Child labour and education policy in the presence of imperfect and asymmetrical information.*

Alessandro Cigno
University of Florence, CESifo, CHILD and IZA

August 26, 2009

Abstract

We use a principal-agent model to show that the government should make school enrollment compulsory for all children. Covert child labour should be dealt with by giving school children a subsidy decreasing in parental income, and increasing in school results. The cost of the policy should be recovered by taxing the school leavers’ enhanced earnings.

Key words: overt and covert child labour, education, learning ability, education subsidies, income tax, principal-agent.

JEL: D82, H21, H31, I28.

1 Introduction

Child labour can be approached either as a human rights issue, or as a forgone investment opportunity. The first approach starts from the premise that a child has a fundamental right to be sheltered from premature work exposure, and carries the implication that the rest of humanity has a moral duty to make this right effective by providing the necessary resources. The second approach starts from the premise that education is a form of investment, because it raises (among other things) one’s future earning capacity, and prompts the following question. If a child (or her parents on her behalf) could trade in full contingent markets, and the decision to invest in an education were thus independent of current family income and tradeable assets, would the investment take place? If the answer is negative, there is no policy issue. If it is positive, but the investment is not carried out because the family is not rich enough to finance it with its own resources, and cannot borrow against the child’s expected future earnings, there is a case for public intervention on both equity and efficiency grounds.¹

*Paper to the UCW-University of Galatasaray Seminar on Child Labour, Education and Youth Employment, Istanbul, 8-9 October 2009

¹The possible inefficiency of child labour under conditions of certainty is established in Baland and Robinson (2000). Pouliot (2006) extends the analysis to the case of uncertainty.
Provided that the government, unlike ordinary citizens, can borrow against the additional tax revenue that an education policy is expected to generate in the future, there will be no need for this policy to be subsidized. There can be an argument for it to be subsidized out of general government revenue only if there is a positive education externality. That is the stance we take in the present paper.

As the consequences of an education externality are well understood, we will assume that there is no such externality, and focus on the policy implications of credit rationing under uncertainty and asymmetric information. Uncertainty arises from the fact that learning ability varies across children, and is only imperfectly observable before the education process begins. Assuming, however, that parents are no better informed than the government about their own children’s innate ability to learn, there will be no adverse selection. Furthermore, as the expected benefit is the same for all children, the educational investment will be either socially optimal for all children, or optimal for none of them. We assume the former. The informational asymmetry arises from the fact that child labour may be either overt or covert. The former takes place mostly in factories or plantations, the latter mostly in the child’s own home, family farm or family business. Being observable, overt child labour can be outlawed, and effectively extirpated if the government is so minded. Covert child labour, by contrast, is unobservable, and cannot be extirpated by command. Something similar may be said about education. School enrollment is necessary, but not sufficient for a child to receive an education. For the latter to be effective, the child must actually attend school, be alert during lessons, and do her homework. If a child is enrolled for school, but works for her parents at the same time, she will tend to miss classes on grounds of actual or alleged ill health more often, and be more prone to fall asleep during lessons when she does attend, than children who do not work. She will also do little or no homework. While school enrollment is observable and thus enforceable, regular school attendance and homework are difficult to monitor and thus to enforce. There may then be a moral hazard problem.

The policy optimization has a principal-agent structure. In the present context, the principal is the government. Assuming that school-age children are under parental control, the agents are the parents. In the logic of agency problems, if an action is either unobservable or very costly to observe, the agent must be given the incentive to undertake it at the level desired by the principal. By contrast, if an action is observable, it does not make sense for the principal to offer costly incentives, because the same result can be costlessly achieved using a "forcing contract" (i.e., threatening the agent with a sufficiently severe

---

2 The connection between child labour and credit rationing are examined theoretically and empirically in Ranjan (2001).

3 Cooking ad cleaning, looking after younger siblings, fetching fuel and water, etc. does not generate income directly, but allows adult family members to do so. See Cigno and Rosati (2005), and references therein.

4 The peculiarities of the government-parent-child relationship raise issues and yield insights not present in the usual IO or optimal taxation applications; see Cigno (2009).
penalty if he does not comply). In our context, the actions falling into the first category are parental transfers to school-age children, covert child labour and the amount of time a child spends studying. Those falling into the second are overt child labour, and school enrollment. The model has the following time structure. At date 0, the government announces its policy. At date 1, children reach school age, and their parents make decisions. Education results are known at date 2. At date 1, parents cannot borrow against their children’s expected date-2 earnings. The real interest rate is normalized to zero.

2 Households

There are \( n \) households indexed \( i = 1, 2 \ldots n \), where \( n \) is a large number. Each household consists of a couple of parents, and one child. For briefness, we will refer to the child in the \( i \)th household as \( i \), and to the couple as \( i \)'s parents. Ex post, \( i \)'s utility will be

\[
U^i = u(c^i_1) + u(c^i_2) ,
\]

where \( c^i_d \) denotes the child's consumption at date \( d = 1, 2 \). Assuming descending altruism, that of \( i \)'s parents will be

\[
V^i = v(C^i) + \rho U^i , \quad 0 < \rho < 1 ,
\]

where \( C^i \) denotes parental consumption. The functions \( u(\cdot) \) and \( v(\cdot) \) are increasing and concave, with \( u'(0) = v'(0) = \infty \). As idleness does not figure in either of these functions, \( i \)'s date-1 time endowment (normalized to unity) will be totally absorbed by study and work activities.\(^5\) The date-2 endowment (also normalized to unity) will be entirely spent working.

At date 1, \( i \) receives a transfer \( m^i \) (positive, negative or zero) from his parents. If \( i \) engages in overt child labour, he will be remunerated at the child wage rate \( w_c \) at date 1, and at the unskilled adult wage rate \( w_u \) at date 2. Both \( w_c \) and \( w_u \) are defined net of income tax. If \( i \) enrolls for school, she will receive a transfer \( s^i \) from the government at date 1. We may think of \( s^i \) as the difference (positive, negative or zero) between a nonnegative student grant \( g^i \),\(^6\) and a positive tuition fee, equal to average tuition cost, \( p \),

\[
s^i \equiv g^i - p, \quad g^i \geq 0, \quad p > 0.
\]

If \( i \) works for her parents, she will get no wage, but her family income will increase by the amount \( z(1 - e^i) \), where \( e^i \) is her study time, varying in the closed interval \( E = [0, 1] \), and \( z(\cdot) \) is a revenue function, increasing and concave,

---

\(^5\)That is only a convenient simplification. In reality, up to quarter of school-age children in many developing countries is reported neither working nor studying. But this can be explained without putting idleness in the utility function by allowing for fixed costs of access to work and school; see Cigno and Rosati (2005).

\(^6\)Having assumed that parents are credit rationed, we must assume that, if the grant is conditional on the final school result, known only at date 2, \( i \) will receive partial advances on \( s^i \) between dates 1 and 2, as partial results come become available, and the final balance at date 2.
with \( z(0) = 0 \) and \( z'(0) = \infty \). At date 2, she will earn the skilled wage \( w_u + x^i \), where \( x^i \) is her individual skill premium, and make an additional income-tax payment \( t^i \). As the policy optimization will determine only the difference between the income-tax rate applicable to \( w_u \), and that applicable to \( x^i \), we will normalize the former to zero.

While \( w_u \) and \( w_u \) are known at date 1, and the same for every \( i \), the individual skill premium, \( x^i \), reflects \( i \)'s final school result.\(^7\) The latter reflects in turn not only the amount of time that \( i \) studied, \( e^i \), but also her learning ability. This trait is not directly observable. Given \( e^i \), it can be inferred from \( x^i \). As \( x^i \) will be revealed only at date 2, however, \( e^i \) is a risky investment. We assume that \( x^i \) is i.i.d. over the closed interval \( X = [0, \pi] \in R^+ \), with density \( f(x^i|.) \) conditional on \( e^i \), and that the cumulative distribution of \( x^i \), \( F(x^i|.) \), associated with a higher \( e^i \) first-order stochastically dominates (FOSD) the one associated with a lower \( e^i \),

\[
F_{e^i}(x^i|e^i) \leq 0. \tag{2}
\]

For each value of \( e^i \), there will be values of \( x^i \) such that this condition holds as an inequality. We further assume that \( (f_u/f) \) is increasing in \( x^i \), and that \( F(x^i|e^i) \) is convex in \( e^i \).\(^8\) Assuming that \( f(x^i|.) \) is common knowledge, there is no adverse-selection problem. If \( e^i \) is private information, however, there will be a moral-hazard one.

If \( i \) does not enroll for school, his future is certain. Assuming \( z(1 - e^i) \leq (1 - e^i) w_c \) for any \( e^i \), \( i \) will supply all his date-1 time endowment to the child labour market. His parents will then choose \( m^i \) to maximize

\[
V^i = v(Y^i + w_c - m^i) + \rho \left[ u(m^i) + u(w_u) \right],
\]

where \( Y^i \), varying in the closed interval \( K = [0, \bar{Y}] \in R^+ \), is their own income.

The solution will satisfy the first-order condition

\[
-v'(Y^i + w_c - m^i) + \rho u'(w_c + m^i) = 0. \tag{3}
\]

If \( i \) enrolls for school, by contrast, her future is uncertain. Allowing for the possibility that \((s^i, t^i)\) depends on \( x^i \), her lifetime utility will be a random variable with expected value

\[
E(U^i) = \int_{x^i} \left[ u(m^i + s^i) + u(w_u + x^i - t^i) \right] f dx^i. \tag{4}
\]

Her parents’ will then choose \((e^i, m^i)\) conditionally on \((s^i, t^i)\) to maximize

\[
E(V^i) = \int_{x^i} v(Y^i + z(1 - e^i) - m^i) f dx^i + \rho E(U^i). \tag{5}
\]

\(^7\)We may think of the latter as of an actual score, or as of the number of school years successfully completed by \( i \).

\(^8\)By ensuring that \( i \)'s, and his parents’, expected utility is concave in \( e^i \), these standard (monotone-likelihood-ratio and convexity-of-distribution-function) conditions will allow us to substitute \( i \)'s first-order condition for the incentive-compatibility constraint on the choice of \( e^i \) in the policy optimization problem.
Using standard procedures, it can be shown that \( E(V) \) is concave in \( e^i \). The solution will satisfy the first-order conditions

\[
z'(1 - e^i) = \int x^i \left[ \frac{1}{\rho} v(Y^i + z(1 - e^i) - m^i) + u(m^i + s^i) + u(w_u + x^i - t^i) \right] f_{x^i} dx^i
\]

and

\[
-v'(Y^i + z(1 - e^i) - m^i) + \rho u'(m^i + s^i) = 0
\]

for each possible realization of \( x^i \).

Let \( y^i \) denote \( i \)'s date-1 income before the parental transfer, so that \( y^i = w_c \) if \( i \) does not enroll, and \( y^i = s^i \) if she does. Using either (3) or (7), we may then express the parental transfer as a function of the parents’ and the child’s own date-1 incomes,

\[
m^i = m(Y^i, y^i),
\]

with

\[
\frac{\partial m_1}{\partial k} = \frac{v''}{v'' + \rho u''} \quad \text{and} \quad \frac{\partial m_1}{\partial I} = -\frac{\rho u''}{V'' + \rho u''}
\]

\[
m_y = \frac{v''(C^i)}{v''(C^i) + \rho u''(c^i)} > 0
\]

and

\[
m_y = -\frac{\rho u''(c^i)}{v''(C^i) + \rho u''(c^i)} < 0.
\]

Assuming \( u'' = v'' = 0 \), both \( m_y \) and \( m_y \) are constant. For \( Y^i \) sufficiently small relative to \( y^i \), \( m(Y^i, y^i) \) will be negative. In other words, the parents will receive money from their child.

### 3 Laissez faire

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

\[
s^i = -p, \quad t^i = 0.
\]

Let the \( LF \) subscript denote the *laissez-faire* value of a variable. As \((s^i, t^i)\) does not depend on \( x^i \) in this case, the pay-off of enrolling \( i \) for school is

\[
\pi^S(Y^i, p) = \max_{(e^i, m^i)} v(Y^i + z(1 - e^i) - m^i) + \rho[u_1(m^i - p) + \int x^i u(w_u + x^i) fdx^i].
\]

That of not enrolling him is

\[
\pi^W(Y^i) = \max_{m^i} v(Y^i + w_c - m^i) + \rho[u_1(m^i - p) + u(w_u)].
\]
The child will be enrolled for school if and only if \( \pi^S(Y^i, p) \) is at least as large as \( \pi^W(Y^i) \). There will be a threshold value of \( Y^i, \tilde{Y}_{LF} \), defined by

\[
\pi^S(\tilde{Y}_{LF}, p) = \pi^W(\tilde{Y}_{LF}),
\]

below which \( i \) is not enrolled for school. As \( \tilde{Y}_{LF} \) is the same for all \( i \), the children excluded from education will be those with the poorest parents, not those with the lowest learning ability.

Proposition 0. In laissez faire, children from poor families may be excluded from school even if their learning ability is high.

4 Education policy

Let the government’s preferences be represented by the Benthamite social welfare function

\[
SW = \sum_{i=1}^{n} \int_{x^i} v(Y^i + z(1 - e^i) - m^i) f^i dx^i + \rho \sum_{i} \int_{x^i} \left[ u(m^i + s^i) + u(w_i + x^i - t^i) \right] f^i dx^i,
\]

where \( f^i = f(x^i|e^i) \). As the expected utility of \( E(V^i) \) is concave in \( e^i \), \( SW \) is concave in \( e^i \) too. Thanks to the i.i.d assumption, the socially optimal \( (s^i, t^i) \) can depend only on \( x^i \) (not on any \( x^j \neq x^i \)). As \( n \) is large, the government faces no uncertainty over the total it will have to pay out in educational subsidies at date 1, and the total it will get back in taxes at date 2. We can then write the government’s education budget constraint as

\[
\sum_{i=1}^{n} \int_{x^i} (s^i + p - t^i) f dx^i = 0.
\]

As the expected gain from education is the same for every \( i \), and assuming

\[
E(x^i) \geq p,
\]

the government will (use a forcing contract to) make school enrollment compulsory for all children, and then choose \( (e^i, m^i) \) for each \( i \) so as to maximize (11), subject to (13) and, if \( (e^i, m^i) \) is private information, also to the incentive-compatibility constraints (6) and (8).

4.1 First best

In first best, the government observes \( (m^i, e^i) \). As (13) is the only constraint, the allocation will then satisfy the first-order conditions
\[
[u'(m^i + s^i) - \lambda] f^i = 0
\]
and
\[
[-u'(w_u + x^i - t^i) + \lambda] f^i = 0
\]
for each possible realization of \(x^i\), and
\[
\int_{x^i} [v(Y^i + z(1 - e^i) - m^i) + \rho(u(m^i + s^i) + u(w_u + x^i - t^i))] f_{e_i}^i dx^i + \lambda \int_{x^i} (t^i - s^i - p) f_{e_i}^i dx^i = 0,
\]
where \(\lambda\) is the Lagrange-multiplier of the education budget.

Let the \(FB\) subscript denote the first-best value of a variable. In the Appendix, we demonstrate the following.

**Proposition 1.** In first best,

(i) all school-age children enroll for school, and study for the same amount of time, \(e_{FB} < 1\);

(ii) child consumption is equalized across children, dates, and states of the world,
\[
c_1^i = c_2^i = c_{FB};
\]

(iii) parental consumption is equalized across parents and states of the world,
\[
C^i = C_{FB}.
\]

Therefore, the first-best policy uses personalized lump-sum subsidies and taxes to achieve consumption smoothing, perfect equity, and full insurance against the risk that a child will turn out to have low learning ability. In standard principal-agent models, the full-insurance property comes from the assumption that the principal is less risk-averse than the agents. Here, by contrast, it is due to the fact that the principal (the government) does not face budget uncertainty. As a child’s learning ability is not known in advance, all children are required to supply the same amount of study time. Even in first best, children will do at least a small amount of unpaid work for their parents.\(^9\)

4.2 Second best

In second best, the government does not observe \((m^i, e^i)\). As the maximization of (11) is then subject also to (6) and (8), the first-order conditions are
\[
[u'_i(m^i + s^i) - \lambda] f + \mu_i u'(m^i + y^i) f_{e_i} = 0,
\]
\(^9\)This result derives from our assumption that \(z'(0) = \infty\). Were we to drop this assumption, we could have a corner solution with \(e_{FB} = 0\). But, it hardly makes a difference whether the children work a little, or not at all.
\[
\left[-u'(w_u + x^i - t^i) + \lambda\right] f - \mu^i u'(w_u + x^i - t^i) f_{e^i} = 0 \quad (18)
\]

and
\[
\int_{x^i} \left[v(Y^i + z(1 - e^i) - m^i) + \rho (u(m^i + y^i) + y^i) + u (w_u + x^i - t^i))\right] f_{e^i} dx^i + \lambda \int_{x^i} (t^i - s^i - p) f_{e^i} dx^i
\]
\[
+ \mu^i \left\{ \int_{x^i} \left[v(Y^i + z(1 - e^i) - m^i) + \rho (u_1(m^i + y^i) + u (w_u + x^i - t^i)) - z_{1e^i}^i\right] f_{e^i} dx^i \right\} = 0,
\]
where \(\lambda\) is the Lagrange-multiplier of (13), \(\mu^i\) that of (6), and \(m^i = m(Y^i, s^i)\).

Let the SB subscript denote the second-best value of a variable. In the Appendix, we demonstrate what follows.

**Proposition 2.** In second best,

(i) all school-age children enroll for school, and study for the same amount of time, \(e_{SB} < 1\);

(ii) at date 1, \(i\) receives a public subsidy \(s_{SB}^i = \varphi_1(Y^i) + \varphi_2(x^i)\), where \(\varphi_1(.)\) is a decreasing function, and \(\varphi_2(.)\) an increasing one;

(iii) at date 2, \(i\) pays a tax \(t_{SB}^i = \tau(x^i)\), where the function \(\tau(.)\) is increasing up to a certain point, but may be decreasing beyond that;

(iv) at each date, expected child consumption is equalized across children;

(v) expected parental consumption is equalized across parents.

Therefore, the second-best policy applies the same subsidy and tax schedules to all households. Both are the result of a compromise between either equity or insurance and incentive considerations. The subsidy has thus two components, one decreasing in parental income (increasing in "need"), and the other increasing in scholastic performance ("merit"). The tax, by contrast, is increasing in the skill premium (or, equivalently, scholastic performance) at low realizations of this random variable where the marginal utility of consumption is very high, and insurance considerations prevail, but possibly decreasing at high ones, where incentive considerations may be paramount. As a consequence, the second-best policy cannot provide full insurance, and can only equalize expected consumption.

### 5 Conclusion

If the expected benefit of education, the same for all children, is at least as large as the cost, the government should make school enrollment compulsory. If it could observe the amount of time a child spends studying, and the amount of money she receives from her parents, it should prescribe those too, and then use personalized lump-sum subsidies and taxes to achieve perfect equity and consumption smoothing, and provide households with perfect insurance against the
risk that a child will turn out to be of low learning ability. In the more realistic hypothesis that study time and parental transfers are private information, equity and insurance would have to be traded against incentive considerations. The government should then subsidize children while at school in accordance with a common schedule, decreasing in parental income (increasing with "need"), and increasing in school results ("merit"). When the children leave school, the government should tax their earnings in accordance with another common schedule, this time increasing in their earnings, at least up to a point, and then possibly decreasing. Because of the need to provide incentives, this policy would achieve neither perfect equity nor perfect insurance.

References

Cigno, A. (2009), Agency in Family Policy, CESifo WP 2664

Appendix

A1. Proof of Proposition 1

Part (i). In view of (2), the cumulative distribution of $x^i$ for any $e^i > 0$ will FOSD the cumulative distribution of $x^i$ for $e^i = 0$. As $e^i$ can be positive only if $i$ is enrolled for school, and recalling that the government faces no budget risk, it then follows that social welfare cannot be at a maximum if any child is not enrolled. As the cumulative distribution function is the same for every $i$, all children have the same expected skill premium. Therefore, they will be required to study for the same amount of time, $e_{FB}$. As $z^i (1 - e^i) = \infty$, $e_{FB} < 1$.

Parts (ii) and (iii). As $\lambda$ is the same for every $i$, (14) and (15) imply

$$ m^i + s^i = w_u + x^i - t^i = c_{FB}. $$

A2. Proof of Proposition 2

Part (i). The argument is the same as for part (i) of Proposition 1.

Parts (ii) and (iii). Condition (17) may be re-written as

$$ \frac{\lambda}{u'_1 (m(Y^i, s^i) + s^i)} = 1 + \mu^i \frac{f_{e^i}(x^i | e^i)}{f(x^i | e^i)}. \quad (20) $$

As $\mu^i$ is positive, $s^i$ depends not only on $Y^i$, but also on $x^i$. The LHS is increasing in $m(Y^i, s^i) + s^i$ because $u'_1(\cdot)$ is a decreasing function. The RHS is increasing in $x^i$ for the MLR assumption. As $m_{Y^i} > 0$ and $-1 < m_{x^i} < 0$,
the second-best $s^i$ is then decreasing in $Y^i$, and increasing in $x^i$. As $Y^i$ does not figure on the RHS of (20), and given that $m_{Y^i}$ and $m_{x^i}$ are constant, the second-best subsidy schedule is additively-separable in $Y^i$ and $x^i$. Condition (18) may be similarly re-written as

$$\frac{\lambda}{u'_2(w_u + x^i - t^i)} = 1 + \mu_i \frac{f_{x^i}(x^i|e^i)}{f(x^i|e^i)}. \quad (21)$$

As the LHS of (21) is increasing in $x^i$ and decreasing $t^i$, and the RHS increasing in $x^i$, the effect of $x^i$ on $t^i$ has ambiguous sign. In view of the fact that $u'_2(w_u + x^i - t^i)$ is decreasing in $x^i$, the effect will be positive and large at low realizations of $x^i$, where the marginal utility of date-2 consumption is large, but will small or negative at high ones, where this marginal utility is small.

**Parts (iv) and (v).** As the RHS of (20) and (21) are increasing in $x^i$, and the same for every $i$, the LHS also will be increasing in $x^i$, and the same for every $i$. 

10