How to Deal with Covert Child Labor and Give Children an Effective Education, in a Poor Developing Country*

Alessandro Cigno†

Abstract

Because credit and insurance markets are imperfect, and given that intrafamily transfers and the way children use their time outside school hours are private information, the second-best policy makes school enrollment compulsory, forces overt child labor below its efficient level (if positive), and uses a combination of need- and merit-based grants, financed by earmarked taxes, to relax credit constraints, redistribute, and insure. Existing conditional cash transfer schemes can be made to approximate the second-best policy by incorporating these principles in some measure.

JEL Codes: D82, H21, H31, I28, J24

Keywords: child labor, education, uncertainty, moral hazard, optimal taxation

Developing country governments and international development agencies have long been aware that human capital accumulation, more than physical capital accumulation, is the mainspring of economic and civil progress. But many children in poor developing countries fail to complete even primary education, and some do not go to school at all. The reasons are well known.[1] Baland and Robinson (2000) demonstrate that child labor will be inefficiently high if parents are either credit or bequest constrained.[2] Evidence that parent inability to borrow discourages education and encourages child labor is reported by a host of researchers, including Jacoby (1994) and Fuwa and others (2009). Loury (1981) and Pouliot (2006) demonstrate that parent inability to insure against the risk of a low return causes education investment to be inefficiently low and child labor to be inefficiently high, even when credit is not rationed and bequests are interior.[3] Ram and Schultz (1979) and Jacoby and Skoufias (1997) find evidence that parent inability to insure against the risk of a low return

*Paper to the 2012 SIE meeting, Matera.

†Alessandro Cigno (cigno@unifi.it) is a professor of economics at the University of Florence.

[1] For a systematic exposition, see Cigno and Rosati (2005).

[2] Cigno (1993, 2006) shows, however, that the problem is alleviated when a set of self-enforcing, renegotiation-proof family rules oblige working-age family members to support their young children and elderly parents.

[3] According to Levhari and Weiss (1974), the return to education is uncertain because a child's learning ability is fully revealed only after the education investment is carried out. For

discourages education investment in developing countries. Parent incomes may also be uncertain. Evidence in Beegle, Dehejia, and Gatti (2006), that parents respond to a negative income shock by making their children work more, suggests that households cannot insure against that kind of risk either. Fitzsimons (2007) reports, however, that parents respond in this way to a downturn not in their own income, but in village aggregate income, suggesting that idiosyncratic income shocks are neutralized by informal insurance arrangements at the local level.[4]

Because idiosyncratic shocks even out on average, governments face less risk than do individual households. Partly because of this lower risk, governments also have easier access to international money markets than do most citizens. Thus in imperfect domestic credit and insurance markets there is an efficiency argument for governments to lend to and insure parents of school-age children. Given evidence of diminishing absolute risk aversion in an education context (Kodde 1986, for example) and given an unequal distribution of parent wealth, there is also an equity argument. Efficiency-enhancing policies are politically easier to implement when they do not involve redistribution, but it is difficult to see how redistribution could be avoided. Even if the government could finance the education of poor children entirely by borrowing against their future tax payments, insuring families against the risk of a low return to education would still imply redistribution from rich to poor school leavers. Similarly, insuring families against the risk of a downturn in parent income would involve redistribution from rich parents to poor parents.[5] The tension between equity and efficiency will be minimal for a small-scale project, especially if financed largely by international aid,[6] but not for a large-scale one. It is also difficult to imagine that any project, large or small, could be supported by the international community forever.

Information asymmetries give rise to another set of problems. In developing countries many children work, but much of what they do is invisible to the government. A small fraction of this covert child labor involves physically damaging or morally degrading activities—the "worst forms" of child labor—which national governments are committed by international treaty to eradicate. But most covert child labor consists of activities conducted for and under the direct supervision of the child's parents (such as helping in the home, working on the family farm, and contributing to the family business).[7] While comparatively harmless in themselves, these activities conflict with education and thus have

---

evidence of that, see Belzil and Hansen (2002). According to Razin (1976), the uncertainty concerns the rental price of the human capital accumulated through education. In developing countries the uncertainty also concerns the length of time for which the future adult will be able to enjoy the benefit.

[4] Evidence of such arrangements in a developing country is reported by Besley (1995) and Townsend (1994), among others.

[5] See Johnson (1987).

[6] It will not even arise when education is privately financed by migrant remittances. Dessy and Rambeloma (2010), Epstein and Kahana (2008), and Hanson and Woodruffs (2003), report evidence that such remittances reduce child labor in the families left behind.

[7] See Cigno and Rosati (2005).

an opportunity cost in terms of forgone future earnings. The government might want to regulate them, but it cannot because they are private information; this creates a moral hazard problem.

Similar considerations apply from an education standpoint. Scholastic performance depends not only on how much time children spend attending school, but also on how much time they spend doing homework and on how alert and well rested they are when doing both. A child who falls asleep during lessons and does not find the time or is too tired to give homework the necessary attention will have poorer results in school than will a child of the same learning ability who comes to school well rested and with homework conscientiously done. Because school enrollment and school attendance are common knowledge but much of what children do outside school is private information, another moral hazard problem exists. And a similar problem arises from the fact that intrafamily transfers are private information, because the government cannot be sure that a public subsidy intended for children does not end up as extra consumption for the parents.

This article describes the second-best policy and compares it with two benchmarks—a low one represented by a laissez-faire policy and a high one represented by the first-best policy—in a situation where parents can neither borrow nor insure, where parents are better informed than the government about their children's time allocation, and where the government does not observe intrafamily transfers. The analysis assumes that the expected return to education is positive.[8] Fasih (2008) reports evidence of high returns to education, especially in low- and middle-income countries. The same study reports that these returns are lower for poor children than for rich children. That may be a sign that the poor can afford only poor-quality schools,[9] but the line of reasoning in this article suggests another (not necessarily rival) explanation: that children in poor families study less, or less effectively, per year of school enrollment or day of school attendance than do children in rich families. The worst forms of child labor raise moral issues that transcend the materialistic calculations underlying this article[10] and are thus omitted from the formal analysis, but the analysis argues that the proposed policy reduces child labor in all its forms.

The policy optimization has an optimal taxation, or principal-agent, format.[11] This type of analysis does not appear to have been attempted before in the context of a poor developing country. Given the context, "school age" is taken to mean primary school age. Assuming that children in that age range are under parent control, parents, rather than the children themselves, are taken to be the agents and are modeled as risk-averse, expected-utility maximizers.[12]

[8] For evidence of a causal effect of education on future earnings, see Card (1999) and Oreopoulos (2006).

[9] On the subject, see Alderman, Orazem, and Paterno (2001).

[10] See Dessy and Pallage (2005) for a strictly economic analysis.

[11] For a survey of the ways in which the approach can be used in a family policy context, see Cigno (2011). For an application in higher education, see Cigno and Luporini (2009).

[12] That is not the only possible representation of individual behavior in the face of uncertainty. In prospect theory (Kahneman 2003) individuals are assumed to be risk averse in the domain of gains and risk lovers in that of losses. Although this alternative approach has

Because the implications of an education externality are well understood and allowing for it would merely reinforce the argument for public intervention, the analysis excludes it (but finds that the policy itself gives rise to a fiscal externality). Because the argument for having the policy financed out of general tax revenue is weak in the absence of an education externality, and assuming that international aid cannot go on forever, the constraint that the policy must be self-financing is imposed. Section I lays down the technical assumptions and characterizes parent decisions. Section II examines the laissez-faire equilibrium. Section III derives the first- and second-best policies. And section IV discusses actual policy practice (including conditional cash transfer schemes) in the light of the theoretical findings.

I. Technical Assumptions and Parent Decisions

There are a large number of families, $i = 1, 2...n$, each consisting of a couple with a given number of children, the same for every family and normalized to unity. The assumption that all parents have the same number of children is less than realistic, but the normative implications of departing from it have been examined in depth elsewhere[13] and do not impinge on the points at issue here. Learning ability is randomly distributed across children and imperfectly observable until the education investment is carried out. Parent income varies exogenously across families and is observable by the government. Later in the analysis income will be allowed to be either uncertain or private information.

There are two periods, $t = 1, 2$. Children are alive in both, parents only in the first. For brevity, the child in the ith family is referred to as i. Ex post, i's utility will be $U_i = u(c_{i1}) + u(c_{i2})$, where c_t^i denotes i's consumption in period t. Assuming descending altruism, the ex post utility of i's parents may be written as $V_i = v(a_i) + \beta U_i$, $0 < \beta < 1$, where a_i denotes parent consumption and β is a measure of altruism. The functions $u(.)$ and $v(.)$ are assumed to be increasing and concave, with $u'(0) = v'(0) = \infty$. In an uncertain environment concavity implies risk aversion. Assuming that marginal utility becomes very large as consumption approaches zero implies that subsistence consumption is

some empirical justification, it is not followed here for two reasons. First, in a situation where most people live slightly above the subsistence level, little risk-loving behavior is likely to be observed. Second, the policymaker may not approve of such behavior and may consequently maximize an objective function that is not a mere aggregation of the individual ones. Kanbur, Pirttilä, and Tuomala (2008) show that if the government corrects for what it considers an aberrant behavior, the solution to an optimal taxation problem with moral hazard may have the same properties as if the agents were risk-averse, expected-utility maximizers.

[13] If the number of children is exogenous and parent income or work effort are private information, the optimal income tax rate is zero for the family with the highest income. If the number of children is exogenous but varies across families as in Cremer, Dellis, and Pestieau (2003), the optimal policy will redistribute in favor of families with more children. Neither of these properties necessarily applies, however, if the number of children is endogenous, as in Cigno (2001) and Balestrino, Cigno, and Pettini (2002): the first one (no distortion at the top) because children's visibility makes mimicking much harder, the second one (children reduce tax liability) because children yield utility.

normalized to zero. Because utility does not depend on time allocation, this implies that leisure is not a good and that work does not yield direct disutility. This may be justified by saying that, with the worst forms of child labor out of the picture and consumption likely to be low, the marginal utility of income is likely to be higher than that of leisure.

In period 1 a child may or may not be enrolled at school. Enrollment has a fixed cost p, equal to the average cost of tuition.[14] The previous section argued that effective education time is increasing in school attendance, homework, and rest time and decreasing in work time. If i is enrolled, effective education time will be positive (or there would be no point in paying p). To simplify, effective education time is measured as the amount of time that the child does not work. If i is not enrolled, effective education time will be zero.[15]

Child labor may be overt or covert. Overt child labor consists of work done for an employer other than the child's parents and carries a wage. Covert child labor involves either participating in a family-run, income-generating activity, such as farming or retailing, or replacing the child's parents in performing household chores such as cooking, cleaning, fetching water, and gathering fuel.[16] Although neither form of covert child labor carries a wage, the former produces income directly, and the latter indirectly by allowing the child's parents to spend more time raising income. Let e_i denote i's effective education, and let L_i denote i's overt labor. Normalizing a child's time endowment to unity, i's covert child labor is then $(1 - e_i - L_i)$.

Because children can contribute to the production of family income, parent income in period 1 is defined as family income if child labor in both its forms are zero. Let y_i denote parent income in family i. The income generated by overt child labor is $L_i w_1$, where w_1 is the child wage rate, and income generated by covert child labor is $z(1 - e_i - L_i)$, where $z(.)$ is a revenue function, increasing and concave, with $z(0) = 0$ and $z'(0) = \infty$. By definition of revenue function, $z(1 - e_i - L_i)$ is the maximum amount of income that the family can produce with $(1 - e_i - L_i)$ units of i's time by optimally allocating this time between direct participation in income-raising activities conducted by i's parents and replacement of i's parents in performing household chores. Concavity reflects diminishing marginal rates of technical substitution between adult and child work. Assuming that marginal revenue gets very large as covert child labor gets very small is realistic in a poverty context like the present one and ensures that such labor will never be zero. In period 2 i will earn $w_2 + x_i$, where w_2 denotes

[14] Tuition fees are usually per year (or shorter period such as the semester) of school enrollment, so the total a family spends for a child's tuition reflects the number of years for which the child is enrolled at school but not the number of days for which the child actually attends school or the number of hours during which the child studies at home, in each of those years. This lumpiness of tuition fees is accounted for by treating p as a constant.

[15] In many developing countries a substantial minority of school-age children is reported as neither working nor studying. This can be explained without introducing leisure by allowing for the existence of fixed costs of access to school and work; see Cigno and Rosati (2005).

[16] See Cigno and Rosati (2005) for an analysis of the incidence of these activities and for the effects of making fetching water and gathering fuel unnecessary by providing homes with electricity and running water.

the income of an unskilled adult and x_i denotes the individual skill premium. If i does not enroll at school, x_i will be zero. In period 1, i receives a transfer, m_i, from i's parents,[17] and another, γ_i, from the government. In period 2 i will make a transfer to the government, θ_i. All these transfers can be positive, negative, or zero.

Parents make their decisions in period 1, after the government has announced its policy. Anticipating a result to be obtained in section III, γ_i is taken to be a function of y_i, and θ_i to be a function of x_i. While w_1, w_2, and γ_i are certain, x_i and consequently θ_i are uncertain. Because e_i must be chosen in period 1, education is a risky investment. The supplementary assumption (to be relaxed later) is made that x_i is independent and identically distributed over the closed interval $[0, \overline{x}] \in R^+$ with density $f(.|e_i)$ conditional on e_i and $f(.|0) = 0$. To simplify the notation, x_i is used to measure both the final school result and the skill premium.[18] The cumulative distribution of x_i, $F(x_i|e_i)$, associated with a higher e_i, first-order stochastically dominates the one associated with a lower e_i,

$$F_{e_i}(x_i|e_i) \leq 0. \tag{1}$$

In other words, the more i studies and the less i works, the more of a chance i has of getting good marks and thus of attracting a high skill premium. For each e_i there will be values of x_i such that equation (1) holds as an inequality. The standard convexity of distribution function assumption, that $F\left(x^i|e_i\right)$ is convex in e_i, and monotone likelihood ratio assumption, that $\frac{f_{e_i}(.|e_i)}{f(.|e_i)}$ is increasing in x_i,[19] are made, which allow the first-order approach to be adopted.

If i enrolls for school and overt child labor is not regulated by the government, i's parents will choose the (e_i, L_i, m_i) that maximizes $E(V_i) \equiv v_i + \beta\left(u_{i1} + \int_{x_i} u_{i2} f^i dx_i\right)$, where $v_i \equiv v(y_i + z_i - m_i)$, $z_i \equiv z(1 - L_i - e_i)$, $u_{i1} \equiv u(m_i + w_1 L_i + \gamma_i - p)$, $u_{i2} \equiv u(w_2 + x_i - \theta_i)$, and $f^i \equiv f(x_i|e_i)$, subject to

$$e_i \geq 0, \tag{2}$$

$$L_i \geq 0 \tag{3}$$

and

$$1 - e_i - L_i \geq 0. \tag{4}$$

[17] One might be tempted to simplify the analysis by taking the utility aggregation problem as solved and viewing V_i as a family welfare function. This would allow intrafamily transfers to be left out and all costs and revenues to be treated as pertaining to the family as a whole, but doing so would be misleading, because, as Baland and Robinson (2000) show, transfers from parents to children may be inefficiently low.

[18] Using one random variable with density conditional on study time to represent the school result and another with density conditional on the school result to represent the skill premium would make no substantive difference to the results so long as both are independent and identically distributed and the skill premium is not conditional on some decision variable.

[19] This property might not hold if x_i depended on systemic factors and $(1 - e_i)$ depended on employment opportunities. In the present context, however, it seems reasonable to assume that there is nothing to stop w_i from falling low enough to clear the (overt) child labor market and that there will also be plenty of opportunities for covert child labor.

Because equation (4) will never be binding for the restrictions imposed on the revenue function, the first-order conditions are

$$-v_i' z_i' + \beta \int_{x_i} u_{i2} f_{e_i}^i dx_i + \xi_i = 0 \quad \text{for } e^i \tag{5}$$

$$-v_i' z_i' + \psi_i + \beta u_{i1}' w_1 = 0 \text{ for } L^i \tag{6}$$

and

$$-v_i' + \beta u_{i1}' = 0 \text{ for } m^i \tag{7}$$

where ξ_i is the Lagrange multiplier of equation (2) and ψ_i that of equation (3). If i does not enroll, e_i cannot be positive. Again assuming that overt child labor is free to vary, i's parents will then choose (L_i, m_i) to maximize $V(L_i, m_i) \equiv v(y_i + z(1 - L_i) - m_i) + \beta [u(m_i + w_1 L_i) + u(w_2)]$, subject to equation (3) − equation (4). The solution will satisfy equation (6) − equation (7) for $p \equiv e_i \equiv 0$. If L_i is regulated by the government, equation (6) need not hold. Irrespective of whether i is enrolled and L_i is regulated, it is clear from equation (7) that m_i is decreasing in γ_i. In other words, public transfers crowd out private transfers.

II. Laissez-Faire Equilibrium

Under laissez-faire school enrollment is not compulsory, overt child labor is free to vary, and $\gamma^i \equiv \theta^i \equiv 0$. The payoff of enrolling i at school is

$$\pi^S(y_i, p) \equiv \max_{(L_i, e_i, m_i)} E(V_i), \text{ subject to equation (2)} - \text{equation (4)} \tag{8}$$

and the payoff of not enrolling is

$$\pi^W(y_i) \equiv \max_{(L_i, m_i)} V(L_i, m_i), \text{ subject to equation (2)} - \text{equation (4).} \tag{9}$$

The child will enroll if $\pi^S(y_i, p)$ is at least as large as $\pi^W(y_i)$. There is then a threshold value of y_i, $\widetilde{y}$, defined by $\pi^S(\widetilde{y}, p) = \pi^W(\widetilde{y})$, below which i will not be enrolled. $\widetilde{y}$ is the same for every i, because the expected return to education is the same for all of them, so if any children are not enrolled, it will be those whose parents have a low income. This result differs from the one in Ranjan (2001), where children's learning ability is assumed to be directly observable ex ante and the threshold is consequently lower for parents of high-ability children than for parents of low-ability children.

Given that ξ_i will be zero if e_i is positive, equation (5) implies

$$\text{either } e_i = 0 \text{ or } v_i' z_i' = \beta \int_{x_i} u_{i2} f_{e_i}^i dx_i. \tag{10}$$

Therefore, either e_i is zero and i is not enrolled or e_i is positive and increasing in y_i. Taken together with equation (7) and given that ψ_i will be zero if L_i is positive, equation (6) similarly implies

$$\text{either } L_i = 0 \text{ or } z'(1 - L_i - e_i) = w_1. \tag{11}$$

Therefore, L_i is either zero, or positive and increasing in w_1. It is then clear that overt child labor is the same in all families.[20] What differs is effective education and total (overt plus covert) child labor.

> *Proposition 1. Under laissez-faire children from very poor families are not enrolled at school; children from less poor families are enrolled, but their effective education increases with parent income; and overt child labor is either zero or increasing in the child wage rate.*

The second part of this proposition provides a possible explanation for the empirical finding that poor children get a smaller increase in their future income in return for an extra year of school enrollment or an extra day of school attendance than do rich children, because it says that poor children receive less effective education during that extra year or day than do rich children.

III. First- and Second-Best Policies

The government's preferences are represented by the Benthamite social welfare function,

$$SW = \sum_{i=1}^{n} E(V_i). \tag{12}$$

Because there are many parents and children and because risks are assumed to be uncorrelated, the government does not face any uncertainty about its tax revenue and thus has easier access to international credit than individual citizens. The usual "small country" assumption—that the real interest rate is constant and normalized to zero—is made. Because the expected return to education is the same for every i, and assuming it is positive, the government will then make school enrollment compulsory. Because the optimization can determine only relative tax rates, the tax on w_2 is normalized to zero, and the socially optimal values of γ_i and θ_i are investigated.

Because the government does not face budget uncertainty, it will choose $(e_i, L_i, m_i, \gamma_i, \theta_i)$, for $i = 1, 2, ...n$, to maximize equation (12), subject to the budget constraint,

$$\sum_{i=1}^{n} \left(\gamma_i - \int_{x_i} \theta_i f^i dx_i \right) = 0, \tag{13}$$

and equation (2) − equation (4). If (e_i, m_i) is private information, the maximization will be subject also to incentive-compatibility constraints. Because

[20] It would vary across families if the $z()$ function did (for example, if the return to covert child labor were higher in a farming family that owns land than in one that does not).

$E(V_i)$ is concave in (e_1, L_i, m_i), SW will be concave in it too. For the independent and identically distributed assumption, the optimal (γ_i, θ_i) can depend only on (e_i, m_i, x_i, y_i) and not on any (e_j, m_j, x_j, y_j) for $j \neq i$.

First-Best Policy

Under the first-best policy the government prescribes (e_i, L_i, m_i) and designs personalized lump-sum transfers, (γ_i, θ_i), for each i. Because there are no incentive-compatibility constraints, and denoting the Lagrange multiplier of equation (13) as λ, the first-order conditions are, respectively, equation (6) for L_i, equation (7) for m_i,

$$-v_i' z_i' + \int_{x_i} (\beta u_{i2} + \lambda \theta_i) f_{e_i}^i dx_i + \xi_i = 0 \text{ for } e_i, \tag{14}$$

$$\beta u_{i1}' - \lambda = 0 \text{ for } \gamma_i \tag{15}$$

and, at each possible realization of x_i,

$$-(\beta u_{i2}' - \lambda) f^i = 0 \text{ for } \theta_i. \tag{16}$$

Because equation (11) must still hold, it is clear that the first-best L_i is the same for every i, $L_i = L^{FB}$, and not necessarily zero. The first-order condition on e_i is not the same as under laissez-faire because it takes account of the expected marginal benefit of tax revenue for society as a whole, $\lambda \int_{x_i} \theta_i f_{e_i}^i dx_i$. Given this fiscal externality, e_i will be larger than under laissez-faire for every i.

In view of equations (7), (15), and (16), it is also clear that $a_i = a^{FB}$, $c_{i1} = c_{i2} = c^{FB}$, and $m_i = m^{FB}$. Because this implies that parent income is equalized across families and children are ex ante identical, the first-best level of e_i is the same for every i, $e_i = e^{FB}$.

Proposition 2. Under the first-best policy the government uses lump-sum taxes and subsidies to achieve perfect equity, perfect consumption smoothing, and full insurance; all school-age children allocate their time in the same way; overt child labor is either zero or increasing in the child wage rate; each school-age child receives more effective education than under laissez-faire.

The last part of this proposition implies that the laissez-faire level of effective education is inefficiently low.

Second-Best Policy

Under the second-best policy, (e_i, m_i) is private information. According to the logic of optimal taxation, the government will then make school enrollment compulsory, fix L_i, and influence parent decisions by announcing how γ_i and θ_i will be related to the information available in the relevant period. Because γ_i is payable in period 1, it can depend only on y_i. Because θ_i is payable in period 2, it can also depend on x_i. If it seems odd that a benevolent government might actually oblige children to do a certain amount of paid work, think of the second-best value of L_i as a legal maximum. Because of the potential moral hazard problem, the maximization of equation (12) is subject not only to equation (2) − equation (4) and equation (13), but also to the incentive-compatibility constraints represented by equations (5) and (7). Let ϕ_i denote the Lagrange multiplier of equation (5) and μ_i that of equation (7). The first-order conditions are

$$-v_i' z_i' + \int_{x_i} (\beta u_{i2} + \lambda\theta_i) f_{e_i}^i dx_i + \xi_i + \phi_i \left[v_i' z_i'' + v_i'' (z_i')^2 + \beta \int_{x_i} u_{i2} f_{e_i e_i}^i dx_i \right] + \mu_i v_i'' z_i' = 0 \tag{17}$$

for e_i,

$$-v_i' z_i' + \beta u_{i1}' w_1 + \psi_i + \phi_i \left[v_i' z_i'' + v_i'' (z_i')^2 \right] + \mu_i \left[v_i'' z_i' + \beta u_{i1}'' w_1 \right] = 0 \tag{18}$$

for L_i,

$$-v_i' + \beta u_{i1}' + \phi_i v_i'' z_i' + \mu_i \left[v_i'' + \beta u_{i1}'' \right] = 0 \tag{19}$$

for m_i,

$$\beta u_{i1}' - \lambda + \mu_i \beta u_{i1}'' = 0 \tag{20}$$

for γ_i and, at each possible realization of x_i,

$$-\left(\beta u_{i2}' - \lambda\right) f^i - \phi_i \beta u_{i2}' f_{e_i}^i = 0 \tag{21}$$

for θ_i.

Using equation (7), equation (20) can be rewritten as $1 + \mu_i r_i = \frac{\lambda}{v_i'}$, where $r_i \equiv -\frac{u_{i1}''}{u_{i1}'}$ is the Arrow-Pratt measure of absolute risk aversion. As long as r_i is nonincreasing in i's income,[21] and given that v_i' is decreasing in y_i,

$$\gamma_i = \gamma(y_i), \; \gamma' < 0. \tag{22}$$

Condition (21) may be similarly rewritten as $1 + \phi_i \frac{f_{e_i}}{f^i} = \frac{\lambda}{\beta u_{i2}'}$. Because $\frac{f_{e_i}}{f^i}$ is increasing in x_i, and u_{i2}' in θ_i,

$$\theta_i = \theta(x_i), \; \theta' < 0. \tag{23}$$

Because there is nothing to prevent γ_i from falling below zero for some y_i, γ_i can be interpreted as the difference between an education grant equal to p[22] and an earmarked tax increasing in parent income. Similarly, because there is nothing to stop θ_i from being negative for some x_i, θ_i can be interpreted as the

[21] For evidence, see Johnson (1987).

[22] Recall that subsistence consumption is normalized to zero.

difference between another earmarked tax, equal to p, and another education grant, this time increasing in the school result.

Having established that $\gamma(.)$ and $\theta(.)$ are decreasing functions, it is clear that the policy redistributes from the rich to the poor and insures parents and children against the risk of a low return to effective education. Comparing equation (20) with equation (15), and equation (21) with equation (16), however, it is also clear that the policy does not go as far as the first-best policy does. The reason is that redistribution has an efficiency cost, because the government cannot use personalized lump-sum transfers as under the first-best policy. What happens to (e_i, L_i, m_i)? Compare equation (17) with equations (4) and (5) to see that e_i is lower than under the first-best policy. In particular, because $(y_i + \gamma_i)$ is not the same for all i, e_i increases with y_i, as (albeit more slowly than) under laissez-faire. Because γ_i can be negative, it cannot be ruled out that e_i will be lower than under laissez-faire for some i. Because the government can borrow against $\int_{x_i} \theta_i f^i dx_i$, however, γ_i will be negative only if y_i is very high. Because the government is also insuring parents against the risk of a low return to effective education, it is thus unlikely that e_i will be lower than under laissez-faire for any i. Comparing equation (18) with equation (6) also shows that L_i will be no higher than under either the first-best policy or laissez-faire. The intuition is that if w_1 is high enough for the efficient L_i to be positive, imposing a ceiling on overt child labor will distort the allocation of i's total working time between overt and covert labor and thus make work as a whole less attractive than education. Comparing equation (19) with equation (7) shows that m_i will be lower than under either the first-best policy or laissez-faire.

> *Proposition 3. Under the second-best policy school enrollment is compulsory, and all school-age children, with the possible but unlikely exception of those from very rich families, receive more effective education than under laissez-fare; the government uses a net subsidy decreasing in parent income, and a net tax decreasing in the individual skill premium, to redistribute and insure, but stops short of perfect equity, full insurance, and perfect consumption smoothing; if the efficient level of overt child labor is positive, the government sets a limit, lower than the efficient level, on the amount of paid work a child can legally do.*

The implications of relaxing some of the assumptions made so far can be intuited without formal analysis. Suppose that the returns to education investment have an aggregate as well as an idiosyncratic component. Because aggregate risks cannot be insured against by redistributing within cohorts, the government must use its ability to borrow and lend on the international credit market to redistribute not only within, but also between, cohorts. A similar argument applies to parent incomes. If the shocks to parent income are purely idiosyncratic, the policy prescription remains qualitatively the same, because redistributing from rich parents to poor parents will insure families against the risk of a downturn in that income. The prescription also remains the same when the shocks have an area component, because the policy redistributes not only within, but also between, areas. If the shocks have a countrywide component, however, the government must use its ability to borrow and lend on the international credit market to redistribute not only within, but also between, cohorts (as in the case where the aggregate shocks concern the return to education investment).

IV. Policy Practice in the Light of Theoretical Findings

Under laissez-faire, if credit and insurance markets are imperfect or contracts between parents and young children are unenforceable, effective education is inefficiently low. If parent income is below a certain threshold, the child will not enroll at school. Above that threshold, the child will enroll, but children from poor families will receive less effective education than children from rich families. This prediction is consistent with evidence in Ram and Schultz (1979), Jacoby (1994), Jacoby and Skoufias (1997), Belzil and Hansen (2002), and Fuwa and others (2009) that inability to borrow and insure reduces education investment. It is consistent also with evidence, surveyed in Fasih (2008), that the return to measurable education inputs such as school enrollment or attendance is positive and particularly large in low- to middle-income countries but lower for poor children than for rich children. Because the amount of effective education that children receive in a year of school enrollment or day of school attendance is lower if they come from a family with low parent income than if they come from one with high parent income, the return to enrollment or attendance will in fact understate the return to effective education of poor children relative to that of rich children. This explanation does not conflict with other possible explanations, such as that poor children have access to poor quality schools only.

The optimal (first- or second-best) policy relaxes the credit constraint on education investment by giving parents an advance on the expected return and provides insurance against the risk of a low return by redistributing from lucky to unlucky school leavers. Because it redistributes from rich parents to poor parents, it will also reduce inequality and, if parent income is uncertain, provide insurance against the risk of a downturn in that income. The first-best policy uses personalized lump-sum transfers to achieve perfect equity, full insurance, and perfect consumption smoothing. Because children are ex ante identical, all parents enroll their children at school and give each child the same efficient amount of effective education. The second-best policy also redistributes and insures. Because it cannot use personalized lump-sum transfers, however, it stops short of perfect equity, full insurance, and perfect consumption-smoothing. It also raises effective education above the laissez-faire level for most children, but not to the efficient level. Although the worst forms of child labor are outside the scope of this analysis, a policy that encourages effective education will discourage all forms of covert child labor, including the worst ones.

Under the second-best policy school enrollment is compulsory (whereas under the first-best policy, it does not need to be compulsory because it is in the interest of all parents to send their children to school). If the child wage rate is high enough for the efficient level of overt child labor to be positive, the government will also impose a legal ceiling, lower than the efficient level, on such labor. This distorts the mix of overt and covert child labor and thus makes child labor as a whole less attractive relative to education. Furthermore, the government makes a transfer decreasing in parent income to every school child and exacts a transfer decreasing in the individual education result from every school leaver. The first transfer can be interpreted as the difference between a need-based education grant, covering maintenance and tuition, and an earmarked tax increasing in parent income. Similarly, the second transfer can be interpreted as the difference between an earmarked tax, equal to (the capitalized value of) the need-based education grant, and a merit-based education grant increasing in the school result. The first transfer may be negative for school-age children from families with high parent income. If the expected return to effective education is high enough, however, this transfer may be positive for anyone. The second transfer may be negative for school leavers with a high education result. In the model the first transfer occurs at the beginning of the education process, and the second at the end. In practice, however, the government could deliver the need-based grant and collect the tax on parent income in installments over the education period. Similarly, it could deliver the merit-based grant in installments over the education period, as partial results become available, and collect the tax on school leavers, again in installments, as the individual skill premia gradually unfold.

The analysis here is tailored for a poor developing country; it may be interesting to compare the results with those of a model tailored for a rich developed economy. Hanushek, Leung, and Yilmaz (2003) use a calibrated general equilibrium model to assess the welfare effects of a range of policy instruments, including need- and merit-based education grants, under the assumptions that child labor is out of the question and that parents are rich enough to be risk neutral (or, equivalently, that there is a well developed insurance market). In such a world, education subsidies generally perform worse than other forms of redistribution, and a merit-based education grant can be justified only in the presence of an education externality (while the current study finds that it is optimal anyway). These differences highlight the importance of the stage of development in designing education policy.

In most countries primary school enrollment is compulsory, and work at a very young age is forbidden (though enforcement is not always effective). In poor developing countries education is subsidized only through the price of school enrollment, if at all. Is that better than nothing? That question is best answered in two steps. First, starting from laissez-faire, would compulsory school enrollment raise social welfare? The answer is no, because it would oblige all parents, including those who would not let their children study anyway, to bear the tuition cost. Forbidding child labor instead, or on top, of that would also reduce welfare, because the ban would apply only to overt child labor, thereby distorting time allocation. Second, given compulsory enrollment, and with or without a ban on child labor, would a price subsidy raise welfare? If the subsidy is financed by a poll tax, the policy will affect welfare to the extent that the number of children varies across families. If all families had the same number of children, the policy would have no effect, because the parents would be taking a lump-sum subsidy with one hand and giving it back with the other. If the number varies exogenously across families, the policy will affect welfare, but the effect will be positive only if the marginal utility of income is higher in families with many children than in families with few children (in other words, if income and fertility are not positively correlated). If fertility is endogenous, the policy could actually reduce welfare, because it will trigger a substitution of quantity for quality of children; see Cigno (1986). If the subsidy is financed by a tax increasing in parent income, the net transfer schedule will look almost like $\gamma(.)$, though not exactly the same because a price subsidy cannot be larger than the price, and may thus be insufficient for a second-best policy. In any case, a second-best policy would also require some form of insurance against the risk of a low return to effective education. In other words, the $\theta(.)$ schedule is needed too.

Finally, considerable attention has been given to schemes that effectively pay children to attend school, such as Mexico's Programa de Educación, Salud y Alimentación.[23] Skoufias and Parker (2001) find evidence that such schemes encourage school attendance and discourage child labor. If the nonobservable determinants of effective education were positively correlated with the observable ones, one could be confident that offering transfers conditional on observable determinants would encourage nonobservable determinants. But there is evidence that the correlation is actually negative. Ravallion and Woodon (2000) report that the increase in school attendance elicited by an enrollment subsidy is four to eight times larger than the corresponding reduction in child labor. Consistent with this finding, Fuwa and others (2009) estimate that a credit constraint reduces average school attendance by 60 percent, but raises child labor by double that percentage. Why? The answer given by the current study's model is that paying a child to attend school triggers a substitution away from not only labor but also homework and rest. This has an efficiency cost and may actually reduce effective education time. The model further shows that paying a child to attend school will crowd out parent transfers (parents will give their children less money or take more money away from them).

In the light of these theoretical results and empirical findings, cash transfers should be made conditional not only on the child attending school, but also on the child doing no more than a certain amount of overt labor (less than the efficient amount, if that is positive). Furthermore, cash transfers to children in the scheme should be increasing in education results and decreasing in parent income. Such corrections would improve the scheme but would not be enough for a second-best policy, because the parents would still get no insurance against

[23] For a comprehensive exposition, see Fiszbein, Schady, and Ferreira (2009).

the risk of a low-skill premium, let alone against the risk of a negative shock to their own income. Of course, the distance from the second best will be even greater if parent income is private information or if overt child labor is not overt after all, because it will then be impossible to make cash transfers conditional on either or both of these variables.

The optimal taxation approach adopted in this study gives new insights into how best to discourage labor at a very young age and provide all children an effective education, in poor developing countries. One such insight is that subsidizing school attendance without rewarding school attainment at the same time is not optimal and may even be counterproductive. Another is that, in a second-best perspective, it is optimal to force overt child labor below its efficient level, if this is positive, despite the fact that (indeed, precisely because) covert child labor cannot be similarly regulated. The analysis lends support to the notion that school enrollment should be made compulsory but not necessarily to the notion that overt child labor should be banned. These results should help in designing a universal education system and in improving partial forms of public intervention such as conditional cash transfers.

References

Alderman, H., P. F. Orazem, and E. M. Paterno. 2001. "School Quality, School Cost, and the Public/Private School Choices of Low-Income Households in Pakistan." *Journal of Human Resources* 36: 304–26.

Baland, J. M., and J. A. Robinson. 2000. "Is Child Labor Inefficient?" *Journal of Political Economy* 108: 663–79.

Balestrino, A., A. Cigno, and A. Pettini. 2002. "Endogenous Fertility and the Design of Family Taxation." *International Tax and Public Finance* 9: 175–93.

Beegle, K., R. H. Dehejia, and R. Gatti. 2006. "Child Labor and Agricultural Shocks." *Journal of Development Economics* 81: 80–96.

Belzil, C., and J. Hansen (2002). "Unobserved Ability and the Return to Schooling." *Econometrica* 70: 2075–91.

Besley, T. 1995. "Non-Market Institutions for Credit and Risk-Sharing in Low-Income Countries." *Journal of Economic Perspectives* 9: 115–27.

Card, D. 1999. "The Causal Effect of Education on Earnings." In O. Ashenfelter and D. Card, eds., *Handbook of Labor Economics.* Vol. 3. Amsterdam: North Holland.

Cigno, A. 1986. "Fertility and the Tax-Benefit System: A Reconsideration of the Theory of Family Taxation." *Economic Journal* 96: 1035–51.

Cigno, A. 1993. "Intergenerational Transfers Without Altruism: Family, Market and State." *European Journal of Political Economy* 9: 505–18.

Cigno, A. 2001. "Comparative Advantage, Observability, and the Optimal Tax Treatment of Families With Children." *International Tax and Public Finance* 8: 455–70.

Cigno, A. 2006. "A Constitutional Theory of the Family." *Journal of Population Economics* 19: 259–83.

Cigno, A. 2011. "Agency in Family Policy: A Survey." *CESifo Economic Studies* 57: 305–31.

Cigno, A., and A. Luporini. 2009. "Scholarships or Student Loans? Subsidizing Higher Education in the Presence of Moral Hazard." *Journal of Public Economic Theory* 11: 55–87.

Cigno, A., and F. C. Rosati. 2005. *The Economics of Child Labour.* New York and Oxford: Oxford University Press.

Cremer, H., A. Dellis, and P. Pestieau. 2003. "Family Size and Optimal Income Taxation." *Journal of Population Economics* 16: 37–54.

Dessy, S. E., and S. Pallage. 2005. "A Theory of the Worst Forms of Child Labor." *Journal of Development Economics* 115: 68–87.

Dessy, S. E., and T. Rambeloma. 2010. "Is Temporary Emigration of Unskilled Workers a Solution to the Child Labor Problem?" Cahiers de Recherche 1037. Centre Interuniversitaire sur le Risque, les Politiques Économiques et l'Emploi, Montréal, Canada.

Epstein, G. S., and N. Kahana. 2008. "Child Labor and Temporary Emigration." *Economics Letters* 99: 545–48.

Fasih, T. 2008. *Linking Education Policy to Labor Market Outcomes.* Washington, DC: World Bank.

Fiszbein, A., N. Schady, and F. H. G. Ferreira. 2009. *Conditional Cash Transfers: Reducing Present and Future Poverty.* Washington, DC: World Bank.

Fitzsimons E. 2007. "The Effects of Risk on Education in Indonesia." *Economic Development and Cultural Change* 56: 1–25.

Fuwa, N., S. Ito, K. Kubo, T. Kurosaki, and Y. Sawada. 2009. "How Does Credit Access Affect Children's Time Allocation? Evidence from Rural India." Discussion Paper 183. Japan External Trade Organization, Institute of Developing Economies, Chiba, Japan.

Hanson, G. H., and C. Woodruff. 2003. "Emigration and Education Attainment." University of California–San Diego.

Hanushek, E. A., C. K. Y. Leung, and K. Y. I. Yilmaz. 2003. "Redistribution through Education and Other Transfer Mechanisms." *Journal of Monetary Economics* 50: 1719–50.

Jacoby, H. 1994. "Borrowing Constraints and Progress Through School: Evidence from Peru." *Review of Economics and Statistics* LXXVI: 151–60.

Jacoby, H., and E. Skoufias. 1997. "Risk, Financial Markets and Human Capital in Developing Countries." *Review of Economic Studies* 64: 311–35.

Johnson, W. R. 1987. "Income Redistribution as Human Capital Insurance." *Journal of Human Resources* 22: 269–80.

Kahneman, D. 2003. "Maps of Bounded Rationality: Psychology for Behavioral Economics." *American Economic Review* 93: 1449–75.

Kanbur, R., J. Pirttilä, and M. Tuomala. 2008. "Moral Hazard, Income Taxation and Prospect Theory." *Scandinavian Journal of Economics* 110: 321–37.

Kodde, A. D. 1986. "Uncertainty and the Demand for Education." *Review of Economics and Statistics* 68: 460–67.

Levhari D., and Y. Weiss. 1974. "The Effect of Risk on the Investment in Human Capital." *American Economic Review* 64: 950–63.

Loury, G. 1981. "Intergenerational Transfers and the Distribution of Earnings." *Econometrica* 49: 843–67.

Oreopoulos, P. 2006. "Average Treatment Effects of Education when Compulsory School Laws Really Matter." *American Economic Review* 96: 152–75.

Pouliot, W. 2006. "Introducing Uncertainty in Baland and Robinson's Model of Child Labour." *Journal of Development Economics* 79: 264–72.

Ram, R., and T. Schultz. 1979. "Life Span, Health, Savings and Productivity." *Economic Development and Cultural Change* 13: 399–421.

Ranjan, P. 2001. "Credit Constraints and the Phenomenon of Child Labor." *Journal of Development Economics* 64: 81–102.

Ravallion, M. and Q. Woodon. 2000. "Does Child Labour Displace Schooling? Evidence on Behavioural Responses to an Enrolment Subsidy." *Economic Journal* 110:c158-c175.

Razin, A. 1976. "Lifetime Uncertainty, Human Capital and Physical Capital." *Economic Enquiry* 14: 439–48.

Skoufias, E., and S. Parker. 2001. "Conditional Cash Transfers and the Impact on Child Work and Schooling. Evidence from Mexico's Progresa Program." *Economia* 2: 45–96.

Townsend, R. 1994. "Risk and Insurance in Village India." *Econometrica* 62: 539–92.