, February 2010

2940 Adam Isen and Betsey Stevenson, Women’s Education and Family Behavior: Trends in Marriage, Divorce and Fertility, February 2010

2941 Peter Debaere, Holger Görg and Horst Raff, Greasing the Wheels of International Commerce: How Services Facilitate Firms’ International Sourcing, February 2010

2942 Emanuele Forlani, Competition in the Service Sector and the Performances of Manufacturing Firms: Does Liberalization Matter?, February 2010

2943 James M. Malcomson, Do Managers with Limited Liability Take More Risky Decisions? An Information Acquisition Model, February 2010

2944 Florian Englmaier and Steve Leider, Gift Exchange in the Lab – It is not (only) how much you give …, February 2010

2945 Andrea Bassanini and Giorgio Brunello, Barriers to Entry, Deregulation and Workplace Training: A Theoretical Model with Evidence from Europe, February 2010

2946 Jan-Emmanuel De Neve, James H. Fowler and Bruno S. Frey, Genes, Economics, and Happiness, February 2010

2947 Camille Cornand and Frank Heinemann, Measuring Agents’ Reaction to Private and Public Information in Games with Strategic Complementarities, February 2010

2948 Roel Beetsma and Massimo Giuliodori, Discretionary Fiscal Policy: Review and Estimates for the EU, February 2010

2949 Agnieszka Markiewicz, Monetary Policy, Model Uncertainty and Exchange Rate Volatility, February 2010

2950 Hans Dewachter and Leonardo Iania, An Extended Macro-Finance Model with Financial Factors, February 2010

2951 Helmuth Cremer, Philippe De Donder and Pierre Pestieau, Education and Social Mobility, February 2010

2952 Zuzana Brixiová and Balázs Égert, Modeling Institutions, Start-Ups and Productivity during Transition, February 2010

2953 Roland Strausz, The Political Economy of Regulatory Risk, February 2010
Sanjay Jain, Sumon Majumdar and Sharun W. Mukand, Workers without Borders? Culture, Migration and the Political Limits to Globalization, February 2010

Andreas Irmen, Steady-State Growth and the Elasticity of Substitution, February 2010

Bengt-Arne Wickström, The Optimal Babel – An Economic Framework for the Analysis of Dynamic Language Rights, February 2010

Stefan Bauernschuster and Helmut Rainer, From Politics to the Family: How Sex-Role Attitudes Keep on Diverging in Reunified Germany, February 2010


Betsey Stevenson, Beyond the Classroom: Using Title IX to Measure the Return to High School Sports, February 2010

R. Quentin Grafton, Tom Kompas and Ngo Van Long, Biofuels Subsidies and the Green Paradox, February 2010

Oliver Falck, Stephan Heblich, Alfred Lameli and Jens Suedekum, Dialects, Cultural Identity, and Economic Exchange, February 2010


Frederick van der Ploeg and Cees Withagen, Is There Really a Green Paradox?, February 2010

Ingo Vogelsang, Incentive Regulation, Investments and Technological Change, February 2010

Jan C. van Ours and Lenny Stoeldraijer, Age, Wage and Productivity, February 2010

Michael Hoel, Climate Change and Carbon Tax Expectations, February 2010

Tommaso Nannicini and Roberto Ricciuti, Autocratic Transitions and Growth, February 2010

Sebastian Brauer and Frank Westermann, A Note on the Time Series Measure of Conservatism, February 2010

Wolfram F. Richter, Efficient Education Policy – A Second-Order Elasticity Rule, February 2010

Tomer Blumkin, Yoram Margalioth and Efraim Sadka, Taxing Children: The Redistributive Role of Child Benefits – Revisited, February 2010