

IZA DP No. 5663

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

Alessandro Cigno

April 2011

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

**Alessandro Cigno**

*University of Florence  
and IZA*

Discussion Paper No. 5663  
April 2011

IZA

P.O. Box 7240  
53072 Bonn  
Germany

Phone: +49-228-3894-0  
Fax: +49-228-3894-180  
E-mail: [iza@iza.org](mailto:iza@iza.org)

Any opinions expressed here are those of the author(s) and not those of IZA. Research published in this series may include views on policy, but the institute itself takes no institutional policy positions.

The Institute for the Study of Labor (IZA) in Bonn is a local and virtual international research center and a place of communication between science, politics and business. IZA is an independent nonprofit organization supported by Deutsche Post Foundation. The center is associated with the University of Bonn and offers a stimulating research environment through its international network, workshops and conferences, data service, project support, research visits and doctoral program. IZA engages in (i) original and internationally competitive research in all fields of labor economics, (ii) development of policy concepts, and (iii) dissemination of research results and concepts to the interested public.

IZA Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be available directly from the author.

## **ABSTRACT**

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

As credit and insurance markets are imperfect, and given that intra-family transfers, and the way a child uses her time outside school hours, are private information, the second-best policy makes school enrollment compulsory, forces overt child labour 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.

#### **NON-TECHNICAL SUMMARY**

The optimal taxation approach adopted in the present paper gives us new insights into how best to discourage labour at a very young age, and give 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 labour below its efficient level, if this is positive, despite the fact that (indeed, precisely because) covert child labour cannot be similarly regulated. The analysis lends support to the notion that school enrollment should be made compulsory, but not necessarily to the one that overt child labour should be banned. These results should be of help, not only in designing a universal education system, but also in improving partial forms of public intervention such as conditional cash transfers.

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

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

Corresponding author:

Alessandro Cigno  
Dipartimento di Studi sullo Stato  
Università di Firenze  
Via delle Pandette 21  
50127 Firenze  
Italy  
E-mail: [cigno@unifi.it](mailto:cigno@unifi.it)

# 1 Introduction

Developing country governments and international development agencies have long been aware that human, more than physical capital accumulation, is the mainspring of economic and civil progress. Yet, many children in poor developing countries fail to complete even the primary education cycle, and some do not go to school at all. The reasons are well known.<sup>1</sup> Baland and Robinson (2000) demonstrate that child labour will be inefficiently high if parents are either credit or bequest constrained.<sup>2</sup> Evidence that parental inability to borrow discourages education, and encourages child labour, is reported by a host of authors, including Jacoby (1994) and Fuwa et al. (2009). Loury (1981) and Pouliot (2006) demonstrate that parental inability to insure against the risk of a low return causes educational investment to be inefficiently low, and child labour inefficiently high, even if credit is not rationed, and bequests are interior.<sup>3</sup> For evidence that parental inability to insure against the risk of a low return discourages education in developing countries, see Ram and Schultz (1979), and Jacoby and Skoufias (1997). Parental incomes also may be uncertain. Evidence in Beegle et al. (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, but in village aggregate income, suggesting that idiosyncratic income shocks are neutralized by informal insurance arrangements at the local level.<sup>4</sup>

As idiosyncratic shocks wash out on average, governments face less risk than individual households. Partly but not only for this reason, governments have also easier access to international money markets than most of their citizens. In the presence of imperfect domestic credit and insurance markets, there is then an efficiency argument for governments to lend to, and insure, parents of school-age children in some form. In view of evidence, in Kodde (1986) and elsewhere, of diminishing absolute risk aversion in an educational context, and given an unequal distribution of parental wealth, there is also an equity argument. It has been argued that efficiency-enhancing policies are politically easier to implement if they do not involve redistribution, but it is difficult to see how they could. Even if the government could finance the education of poor children entirely by borrowing against the future tax payments of these children, insuring families against the risk of a low return to education would still imply redistribution

---

<sup>1</sup>For a systematic exposition, see Cigno and Rosati (2005).

<sup>2</sup>Cigno (1993, 2006) shows, however, that the problem is alleviated if there exists a set of self-enforcing, renegotiation-proof family rules obliging working-age family members to support their young children and elderly parents.

<sup>3</sup>According to Levhari and Weiss (1974), the return to education is uncertain because a child's learning ability is fully revealed only after the educational investment is carried out. For 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 concerns also the length of time for which the future adult will be able to enjoy the benefit.

<sup>4</sup>Evidence of such arrangements in a developing country is reported by, among others, Besley (1995) and Townsend (1994).

from rich to poor school-leavers. Similarly, insuring families against the risk of a downturn in parental income would involve redistribution from rich to poor parents.<sup>5</sup> The tension between equity and efficiency will be minimal in the case of a small-scale project, especially if this is financed in large part by international aid,<sup>6</sup> but not in the case of a large-scale one. It is also difficult to imagine that any project, large or small, could be supported by the international community for ever. Informational 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 labour involves physically damaging or morally degrading activities. These are the "worst forms" of child labour national governments are committed by international treaty to eradicate. But the great bulk of it consists of activities (like helping in the home, working in the family farm, contributing to the family business) conducted for and under the direct supervision of the children's own parents.<sup>7</sup> While comparatively harmless in themselves, these activities conflict with education, and have thus an opportunity-cost in terms of forgone future earnings. The government might want to regulate them, but it cannot because they are private information. There is then a moral hazard problem. Similar considerations apply if we look at the issue from the educational standpoint. Every teacher knows that the scholastic performance of a child of any given learning ability depends not only on how much time the latter spends attending lessons, but also on how much time she spends on her homework, and how alert and well rested she is doing both. If a child falls asleep during lessons, and does not find the time, or is too tired, to give her homework the necessary attention, her school results will not be as good as those of a child of the same learning ability who comes to school well rested, and with her homework conscientiously done. As school enrollment and school attendance are common knowledge, but much of what a child does outside school is private information, we have then another moral hazard problem. A problem of the same nature arises also from the fact that intra-family transfers are private information, because the government cannot then be sure that a public subsidy intended for the children will not end up as extra consumption for the parents.

The present paper sets out to characterize the second-best policy in the face of moral hazard, and to compare it with two benchmarks, a low one represented by *laissez faire*, and a high one represented by the first best policy, in a situation where (a) parents can neither borrow nor insure, and (b) parents are better informed than the government about their children's time allocation, (c) the government does not observe intra-household transfers. The analysis rests on the assumption that the expected return to education is positive.<sup>8</sup> A recent

---

<sup>5</sup>See Johnson (1987).

<sup>6</sup>It will not even arise if education is privately financed by migrant remittances. Dessy and Rangeloma (2010), Epstein and Kahana (2008), and Hanson and Woodruffs (2003), report evidence that such remittances reduce child labour in the families left behind.

<sup>7</sup>See Cigno and Rosati (2005).

<sup>8</sup>For evidence of a causal effect of education on future earnings, see Card (1999) and Oreopoulos (2006).

World Bank survey, Fasih (2008), reports evidence of high returns to education especially in low and middle-income countries. The same study reports, however, that these returns are lower for the poor than for the rich. That may be taken as a sign that the poor can only afford poor-quality schools,<sup>9</sup> but our line of reasoning suggests another (not necessarily rival) explanation, namely that children from poor families study less, or less effectively, per year of school enrollment or day of school attendance, than children from rich families do. As the worst forms of child labour raise moral issues that transcend the materialistic calculations underlying the present paper,<sup>10</sup> we leave them out of the formal analysis, but will argue that the proposed policy reduces child labour in all its forms.

The policy optimization has an optimal taxation, or principal-agent, format.<sup>11</sup> 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 parental control, we take their parents, rather than the children themselves, to be the agents. The latter are modelled as risk-averse, expected-utility maximizers.<sup>12</sup> As the implications of an educational externality are well understood, and allowing for it would merely reinforce the argument for public intervention, the analysis abstracts from it (but we will find that the policy itself gives rise to a fiscal externality). As 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 for ever, we impose the constraint that the policy must be self-financing. Section 2 of the paper lays down the technical assumptions, and characterizes parental decisions. Section 3 examines the *laissez faire* equilibrium. Section 4 derives the first and second best policies. Section 5 discusses actual policy practice (including conditional cash transfer schemes) in the light of the theoretical findings, and concludes.

---

<sup>9</sup>On the subject, see Alderman et al. (2001).

<sup>10</sup>But, see Dessy and Pallage (2005) for a strictly economic analysis.

<sup>11</sup>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 the field of higher education, see Cigno and Luporini (2009).

<sup>12</sup>That is not the only possible representation of individual behaviour 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 some empirical justification, we refrain from going down that road here for two reasons. The first is that, in a situation where most people live little above the subsistence level, we are not likely to observe much risk-loving behaviour. The second is that the policy maker may not approve of such behaviour, and consequently maximize an objective function which is not a mere aggregation of the individual ones. Kanbur et al. (2008) show that, if the government corrects for what it considers an aberrant behaviour, 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.

## 2 Families

There is a large number of families, labelled  $i = 1, 2 \dots n$ . Each family consists 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,<sup>13</sup> and do not impinge on the points at issue here. Learning ability is randomly distributed across children, and imperfectly observable by all concerned until the educational investment is carried out. We take parental income to vary exogenously across families, and to be observable by the government. Later in the paper, we will allow for this income to be either uncertain, or private information.

There are two periods, labelled  $t = 1, 2$ . Children are alive in both, parents only in the first one. For brevity, we refer to the child in the  $i$ th family 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

$$V_i = v(a_i) + \beta U_i, \quad 0 < \beta < 1,$$

where  $a_i$  denotes parental consumption, and  $\beta$  is a measure of altruism. The functions  $u(\cdot)$  and  $v(\cdot)$  are assumed increasing and concave, with  $u'(0) = v'(0) = \infty$ . In an uncertain environment, concavity implies risk aversion. The assumption that marginal utility becomes very large as consumption approaches zero may be taken to imply that subsistence consumption is normalized to zero. As utility does not depend on time allocation, we are implicitly saying that leisure is not a good, and that work does not yield direct disutility. This may be justified by saying that, as the worst forms of child labour are out of the picture, and consumption is likely to be low, the marginal utility of income is likely to be high relative to 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 total cost of tuition.<sup>14</sup> In the last section, we argued that effective education time is increasing in school attendance, homework and rest time, and decreasing in work time. If  $i$  is enrolled, her effective

---

<sup>13</sup>If the number of children is exogenous, and parental income or work effort are private information, the optimal income tax rate is in fact zero for the family with the highest income. If it is exogenous, but varies across families as in Cremer et al. (2003), the optimal policy will redistribute in favour 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 et al. (2002), the first one (no distortion at the top) because the children's visibility makes mimicking much harder, the second one (children reduce tax liability) because children yield utility.

<sup>14</sup>Tuition fees are usually per year (or shorter period like the semester) of school enrollment. Therefore, the total amount of money 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 for which the child studies at home, in each of those years. We account for this lumpiness of tuition fees by treating  $p$  as a constant.

education time will be positive (or there would be no point in paying  $p$ ). To simplify, we measure effective education time as the amount of time that the child does not work. If  $i$  is not enrolled, her effective education time will be zero.<sup>15</sup> Child labour may be overt or covert. The former consists of work done for an employer other than the child's own parents, and carries a wage. The latter involves either participating in a family-run, income-generating activity, such as farming and retailing, or replacing the child's parents in the performance of household chores such as cooking, cleaning, fetching water and gathering fuel, etc.<sup>16</sup> Although neither of these forms of covert child labour 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  $L_i$  her overt labour. Normalizing a child's time endowment to unity,  $i$ 's covert child labour is then  $(1 - e_i - L_i)$ .

As children can contribute to the production of family income, we define parental income in period 1 as family income if child labour in both its forms were zero. Let  $y_i$  denote parental income in family  $i$ . The income generated by overt child labour is  $L_i w_1$ , where  $w_1$  is the child wage rate. That produced by covert child labour is  $z(1 - e_i - L_i)$ , where  $z(\cdot)$  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 the performance of household chores. Concavity reflects diminishing marginal rates of technical substitution between adult and child work. The assumption that the marginal revenue gets very large as covert child labour gets very small is realistic in a poverty context like the present one, and ensures that this kind of labour will never be zero. In period 2,  $i$  will earn  $w_2 + x_i$ , where  $w_2$  denotes the income of an unskilled adult, and  $x_i$  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 his parents,<sup>17</sup> and another,  $\gamma_i$ , from the government. In period 2,  $i$  will make the government a transfer  $\theta_i$ . All these transfers can be positive, negative or zero.

Parents take their decisions in period 1, after the government has announced its policy. Anticipating a result that will obtain in Section 4, we take  $\gamma_i$  to be a function of  $y_i$ , and  $\theta_i$  to be a function of  $x_i$ . While  $w_1$ ,  $w_2$  and  $\gamma_i$  are certain,  $x_i$

<sup>15</sup>In many developing countries, a substantial minority of school-age children is reported as neither working nor studying. But 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).

<sup>16</sup>See Cigno and Rosati (2005) for an analysis of the incidence of these activities, and for the effects of making the last two unnecessary by providing homes with mains electricity, and running water.

<sup>17</sup>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 us to leave intra-family transfers out of the picture, and treat all costs and revenues as pertaining to the family as a whole, but it would be misleading, because we know from Baland and Robinson (2002) that transfers from parents to children may be inefficiently low.

and consequently  $\theta_i$  are uncertain. As  $e_i$  must be chosen in period 1, education is then a risky investment. We make the supplementary assumption (to be relaxed later in the paper) that  $x_i$  is i.i.d. over the closed interval  $[0, \bar{x}] \in R^+$  with density  $f(\cdot|e_i)$  conditional on  $e_i$ , and  $f(\cdot|0) = 0$ . To simplify the notation, we use  $x_i$  to measure both the final school result and the skill premium.<sup>18</sup> 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  $e_i$ ,

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

In other words, the more  $i$  studies, and the less she works, the more of a chance she 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 (1) holds as an inequality. We make the standard convexity-of-distribution-function (CDF) assumption, that  $F(x_i|e_i)$  is convex in  $e_i$ , and monotone-likelihood-ratio (MLR) assumption, that  $\frac{f_{e_i}(\cdot|e_i)}{f(\cdot|e_i)}$  is increasing in  $x_i$ ,<sup>19</sup> which allow us to adopt the first-order approach.

If  $i$  enrolls for school, and overt child labour is not regulated by 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

$$e_i \geq 0, \quad (2)$$

$$L_i \geq 0 \quad (3)$$

and

$$1 - e_i - L_i \geq 0. \quad (4)$$

As (4) will never be binding for the restrictions imposed on the revenue function, the first-order conditions on the choice of  $e^i$ ,  $L^i$  and  $m^i$  are, respectively,

$$-v'_i z'_i + \beta \int_{x_i} u_{i2} f_{e_i}^i dx_i + \xi_i = 0, \quad (5)$$

$$-v'_i z'_i + \psi_i + \beta u'_{i1} w_1 = 0 \quad (6)$$

and

$$-v'_i + \beta u'_{i1} = 0, \quad (7)$$

---

<sup>18</sup>Using a 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 difference of substance to our results so long as both are i.i.d., and the skill premium is not conditional on some decision variable.

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

where  $\xi_i$  is the Lagrange-multiplier of (2), and  $\psi_i$  that of (3). If  $i$  does not enroll,  $e_i$  cannot be positive. Again assuming that overt child labour is free to vary,  $i$ 's parents will then choose  $(L_i, m_i)$  so as 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 (3) – (4). The solution will satisfy (6) – (7) for  $p \equiv e_i \equiv 0$ . If  $L_i$  is regulated by the government, (6) need not hold. Irrespective of whether  $i$  is or is not enrolled, and  $L_i$  is or is not regulated, it is clear from (7) that  $m_i$  is decreasing in  $\gamma_i$ . In other words, public transfers crowd out private transfers.

### 3 Laissez faire

In *laissez faire*, school enrollment is not compulsory, overt child labour is free to vary, and

$$\gamma^i \equiv \theta^i \equiv 0.$$

The pay-off of enrolling  $i$  at school is

$$\pi^S(y_i, p) \equiv \max_{(L_i, e_i, m_i)} E(V^i), \text{ s.t. (2) – (4)}. \quad (8)$$

That of not enrolling her is

$$\pi^W(y_i) \equiv \max_{(L_i, m_i)} V(L_i, m_i), \text{ s.t. (2) – (4)}. \quad (9)$$

The child will be enrolled if and only if  $\pi^S(y_i, p)$  is at least as large as  $\pi^W(y_i)$ . There is then a threshold value of  $y_i$ ,  $\tilde{y}$ , defined by

$$\pi^S(\tilde{y}, p) = \pi^W(\tilde{y}),$$

below which  $i$  will not be enrolled. As  $\tilde{y}$  is the same for every  $i$ , because the expected return to education is the same for all of them, if any children are not enrolled, it will then be those whose parents have a low income. This result differs from the one in Ranjan (2001), where a child's learning ability is assumed to be directly observable *ex ante*, and the threshold is consequently lower for parents of high-ability, than for parents of low-ability children.

Given that  $\xi_i$  will be zero if  $e_i$  is positive, (5) implies

$$\text{either } e_i = 0 \text{ or } v'_i z'_i = \beta \int_{x_i} u_{i2} f_{e_i}^i dx_i. \quad (10)$$

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

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

Therefore,  $L_i$  is either zero, or positive and increasing in  $w_1$ . It is then clear that overt child labour is the same in all families.<sup>20</sup> What differs is effective education and total (overt plus covert) child labour.

*Proposition 1. In laissez faire,*

- (i) *children from very poor families are not enrolled at school;*
- (ii) *children from less poor families are enrolled, but their effective education increases with parental income;*
- (iii) *overt child labour is either zero, or increasing in the child wage rate.*

The second part of this proposition provides a possible explanation for the empirical finding, cited in Section 1, 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 rich children do, because it says that the former receive less effective education, in that extra year or day, than the latter do.

## 4 Government

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

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

As parents and children are many, and having assumed that risks are uncorrelated, the government does not face any uncertainty about its tax revenue. For this and other reasons, the government has easier access to international credit than any of its citizens taken individually. Making the usual "small country" assumption, we will treat the real interest rate as a constant, and normalize it to zero. As the expected return to education is the same for every  $i$ , and assuming this to be positive, the government will then make school enrollment compulsory. As the optimization can determine only relative tax rates, we will normalize the tax on  $w_2$  to zero, and look for the socially optimal values of  $\gamma^i$  and  $\theta^i$ .

As it does not face budget uncertainty, the government will choose  $(e^i, L^i, m^i, \gamma^i, \theta^i)$ , for  $i = 1, 2, \dots, n$ , so as to maximize (12), subject to the budget constraint,

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

and (2) – (4). If  $(e^i, m^i)$  is private information, the maximization will be also incentive-compatibility constraints. As  $E(V^i)$  is concave in  $(e^i, L^i, m^i)$ ,  $SW$  will be concave in it too. For the i.i.d. assumption, the optimal  $(\gamma^i, \theta^i)$  can depend only on  $(e^i, m^i, x^i, y^i)$ , and not also on any  $(e^j, m^j, x^j, y^j)$  for  $j \neq i$ .

<sup>20</sup>It would vary across families if the  $z()$  function did (e.g., if the return to covert child labour were higher in a farming family that owns land, than in one that does not).

## 4.1 First best

In first best, the government prescribes  $(e_i, L_i, m_i)$ , and designs personalized lump-sum transfers,  $(\gamma_i, \theta_i)$  for each  $i$ . As there are no incentive-compatibility constraints, and denoting by  $\lambda$  the Lagrange-multiplier of (13), the first-order conditions on  $L_i$ ,  $m_i$ ,  $e_i$ ,  $\gamma_i$  and  $\theta_i$  are, respectively, (6), (7),

$$-v'_i z'_i + \int_{x_i} (\beta u_{i2} + \lambda \theta_i) f_{e_i}^i dx_i + \xi_i = 0, \quad (14)$$

$$\beta u'_{i1} - \lambda = 0 \quad (15)$$

and, for each possible realization of  $x_i$ ,

$$-(\beta u'_{i2} - \lambda) f^i = 0. \quad (16)$$

As (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 in *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 in *laissez faire* for every  $i$ .

In view of (7), (15) and (16), it is also clear that

$$a_i = a^{FB}, \quad c_{i1} = c_{i2} = c^{FB} \quad \text{and} \quad m_i = m^{FB}.$$

As this implies that parental income is equalized across families, and given that children are ex-ante identical, the first-best level of  $e_i$  is the same for every  $i$ ,

$$e_i = e^{FB}.$$

*Proposition 2. In first best,*

- (i) *the government uses lump-sum taxes and subsidies to achieve perfect equity, perfect consumption smoothing, and full insurance;*
- (ii) *all school-age children allocate their time in the same way;*
- (iii) *overt child labour is either zero, or increasing in the child wage rate;*
- (iv) *each school-age child receives more effective education than in laissez faire.*

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

## 4.2 Second best

In second best,  $(e_i, m_i)$  is private information. According to the logic of optimal taxation, the government will then make school enrollment compulsory, fix  $L_i$ , and influence parental decisions by announcing how  $\gamma_i$  and  $\theta_i$  will be related to the information available in the relevant period. As  $\gamma_i$  is payable in period 1, it can depend only on  $y_i$ . As  $\theta_i$  is payable in period 2, it can depend also 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 (12) is subject not only to (2) – (4) and (13), but also to the incentive-compatibility constraints (5) and (7). Let  $\phi_i$  denote the Lagrange-multiplier of (5), and  $\mu_i$  that of (7). The first-order conditions on  $e_i$ ,  $L_i$ ,  $m_i$ ,  $\gamma_i$  and  $\theta_i$  are, respectively,

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

$$-v'_i z'_i + \beta u'_{i1} w_1 + \psi_i + \phi_i \left[ v'_i z''_i + v''_i (z'_i)^2 \right] + \mu_i [v''_i z'_i + \beta u''_{i1} w_1] = 0, \quad (18)$$

$$-v'_i + \beta u'_{i1} + \phi_i v''_i z'_i + \mu_i [v''_i + \beta u''_{i1}] = 0, \quad (19)$$

$$\beta u'_{i1} - \lambda + \mu_i \beta u''_{i1} = 0 \quad (20)$$

and, for each possible realization of  $x_i$ ,

$$-(\beta u'_{i2} - \lambda) f^i - \phi_i \beta u'_{i2} f_{e_i}^i = 0. \quad (21)$$

Using (7), we can re-write (20) 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. So long as  $r_i$  is non-increasing in  $i$ 's income,<sup>21</sup> and given that  $v'_i$  is decreasing in  $y_i$ , we can then write

$$\gamma_i = \gamma(y_i), \quad \gamma' < 0. \quad (22)$$

Condition (21) may be similarly re-written as

$$1 + \phi_i \frac{f_{e_i}}{f^i} = \frac{\lambda}{\beta u'_{i2}}.$$

<sup>21</sup>See the evidence cited in Section 1, and the discussion in footnote 11.

As  $\frac{f_{e_i}}{f^i}$  is increasing in  $x_i$ , and  $u'_{i2}$  in  $\theta_i$ , we can then write

$$\theta_i = \theta(x_i), \theta' < 0. \quad (23)$$

As there is nothing to prevent  $\gamma_i$  from falling below zero for some  $y_i$ , we can interpret the former as the difference between an education grant equal to  $p$ ,<sup>22</sup> and an earmarked tax increasing in parental income. Similarly, as there is nothing to stop  $\theta_i$  being negative for some  $x_i$ , we can interpret the former as the difference between another earmarked tax, this time equal to  $p$ , and another education grant, this time increasing in the school result.

Having established that  $\gamma(\cdot)$  and  $\theta(\cdot)$  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 (20) with (15), and (21) with (16), however, it is also clear that the policy does not go as far as in first best. The reason is, of course, that redistribution has an efficiency cost, because the government cannot use personalized lump-sum transfers as in first best. Let us see what happens to  $(e_i, L_i, m_i)$ . Comparing (17) with (14) and (5), we can see that  $e_i$  is lower than in first best. In particular, as  $(y_i + \gamma_i)$  is not the same for all  $i$  as in first best,  $e_i$  increases with  $y_i$  as (albeit more slowly than) in *laissez faire*. As  $\gamma_i$  can be negative, we cannot even rule out that  $e_i$  will be lower than in *laissez faire* for some  $i$ . As the government can borrow against  $\int_{x_i} \theta_i f^i dx_i$ , however,  $\gamma_i$  will be negative only if  $y_i$  is very high. As 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 in *laissez faire* for any  $i$ . Comparing (18) with (6), we can also see that  $L_i$  will be no higher than in either first best or *laissez faire*. The intuition is that, if  $w_1$  is sufficiently high for the efficient  $L_i$  to be positive, imposing a ceiling on overt child labour will distort the allocation of  $i$ 's total working time between overt and covert labour, and thus make work as a whole less attractive in comparison with education. Comparing (19) with (7), finally, we can see that  $m_i$  will be lower than in either first best or *laissez faire*.

*Proposition 3. In second best,*

*(i) 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 in laissez faire;*

*(ii) the government uses a net subsidy decreasing in parental 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;*

*(iii) if the efficient level of overt child labour is positive, the government sets a limit, lower than the efficient level, on the amount of paid work a child can legally do.*

<sup>22</sup>Recall that subsistence consumption is normalized to zero.

The implications of relaxing some of the assumptions we have made so far can be intuited without formal analysis. Suppose that the returns to educational investments have an aggregate, as well as an idiosyncratic, component. As aggregate risks cannot be insured against by redistributing within cohorts, the government must then 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 parental incomes. If the shocks to parental income are purely idiosyncratic, the policy prescription remains qualitatively the same, because redistributing from rich to poor parents will insure families against the risk of a downturn in that income. The prescription remains the same also if the shocks have an area component, because the policy redistributes not only within, but also between areas. If the shocks have a country-wide 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 educational investment).

## 5 Discussion

In *laissez faire*, if credit and insurance markets are imperfect, or contracts between parents and young children are not enforceable, effective education is inefficiently low. If parental 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 relatively rich ones. This prediction is consistent with evidence in Ram and Schultz (1979), Jacoby (1994), Jacoby and Skoufias (1997), Belzil and Hansen (2002), and Fuwa et al. (2009), that inability to borrow and insure reduces educational investment. It is consistent also with evidence, surveyed in Fasih (2008), that the return to measurable educational inputs such as school enrollment or attendance is positive and particularly large in low to middle income countries, but lower for poor than for rich children. As the amount of effective education that a child receives in a year of school enrollment, or day of school attendance, is lower if she comes from a family with low, than if she comes from one with high, parental income, the return to enrollment or attendance will in fact understate the return to effective education of rich, relative to that of poor children. This explanation does not conflict with other possible explanations, such as the one that poor children have only access to poor quality schools.

The optimal (first or second-best) policy relaxes the credit constraint on educational 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. By redistributing from rich to poor parents, it also reduces inequality and, if parental income is uncertain, provides 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. As 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. As 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. It is worth noting that, although the worst forms of child labour are outside the scope of the analysis, a policy that encourages effective education will discourage covert child labour in all its forms, including the worst ones.

In second best, school enrollment is compulsory (that is unnecessary in first best, where it is the interest of all parents to send their children to school). If the child wage rate is sufficiently high for the efficient level of overt child labour to be positive, the government will also impose a legal ceiling, lower than the efficient level, on this type of labour. By distorting the mix of overt and covert child labour, this will help to make child labour as a whole less attractive compared with education. Furthermore, the government makes a transfer decreasing in parental income to every school child, and exacts a transfer decreasing in the individual skill premium from every school leaver. The first transfer may be negative for school-age children from families with high parental income, the second for school leavers with a low skill premium. If the expected return to effective education is sufficiently high, however, the first transfer is unlikely to be negative for anyone. We can interpret this transfer as the difference between a need-based education grant, covering maintenance and tuition, and an earmarked tax increasing in parental income. Similarly, we can interpret the other as the difference between an earmarked tax, equal to the need-based education grant, and a merit-based education grant increasing in the school result. In the model, the first transfer occurs "at the beginning", and the second "at the end" of the education process. In practice, the government could deliver the need-based grant, and collect the tax on parental 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.

As the analysis is tailored for a poor developing country, it may be interesting to compare our results with those of a model tailored for a rich developed economy. Hanushek *et al.* (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 assumption that child labour 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 in general perform less well than other forms of redistribution, and a merit-based education grant can be justified only in the presence of an education externality (while we find that it is optimal anyway). These differences highlight the importance of the stage of development in the design of education policy.

Let us now compare our analytical results with actual practice. Primary school enrollment is compulsory, and work at a very young age forbidden, in

most countries (but enforcement is not always effective). In poor developing ones, and education is subsidized, if at all, only through the price of school enrollment. Is that better than nothing? We will answer the question 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 cost of tuition. Forbidding child labour instead or on top of that would also reduce welfare, because the ban would apply only to overt child labour, and consequently distort time allocation. Second, given compulsory enrollment, and with or without a ban on child labour, 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 this number is the same for all families, the policy will have no effect, because the parents will take a lump-sum subsidy with one hand, and give it back with the other. Otherwise, the policy will affect welfare, but this effect will be positive only if the marginal utility of income is higher in families with many, than in families with few children, or the fiscal externality is sufficiently large.<sup>23</sup> If the subsidy is financed by a tax increasing in parental income, the net transfer schedule will look almost like our  $\gamma(\cdot)$ . Not quite, however, because a price subsidy cannot be larger than the price, and may thus be insufficient for a second best. In any case, a second best would require also some form of insurance against the risk of a low return to effective education; in other words, we need the  $\theta(\cdot)$  schedule too. Finally,

Considerable attention has been given to schemes, like Mexico's PROGRESA, which effectively pay children to attend school.<sup>24</sup> Skoufias and Parker (1991) find evidence that such schemes encourage school attendance and discourage child labour. If the non observable determinants of effective education were positively correlated with the observable ones, we could be confident that offering transfers conditional on the latter would encourage the former. 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 labour. Consistently with this finding, Fuwa et al. (2009) estimate that a credit constraint reduces average school attendance by 60 percent, but raises child labour by double that percentage. Why is that so? The answer given by our model is that paying a child to attend school triggers a substitution away from, not only labour, but also homework and rest. This has an efficiency cost, and may actually *reduce* effective education time. The model further tells us that paying a child to attend school will crowd out parental 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 then be made conditional not only on the child attending school, but also on the child doing no more than a certain amount (lower than the efficient one, if that is positive) of overt labour. Furthermore, cash transfers to children in the scheme should

<sup>23</sup>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).

<sup>24</sup>For a comprehensive exposition, see Fiszbein et al. (2009).

be increasing in school results, and decreasing in parental income. Such corrections would improve the scheme, but would not yield a second best, because the parents would still get no insurance against 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 parental income is private information, and "overt" child labour is not overt after all, because it will then be impossible to make cash transfers conditional on either of these variables.

The optimal taxation approach adopted in the present paper has given us new insights into how best to discourage labour at a very young age, and give 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 labour below its efficient level, if this is positive, despite the fact that (indeed, precisely because) covert child labour cannot be similarly regulated. The analysis lends support to the notion that school enrollment should be made compulsory, but not necessarily to the one that overt child labour should be banned. These results should be of help, not only in designing a universal education system, but also 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, 663–679
- Balestrino, A., A. Cigno and A. Pettini (2002), Endogenous fertility and the design of family taxation, *International Tax and Public Finance* 9, 175–193
- 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, 2075–2091
- Besley, T. (1995), Non-market institutions for credit and risk-sharing in low-income countries, *Journal of Economic Perspectives* 9, 115–127
- Card, D. (1999), The causal effect of education on earnings, in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, vol. 3, 1801–63, Amsterdam: North Holland
- Cigno, A. (1986), Fertility and the tax-benefit system: a reconsideration of the theory of family taxation, *Economic Journal* 96, 1035–1051
- (1993), Intergenerational transfers without altruism: family, market and state, *European Journal of Political Economy* 9, 505–518
- (2001), Comparative advantage, observability, and the optimal tax treatment of families with children, *International Tax and Public Finance* 8, 455–470

- (2006), A constitutional theory of the family, *Journal of Population Economics* 19, 259-283
- (2011), Agency in family policy: A survey, *CESifo Economic Studies* 58
- 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?, CIRPEE Cahiers de Recherche N° 1037
- Epstein, G. S. and N. Kahana (2008), Child labor and temporary emigration, *Economics Letters* 99, 545-548
- Fasih, T. (2008), *Linking Education Policy to Labor Market Outcomes*, Washington DC: World Bank Publications
- Fiszbein, A., N. Schady and F. H. G. Ferreira (2009), *Conditional Cash Transfers: Reducing Present and Future Poverty*, Washington DC: World Bank Publications
- 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, IDE Discussion Paper N° 183
- Hanson, G. H. and C. Woodruff (2003), Emigration and education attainment, UCSD mimeo
- 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, 151-160
- and E. Skoufias (1997), Risk, financial markets and human capital in developing countries, *Review of Economic Studies* 64, 311-335
- Johnson, W. R. (1987), Income redistribution as human capital insurance, *Journal of Human Resources* 22, 269-280
- Kahneman, D. (2003), Maps of bounded rationality: Psychology for behavioral economics, *American Economic Review* 93, 1449-1475
- Kanbur, R., J. Pirttilä and M. Tuomala (2008), Moral hazard, income taxation and prospect theory, *Scandinavian Journal of Economics* 110, 321-337
- Kodde, A. D. (1986), Uncertainty and the demand for education, *Review of Economics and Statistics* 68, 460-467

- Levhari D. and Y. Weiss (1974), The effect of risk on the investment in human capital, *American Economic Review* 64, 950-963
- 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, 264-272
- 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
- Razin, A. (1976), Lifetime uncertainty, human capital and physical capital, *Economic Enquiry* 14, 439-448
- Skoufias, E. and S. W. Parker (1991), Conditional cash transfers and their impact on child work and schooling: Evidence from the PROGRESA program in Mexico, IFPRI FCHND Discussion Paper N° 123
- Townsend, R. (1994), Risk and insurance in village India, *Econometrica* 62, 539-592