How to Deal with Covert Child Labour, and Give Children an Effective Education, in a Poor Developing Country: An Optimal Taxation Problem with Moral Hazard.*

Alessandro Cigno†
University of Florence, CESifo, CHILD and IZA

April 27, 2010

Abstract

As the return to education (and possibly also parental income) is uncertain, and given that the work a child does covertly for his own parents, and transfers between parents and children, are private information, the government should make school enrollment compulsory, set a legal limit (decreasing in parental income) on overt child labour, and redistribute across families using a flat-rate education grant, and a tax on parental income. That done, it should use a scholarship increasing in school results, and a tax on the skill premium, to raise the expected return to educational investment, and make it less uncertain.

Key words: child labour, education, uncertainty, moral hazard.

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

1 Introduction

Developing country governments and international development agencies have come to realize that human, rather than physical, capital accumulation is the mainspring of economic progress. At the same time, the international community is putting pressure on poor countries to curb child labour for humanitarian reasons. This raises three issues. The first is that, in a poor developing country, the efficient level of child labour is unlikely to be zero. The second is

*Work on this paper was completed while the author was visiting the Center of Intergenerational Studies of the Institute of Economic Research at Hitotsubashi University. Comments by Moriguki Chiaki, Kurosaki Takashi, Fuji Tomoki, Arimoto Yutaka, and other participants in an IER development economics seminar, are gratefully acknowledged. Remaning errors are the author's responsibility.

†Dipartimento di Studi sullo Stato, Università di Firenze, Via delle Pandette 21, 50127 Firenze. Email: cigno@unifi.it.

that, realistically assuming imperfect credit and insurance markets, and given the impossibility of a legally binding contract between parents and school-age children,[1] parental investment in their children's education may be inefficiently low, and child labour inefficiently high. Baland and Robinson (2000) show that child labour will be inefficiently high if parents are credit rationed, or bequests constrained to be zero (in other words, if parents would like to leave negative bequests or, more generally, make negative transfers to their grown-up children, but have no legal means to do so). Pouliot (2006) shows that, if parents cannot insure against the risk of a low return to their educational investments, child labour will be inefficiently high even if borrowing and bequests are interior. The third issue is one of observability. A very large part of the work children do takes place outside the market, and is private information.[2] A very small (but never small enough) proportion of this work involves physically or morally damaging activities. These are the "worst forms" of child labour, which national governments are committed by international treaty to eradicate. The rest, the great bulk of it, consists of activities (helping in the home, working in the family business) conducted for and under the supervision of the children's own parents.[3] While relatively harmless in themselves, these activities conflict with education,[4] and have thus an opportunity-cost in terms of forgone future earnings. As these activities are not observable by the government, however, they cannot be regulated.

Something similar may be said about education. School enrollment is common knowledge, but the total amount of time a child spends actually studying is not, because this total includes not only school attendance, but also time spent doing her homework, which is private information. The problem with regulating study time is only partly related to the observability of study time itself. If a child goes to sleep during lessons, the time he is asleep cannot be counted as effective study time. A teacher will notice if a child is prone to falling asleep. She will know also if a child is frequently absent (may be on grounds of actual or alleged ill health), or does not do his homework, and may take all this as evidence that the child does a substantial amount of work when he is not at school. If this work is private information, however, there is little she, or the government, can do about it. Therefore, a government may make school enrollment compulsory, but cannot oblige a child to study for any specified amount of time. A way round the problem is offered by schemes (like PROGRESA) which effectively pay children to attend school. As the payment is contingent on attendance, not on total study time, however, that is only a partial solution to the problem (we will show that the payment must depend, at least in part, on verifiable scholastic performance). Besides, as these schemes

[1] Cigno (2006) shows, however, that a self-enforcing, renegotiation-proof family constitution will serve the same purpose.

[2] One can get statistical estimates of it through household surveys, but the information cannot be used for regulatory purposes.

[3] See Cigno and Rosati (2005).

[4] That is a simplification. In developing countries, a large fraction of school-age children is reported by household surveys as neither working nor attending school; see Cigno and Rosati (2005) for estimates and explanations.

are not self-financing, they can only last as long as the general tax payer, and the international community, are willing to pay for them. Therefore, they are probably the best answer in a crisis situation, where a large number of children in a particular area is at risk of engaging in the worst forms of child labour, but cannot be extended to all children in the appropriate age range, and cannot last for ever.

The present paper sets out to derive the second-best policy, and to compare it with two benchmarks, a low one, represented by *laissez faire*, and a high one, represented by the first-best policy. We take the return to educational investment to be uncertain, either because, as in Levhari and Weiss (1974), a child's learning ability is fully revealed only after the investment is made, or because, as in Razin (1976), the rental price of the human capital accumulated through education is not known in advance. Parental income (family income net of the children's contribution) also may be uncertain, and the child labour repercussions of this uncertainty have received some attention in the empirical literature. As the results of introducing this second layer of uncertainty in the model makes little difference to the analytical results, however, we will only discuss the implications informally. We will assume that parents can neither insure, nor borrow against the expected return to their educational investment. Assuming that parents have no informational advantage over the government where this return is concerned,[5] however, there is no adverse selection problem. There is a moral hazard problem, on the other hand, because covert child labour, total study time, and transfers between parents and children, are private information. As the worst forms of child labour raise moral issues that transcend the materialistic calculations underlying the present paper,[6] we abstract from them. As the implications of an educational externality are well understood,[7] we abstract from that too (but we will find that the policy itself gives rise to a fiscal externality).

The policy optimization has an optimal taxation (or principal-agent) format. As we are talking of a poor developing country, school-age realistically means 5-6 to 11-12. It thus seems reasonable to assume that school-age children are under parental control, and thus that the agents are the parents of these children. In the logic of this type of problem, if an action is private information, the government must give agents the incentive to undertake it at the socially optimal (second-best) level. If an action is common knowledge, however, it does not make sense for the government to offer costly incentives, because the same result can be costlessly achieved by threatening the agents with a sufficiently severe penalty.[8] In our context, the actions falling into the first category are parental

[5] A child's learning ability is objectively revealed by education results, which are common knowledge. If there is any informational asymmetry, it will then be about the future state of the labour market. If that is the case, however, the informational advantage will rest with the government, not with the parents.

[6] But, see Dessy and Pallage (2005) for a strictly economic analysis.

[7] See, for example, Hanushek *et al.* (2003)

[8] In principal-agent language, this is called a "forcing contract". For a survey of the ways in which optimal taxation, or principal-agent, concepts are used in a family policy context, see Cigno (2009).

transfers, covert child labour, and the amount of time a child spends studying. Those falling into the second are overt child labour, and school enrollment. Section 2 of the paper sets out the basic assumptions regarding parents and children, and characterizes parental decisions. Section 3 examines the *laissez faire* equilibrium. Section 4 derives the first and second best policy. Section 5 examines the relevant evidence, and discusses the implications of introducing uncertainty about parental income, and aggregate risks. Section 6 discusses the policy results, and concludes.

2 Families

There is a large number of families, labelled $i = 1, 2...n$. Each family consists of a couple of adults, and one child. Neither the parents, nor the children, have access to credit and insurance markets. For brevity, we will refer to the child in the ith family as i. There are two periods, labelled $t = 1, 2$. Parents are alive only in the first period. Children are alive in both periods. In period 1, i is of school age. If she enrolls for school, she will divide her time between study, paid work in the child labour market, and unpaid work for her parents. If she does not enroll, she will divide her time between the two forms of work. The government observes whether i is or is not enrolled at school, and how much time she spends working in the labour market ("overt child labour"), but does not how much time she spends working for her parents ("covert child labour") , or actually studying. In period 2, i will be of working age, and work in the adult labour market.[9]

Ex post, i's utility will be given by

$$U_i = u(c_{i1}) + u(c_{i2}),$$

where c_t^i denotes i's consumption in period t. As this utility depends only on consumption, and not on time allocation, we are implicitly assuming that leisure is not a good, and that work does not yield direct disutility. The first assumption may be justified by saying that, at very low levels of income, the marginal utility of leisure is low too. The second is justified only because we are disregarding the worst forms of child labour. Assuming descending altruism, the ex-post utility of i's parents may written as

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

where a_i denotes parental consumption. The functions $u(.)$ and $v(.)$ are assumed increasing and concave, with $u'(0) = v'(0) = \infty$. The first assumption implies risk aversion, the second that subsistence consumption is normalized to zero. The constant β is a measure of altruism. The income i's parents would

[9] Alternatively, he may decide to continue the family business inherited from his parents. Realistically assuming wage flexibility (remember that we are talking of a poor country), however, his marginal product as a self-employed worker will be equated to his wage rate as a dependent worker.

produce, in period 1, if i studied full time is denoted by y_i. We assume that this income varies across families, but is exogenously given for each particular i. Later in the paper, we will see what happens if y_i is a random variable.

We further assume that i's time endowment, normalized to unity in each period, is entirely absorbed by study and work in period 1,[10] and by work in period 2. The amount of time i spends studying, denoted by e^i, includes not only school attendance, but also homework. If i works in the child labour market for L^i units of time, he earns $w_1 L_i$, where w_1 is the child wage rate. The amount of time he works for his parents is then $(1 - e_i - L_i)$.[11] The amount of income she directly or indirectly generates in this way is $z\,(1 - e_i - L_i)$, where $z\,(.)$ is a revenue function, increasing and concave, with $z\,(0) = 0$. Concavity is a reflection of increasing child fatigue and, if the work activity uses a fixed factor as in the case of farming, also of diminishing marginal productivity. In period 1, i receives two transfers, m_i from her parents, and γ_i from the government. Both transfers may be positive, negative or zero.

Let p denote the price of school enrollment, equal to the average total cost of tuition (later in the paper, we will consider the possibility of a price subsidy). If i does not enroll, e_i cannot be anything other than zero. In period 2, i will then earn the unskilled adult wage, w_2. If she enrolls at school, by contrast, e_i may be positive. At date 2, i will then earn $w_2 + x_i - \theta_i$, where x_i is an individual skill premium, and θ_i a transfer (positive, negative or zero) to the government. While w_1 and w_2 are certain, and the same for everybody, x_i is neither of these things. As x_i is revealed only in period 2, while e_i must be chosen in period 1, education is a risky investment. Anticipating a result that will obtain in Section 4, we take γ_i to be independent of L_i, and θ_i to be contingent on x_i. Parents take their decisions in period 1, after the government has announced its policy.

For the moment, we will assume that x_i is i.i.d. over the closed interval $[0, \overline{x}] \in R^+$ with density $f\,(.|e_i)$ conditional on e_i, and $f\,(.|0) = 0$. The i.i.d. assumption, to be relaxed later in the paper, implies that the uncertainty surrounding x_i may arise from imperfect information about i's learning ability, or about the timing of the job offers i will receive after leaving school, but not from aggregate shocks. To simplify the notation, we use x_i to measure the final school result as well as the skill premium. This implies that the only source of uncertainty is actually the child's learning ability. But, using a random variable with density conditional on study time to represent the school result, and another random variable with density conditional on the school result to represent

[10] That is only a convenient simplification. In many developing countries, a substantial minority of school-age children is reported as neither working nor studying. But this can be explained, without putting leisure in the utility function, by allowing for the existence of fixed costs of access to work and school; see Cigno and Rosati (2005).

[11] Especially if the child is female, this unpaid work may consist of replacing her parents (typically her mother) in the performance of domestic chores, and thus allowing them to spend more time generating income. If the child is male, however, this work is more likely to involve direct participation in an income generating activity run by the family, such as farming, retailing, or joining with his parents in seasonal plantation work (which, in some countries, typically means that the whole family has to move temporarily to a place a long way away from home, and consequently also from school).

the skill premium, would make no difference of substance to our results so long both variables are i.i.d., as we assuming for the moment (and, of course, the skill premium is not conditional also on some decision variable). 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, 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 impose the standard convexity-of-distribution-function (CDF), and monotone-likelihood-ratio (MLR), restrictions that $F\left(x^i|e_i\right)$ is convex in e_i, and $\frac{f_{e_i}(x_i|e_i)}{f(x_i|e_i)}$ increasing in x_i.

If i enrolls for school, and L^i is not regulated by the government, i's parents choose $\left(e^i, L^i, m^i\right)$ so as to maximize

$$E\left(V^i\right) \equiv v_i + \beta\left(u_{i1} + \int_{x_i} u_{i2} f^i dx_i\right),$$

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

$$L_i \geq 0. \tag{2}$$

As $E\left(V^i\right)$ is concave in e_i for the MLR and CDF assumptions, the first-order conditions are

$$\text{either } -v_i' z_i' + \beta w_1 u_{i1}' = 0 \text{ or } L_i = 0, \tag{3}$$

$$-v_i' z_i' + \beta \int_{x_i} u_{i2} f_{e_i}^i dx_i = 0 \tag{4}$$

and

$$-v_i' + \beta u_{i1}' = 0. \tag{5}$$

If i does not enroll, i's parents will choose (L_i, m_i) so as to maximize

$$V(L_i, m_i) \equiv v(y_i + z(1 - L_i) - m_i) + \beta\left[u(m_i + w_1 L_i) + u(w_2)\right],$$

subject to (2). The first-order conditions are then just (3) and (5), with $p \equiv e_i \equiv 0$.

3 Laissez faire

In *laissez faire*, school enrollment is not compulsory, L_i is not regulated, and

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

The pay-off of enrolling i at school is then

$$\pi^S(y_i, p) \equiv \max_{(L_i, e_i, m_i)} E\left(V^i\right), \text{ s.t. } (2). \tag{6}$$

That of not enrolling her is

$$\pi^W(y_i) \equiv \max_{(L_i, m_i)} V(L_i, m_i), \text{ s.t. } (2). \tag{7}$$

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 will then be a threshold value of y_i, $\widetilde{y}$, defined by

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

below which i is not enrolled. As $\widetilde{y}$ is the same for every i, because the expected return to education is the same for all of them, if any children do not get an education, 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 (hence, certain), and the threshold is consequently lower for parents of high-ability, than for parents of low-ability children.

If (2) is not binding, and irrespective of whether y_i is or is not high enough for i to be enrolled at school, (3) and (5) imply

$$z_i' = w_1. \tag{8}$$

The amount of work i does for her parents is then independent of y_i. For y_i no lower than $\widetilde{y}$, e_i will then be increasing, and L_i decreasing, in y_i. For y_i lower than $\widetilde{y}$, e_i will be zero, and L_i constant. By contrast, if (2) is binding, L_i is zero. For y_i no lower than $\widetilde{y}$, e_i will then be increasing, and $(1 - e_i)$ consequently decreasing, in y_i. For y_i lower than $\widetilde{y}$, e_i will be constant.

If parents could trade in perfect credit and insurance markets, the *laissez-faire* level of e_i would be independent of y_i. Then, as the expected return to education is the same for every i, either all children would be enrolled at school, or none would. Say that they all would (in other words, that the expected return is sufficiently high to justify borrowing and buying insurance). Having assumed that parents can neither borrow nor buy insurance, however, educational investment without government intervention will be inefficiently low. As pointed out in Lowry (1981), it would be so even if credit were not rationed. For the investment to be efficient, parents must also be able to buy insurance or, failing that, reduce risk by somehow imposing on their children to redistribute income among themselves (and thus, if the redistribution is to be a private family affair, that the parents have many children).

> *Proposition 1. In laissez faire, study time is inefficiently low. In very poor families, children are not even enrolled at school. In less poor ones, study time increases with parental income.*

4 Government

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

$$SW = \sum_{i=1}^{n} E(V_i). \tag{9}$$

Unlike parents, the government is free to borrow or lend in the international credit market. As parents and children are many, and having assumed that risks are uncorrelated, the government does not face any uncertainty about its tax revenue. Assuming that the expected return is sufficiently high for efficiency to require educational investment, the government will then make school enrollment compulsory. Making the usual "small country" assumption, we will treat the real interest rate as a constant, and normalize it to zero. As the optimization can determine only relative tax rates, we will normalize the one implicit in w_2 to zero (and thus avoid having the account for the revenue).

As it does not face any budget uncertainty, the government will then choose $\left(e^i, L^i, m^i, \gamma^i, \theta^i\right)$, for $i = 1, 2, ...n$, so as to maximize (9), subject to the budget constraint,

$$\sum_{i=1}^{n} \left(\gamma_i - \int_{x_i} \theta_i f^i dx_i\right) = 0, \tag{10}$$

and, if $\left(e^i, m^i\right)$ is private information, also to the incentive-compatibility constraints on these variables for each i. As $E\left(V^i\right)$ is concave in $\left(e^i, L^i, m^i\right)$, SW is concave in it too. For the i.i.d. assumption, the socially optimal $\left(\gamma^i, \theta^i\right)$ can depend only on $\left(e^i, m^i, x^i, y^i\right)$, and not also on any $\left(e^j, m^j, x^j, y^j\right)$, $j \neq i$.

4.1 First best

In first best, all parental decisions are common knowledge. In addition to making school enrollment compulsory, the government will then prescribe (e_i, L_i, m_i), and design personalized lump-sum transfers (γ_i, θ_i) for each i, to re-distribute and insure. As the maximization of (9) is subject only to (2) and (10), and denoting by λ the Lagrange-multiplier of the latter, the first-order conditions are (3), (5),

$$-v_i' z_i' + \int_{x_i} (\beta u_{i2} + \lambda\theta_i) f_{e_i}^i dx_i = 0, \tag{11}$$

$$\beta u_{i1}' - \lambda = 0 \tag{12}$$

and, for each possible realization of x_i,

$$-\left(\beta u_{i2}' - \lambda\right) f^i = 0. \tag{13}$$

As in *laissez faire*, either L_i is zero, or $(1 - e_i - L_i)$ is determined by (8). The first-order condition on e_i is not the same as in *laissez faire*, however, because the marginal opportunity-cost to i's parents, $v_i' z_i'$, is equated not to the expected marginal benefit for the parents themselves, $\beta \int_{x_i} u_{i2} f_{e_i}^i dx_i$, but to the sum of this private benefit, and the fiscal externality represented by the expected marginal benefit of a larger tax revenue for society as a whole, $\lambda \int_{x_i} \theta_i f_{e_i}^i dx_i$.

In view of (5), (12) and (13), it is clear that

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

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

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

and larger than in *laissez faire*. But there is nothing to tell us that e^{FB} will be equal to unity. In other words, even an all-seeing and all-powerful government might not want to do away with child work altogether.

> *Proposition 2. In first best,*
>
> (*i*) *all school-age children attend school, and divide their time between study and work in the same way;*
>
> (*ii*) *each school-age child studies more than in laissez faire, but not necessarily full time;*
>
> *(iii) the government uses lump-sum taxes and subsidies to achieve perfect equity, full insurance, and perfect consumption smoothing.*

4.2 Second best

In second best, (e_i, m_i) is private information. The government will then make school enrollment compulsory, fix L_i, and announce how γ_i and θ_i will be related to the information available at the relevant time. As γ_i is payable in period 1, it can depend only on (L_i, y_i). As θ_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 L_i as a legal maximum.

Because of the potential moral hazard problem, the maximization of (9) is subject not only to (2) and (10), but also to the incentive-compatibility constraints (4) and (5). Denoting by ψ_i the Lagrange-multiplier of (2), by ϕ_i that of (4), and by μ_i that of (5), the first-order conditions are now,

$$-\left(z_i' + \phi_i z_i''\right) v_i' + \beta\left(u_{i1}' + \mu_i u_{i1}''\right) w_1 + \psi_i - \left(\phi_i z_i' + \mu_i\right) v_i'' z_i' = 0, \tag{14}$$

$$-\left[v_i' + \left(\mu_i + \phi_i z_i'\right) v_i''\right] z_i' - \phi_i v_i' z_i'' + \int_{x_i} \left[\left(\beta u_{i2} + \lambda\theta_i\right) f_{e_i}^i - \phi_i \beta u_{i2} f_{e_i e_i}^i\right] dx_i = 0, \tag{15}$$

$$-v_i' + \beta u_{i1}' - \mu_i\left(v_i'' - \beta u_{i1}''\right) - \phi_i z_i' v_i'' = 0, \tag{16}$$

$$\beta\left(u_{i1}' - \mu_i u_{i1}''\right) - \lambda = 0 \tag{17}$$

and, for each possible realization of x_i,

$$-\beta\left(f^i - \phi_i f_{e_i}^i\right) u_{i2}' + \lambda f^i = 0. \tag{18}$$

Let us start by characterizing the second-best $(\gamma_i, \theta_i, L_i)$. Using (5), we can re-write (17) 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 this is non-increasing, and given that v'_i is decreasing in y_i, we can then write

$$\gamma_i = \gamma(y_i), \ \gamma' < 0. \tag{19}$$

Condition (18) may be similarly re-written as

$$\beta\left(1 - \phi_i \frac{f_{e_i}}{f^i}\right) = \frac{\lambda}{u'_{i2}}.$$

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

$$\theta_i = \theta(x_i), \ \theta' > 0. \tag{20}$$

As there is nothing to prevent γ_i from falling below zero for y_i sufficiently small, we can think of this period-1 government transfer as the difference between a flat-rate education grant, and a tax increasing in parental income. As it does not depend on L_i, γ_i is neither a tax nor a subsidy on overt child labour. The intuitive explanation is that L_i is fixed by the government. Irrespective of whether L_i is zero or positive, the tax or subsidy would then be a lump-sum transfer. As this would be subtracted from, or added to, γ_i, it would then have no effect. If it is positive, the tax would be a lump-sum payment. By a similar argument, there would be no point in using a price subsidy to push p below the average total cost of tuition. As there is nothing to stop θ_i being negative for x_i sufficiently small, we can similarly think of this period-2 transfer as the difference between a tax increasing in the individual skill premium, and a scholarship increasing in the final school result. Stepping outside the formal model for a moment, we can think of the scholarship as being paid "at the beginning" of period 2 (the end of period 1), when the final school result is revealed,[12], and of the tax as being collected "in the course" of period 2, as the skill premium gradually unfolds. Having established that $\gamma(.)$ is a decreasing function, and $\theta(.)$ an increasing one, it is clear that the policy redistributes from rich to poor parents, and insures them against the risk of a low return to their educational investments, as in first best. Comparing (17) with (12), and (18) with (13), however, it is also clear that the policy does not go as far as in first best. The intuition is, of course, that the government must now limit its recourse to distortionary policy instruments.

Using the incentive-compatibility constraints, (15) and (16), simplify to

$$-(\mu_i + \phi_i z'_i) v''_i z'_i - \phi_i v'_i z''_i + \int_{x_i} \left[\lambda\theta_i f^i_{e_i} - \phi_i \beta u_{i2} f^i_{e_i e_i}\right] dx_i = 0 \tag{21}$$

[12] If partial results are available in the course of period 1, the child could receive partial payments in the course, and the balance at the end, of that period.

and

$$(\mu_i + \phi_i z_i')\, v_i'' - \mu_i \beta u_{i1}'' = 0. \tag{22}$$

Substituting from (22), and using (5), (14) in turn simplifies to

$$\text{either } z_i' = w_1 - \phi_i z_i'' \text{ or } L_i = 0. \tag{23}$$

If w_1 is sufficiently low, L_i is then zero. Otherwise, $(1 - e_i - L_i)$ will be lower than in either *laissez faire* or first best, and decreasing in y_i (because ϕ_i is). Either way,

$$z_i' > w_1. \tag{24}$$

We now go on to characterize the second-best (e_i, m_i). We can see from (21) – (22) that study time and parental transfers are not equalized across families. Having found that period-1 redistribution does not go as far as in first best, that is hardly surprising. This is as far as the government's first-order conditions will carry us, but we can go a little further by a different route.

Given the policy, (e_i, m_i) is determined by (4) – (5). Taking the total differential, and solving by Cramer, we then find

$$\frac{\partial e_i}{\partial \gamma_i} = \frac{v_i'' z_i' \beta u_{i1}''}{H_i} > 0, \tag{25}$$

$$\begin{aligned}
\frac{\partial e_i}{\partial L_i} &= -\frac{v_i'' z_i' \beta u_{i1}'' \left(z_i' - w_1\right) + v_i' z_i'' \left(v_i'' + \beta u_{i1}''\right)}{H_i} < 0, \\
\frac{\partial m_i}{\partial \gamma_i} &= -\frac{\beta u_{i1}''}{H_i}\left[v_i' z_i'' + v_i'' \left(z_i'\right)^2 + \frac{\partial}{\partial e_i} E\left(MU_{e_i}\right)\right] < 0
\end{aligned} \tag{26}$$

and

$$\frac{\partial m_i}{\partial L_i} = -\frac{\beta u_{i1}'' \left[v_i' z_i'' + v_i'' \left(z_i'\right)^2\right]}{H_i} - \frac{\left(v_i'' z_i' + \beta w_1 u_{i1}''\right)}{H_i} \frac{\partial}{\partial e_i} E\left(MU_{e_i}\right) < 0$$

where H_i is the Hessian determinant (positive at a maximum), and

$$E\left(MU_{e_i}\right) \equiv \int_{x_i} u_{i2} f_{e_i}^i \, dx_i$$

the expected marginal utility of e_i.

It is clear from (25) – (26) that i's parents regard γ_i as a transfer to themselves, and will thus raise or lower e_i according to whether γ_i is positive or negative (relaxes or tightens the credit constraint). Let us see if that helps us to deduce what the policy does for educational investment. On the one hand, as $\gamma(.)$ is decreasing in y_i, the policy then encourages educational investment more in poor than in rich families; in very rich ones, it may actually discourage investment. On the other hand, as $\theta(.)$ is decreasing in x_i, the policy makes education less risky for all parents. All families, except perhaps the very rich ones, will then invest more in their children's education than they would in *laissez faire*, and the increase will be larger in poor than in rich families. In this respect too, disparities will then be smaller than in *laissez faire*, but we cannot be sure that they will disappear altogether. Unless it is zero for all families, L_i is decreasing in y_i.

> *Proposition 3. In second best,*
>
> *(i) all school-age children enroll for school, and all of them, with the possible exception of those with very rich parents, study more than they would in laissez fare;*
>
> *(ii) either no school-age child does any paid work, in which case children with poor parents do more unpaid work than children with rich ones, or children with poor parents do more unpaid work than children with poor ones;*
>
> *(iii) the government uses a net subsidy decreasing in parental income, and a net tax increasing in the skill premium, to redistribute and insure, and will thereby relax the credit constraints of all but the richest parents, but will not achieve perfect equity, full insurance, and perfect consumption smoothing.*

In most countries, primary school enrollment is compulsory, and (at least officially) children up to a certain age are not allowed to work. In many developing countries, however, $\gamma(.)$ is not available. If education is subsidized at all in period 1, it is through p. Is that better or worse than *laissez faire*? We will answer this question in two steps. First, starting from *laissez faire*, would the imposition of school enrollment raise social welfare? The answer is no. The policy would lower social welfare, because it would oblige all parents, including those who will not send their children to school anyway, to pay p. Forbidding child labour would reduce welfare even further, because the ban would apply only to overt child labour, and thus distort time allocation. Second, given compulsory education, and with or without a ban on child labour, would a price subsidy raise welfare? The answer depends on how the subsidy is financed. If it is financed by a poll tax, the policy will have no effect, because parents will effectively take a lump-sum subsidy with one hand, and pay a lump-sum tax of the same amount with the other. If the price subsidy is financed by a tax on parental income, by contrast, we get something that looks almost like our $\gamma(.)$. Not quite like it, because the subsidy cannot be larger than the average total cost of tuition, and that may not be enough for a second best.

5 Extensions and some evidence

So far, we have assumed that the returns to educational investments are i.i.d., thereby implying that the uncertainty surrounding x_i arises entirely from imperfect information about i's learning ability, or the timing of the job offers that i will receive after leaving school. What difference would it make if information about the state of the labour market, or of the economy generally, were imperfect too, and the returns to educational investments were thus subject to aggregate, as well as idiosyncratic shocks? As aggregate risks cannot be insured by redistributing among parents, the government should use its ability to borrow and lend, not only to relax the credit constraints faced by parents, but also to stabilize the expected return to education across generations (i.e., replace the constraint that taxes and subsidies must be matched for each generation, with the one that they must be matched over an infinite time horizon).

We have also assumed non-increasing absolute risk aversion. That is consistent with evidence of diminishing absolute risk aversion in an educational context is reported in Kodde (1986). Our prediction that, in the absence of perfect credit and insurance market, and short of government intervention, educational investment increases with parental income is consistent with the empirical findings of Jacoby (1994). Our result, that the second-best policy raises the aggregate demand for education, not only because it relaxes the credit constraint on educational investment for the average family, but also because it makes the return to this investment less uncertain, is consistent with evidence in Johnson (1987).

The empirical literature has paid a certain amount of attention also to the role of children as a form of insurance in the face of parental income uncertainty (as against uncertainty about the return to education). Beegle et al. (2007) report that, in the absence of formal credit and insurance facilities, parents respond to a downturn in their own income by making their children work more. Fitzsimons (2007), by contrast, finds that parents respond in this way to a downturn in village aggregate income, but not to a downturn in their own income. Our model can be extended to accommodate this type of uncertainty by treating parental income as a random variable with known density. As the density in question is not conditional on any action undertaken by the parents (or, rather, assuming this to be the case, as in the reference literature), however, the implications can be intuited without formal analysis.

If income shocks are purely idiosyncratic, they can be neutralized by mutual insurance arrangements at the local (say, village) level. The results are then unchanged. If income shocks have also an aggregate, village-level component, parents will respond in the way estimated by Fitzsimons. The policy prescription will then change, but only in a matter of interpretation, namely that $\gamma(.)$ now has the additional role of insuring parents against village-level risks. Something similar may be said if the aggregate shocks are country-wide, rather than village-level. The government must then borrow or lend, not only to relax individual credit constraints, and redistribute within cohorts, but also (as in the case where the shocks affect the return to educational investment) to redistribute between cohorts.

6 Conclusion

In *laissez faire*, if parents can neither borrow nor insure, educational investment is inefficiently low. If parental income falls below a certain threshold, the child will not even enroll at school. Above that threshold, study time will increase with parental income. That is consistent with the available evidence. As already established in the theoretical literature, and confirmed by the empirical one, parents will underinvest in their children's education not only if they cannot borrow against the expected return, but also if they cannot insure against the risk of a low return.

The government can relax the credit constraint by effectively giving parents an advance on the expected return of their educational investments, and make this return less uncertain by redistributing from rich (lucky) to poor (unlucky) children. In the light of evidence of diminishing absolute risk aversion in the present context, reducing uncertainty would reduce also inequality among parents. This inequality can be further reduced by redistributing from rich to poor parents. If we extend the model to allow for parental income to be uncertain, redistributing from rich to poor parents also would serve as insurance.

In first best (i.e., if the government costlessly observes all parental decisions), all parents will enroll their children at school, and invest in their education at the efficient level. As children are ex-ante identical, they all study for the same amount of time. But there is nothing to indicate that they will study full time. In a poor country, it may well be efficient for school-age children to do a certain amount of non-harmful work. The policy uses personalized lump-sum transfers to achieve perfect equity, full insurance, and perfect consumption-smoothing.

In second best (defined here to mean that the government observes only the children's enrollment status, and the amount of work they do overtly in the market), school enrollment is compulsory, and there is a legal ceiling, decreasing in parental income, on the amount of overt work a school-age child is allowed to do. The government again redistributes across parents, and across children, but does not achieve perfect equity, full insurance, and perfect consumption-smoothing, because it uses distortionary policy instruments. These include a flat-rate education grant, and a scholarship increasing in the final school result. Overt child labour is not taxed, and the price of school enrollment is not subsidized. But school enrollment is compulsory, and there is a legal ceiling, non decreasing in parental income, on the overt child labour. Most children study more than in *laissez faire*.

Although it abstracts from the worst forms child labour,[13] our analysis refers quite clearly to a poor developing country. It may thus be interesting to compare our results with those of an analysis which relates quite clearly to a rich developed one. Hanushek *et al.* (2003) use a calibrated general equilibrium model to assess the welfare effects of a range of policy instruments under the assumption that child labour is out of the question, risk considerations can be ignored, and there is no moral hazard with regard to the amount of time a child spends studying. The result is that education subsidies generally perform less well than other forms of redistribution, and that a merit-based education like ours is justified only in the presence of an education externality. These differences highlight the importance of the stage of development.

[13] For an analysis encompassing these forms of child labour, see Dessy and Pallage (2005).

It may also be interesting to compare our policy results with what is done in practice. Primary school enrollment is compulsory, and labour at a very young age forbidden (officially at least), in most countries. In poor developing ones, however, education is subsidized, if at all, only through the price of school enrollment. We have argued that making enrollment compulsory, without offering any kind of subsidy, in a situation where parents can neither borrow, nor buy insurance, would push social welfare below the *laissez-faire* level, and that forbidding child labour on top of it would make things even worse. Starting from a situation where school enrollment is compulsory, however, the introduction of a price subsidy would have no effect if this were financed by a poll-tax, but would raise educational investment and social welfare if it were financed by a tax increasing in parental income. Either, the effect would not be as large as that of our policy, because the subsidy could not exceed the cost of tuition, and that might not be enough for a second best.

References

Baland, J. M. and J. A. Robinson (2000), Is Child Labor Inefficient? *Journal of Political Economy* 108, 663-679

Cigno, A. (2006), A Constitutional Theory of the Family, *Journal of Population Economics* 19, 259-283

——— (2009), Agency in Family Policy: A Survey, CESifo Working Paper 2664

——— and F. C. Rosati (2005), *The Economics of Child Labour*, New York and Oxford: Oxford University Press

Beegle, K., Dehejia, R. H., Gatti, R. (2007), Child Labour, Income Shocks, and Access to Credit, *Journal of Development Economics*, forth.

Dessy, S. E. and S. Pallage (2005), A Theory of the Worst Forms of Child Labor, *Journal of Development Economics* 115, 68-87

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

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

Johnson, W. R. (1987), Income Redistribution as Human Capital Insurance, *Journal of Human Resources* 22, 269-280

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

Pouliot, W. (2006), Introducing Uncertainty in Baland and Robinson's Model of Child Labour, *Journal of Development Economics* 79, 264-272

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