Collegio Carlo Alberto

Child labor, school attendance and access to health care services by children: evidence from Ghana

Vincenzo Atella
Mariacristina Rossi

No. 178
December 2010

Carlo Alberto Notebooks

www.carloalberto.org/working_papers

Child labor, school attendance and access to health care services by children: evidence from Ghana[1]

Vincenzo Atella
CEIS – Department of Economics
University of Rome Tor Vergata
CHILD

Mariacristina Rossi
Department of Economics
University of Turin, Italy
CeRP Collegio Carlo Alberto

December 2010

The paper develops a simple two-period model relating child labor, child school attendance and child health care access in LDCs showing that child labor is positively correlated to access to health care services. In fact, higher medical expenditure generates better health and, therefore, higher child productivity. Accumulation of human capital, which generates higher future utility, comes at the expense of current productivity and consumption. The optimal choice of child labor is such that the marginal benefit of schooling is equal to the marginal productivity of child labor, which is enhanced by additional medical expenditure. Under this theoretical set-up we expect medical expenditure and child labor to be positively correlated, with parents caring more for their children if they contribute to household income. We explore these relationships using a micro data set from Ghana LSS for the year 1999. Empirical results confirm the model theoretical predictions.

Keywords: Child labor, Health care demand, Human capital, Latent variables, multivariate Probit, Unobserved heterogeneity, LDCs.
JEL code: I12, J13, J22, J28, J43

1. - Introduction

The last decades have witnessed a large proliferation of research on child labor in Less Developed Countries (LDCs). In particular, the worst forms of child labor, such as hazardous work in mining, have received much of the policy attention, mainly because they are harmful and impede children's future development, both physical and intellectual. However, the consequences of non-hazardous child work on both health and education are less clear, despite the fact that it represents the largest fraction of total child work. In fact,

[1] The authors wish to thank Alberto Holly for kindly providing the GAUSS routine to estimate the simultaneous trivariate Probit model. They also acknowledge exceptional computing support from Jacques Huguenin who helped produce the econometric results. We gratefully acknowledge the financial help from Understanding Children's Work inter-agency project. The usual disclaimers apply. Corresponding author: Vincenzo Atella, Department of Economics, University of Rome Tor Vergata, Via Columbia n.2, 00133 Roma Tel. +39 6 72595635, Fax +39 6 2020687, e-mail: atella@uniroma2.it

non-hazardous work for children is quite often a low-intensity activity and mostly concerns rural activities.[2] According to the International Labour Organization (ILO, 2002), 70% of working children worldwide are involved in rural activities. Child labor in Ghana is in line with the worldwide evidence, with approximately 80% of working children working in 2003 in the rural sector (UCW, statistics portal). Moreover, most working children engage in unpaid family work, by contributing to the production function of the household.

Within child labor research, analysis has mostly focused on the determinants and consequences of child labor, with special emphasis on (i) the relationship between child work decision and school attendance (or school outcomes) and (ii) the relationship between child work and child health status. Although there exists a large body of literature on the relationship between child labor and schooling, the empirical analyses have usually been limited to exploring correlations, rather than causal relationships (one exception is Beegle et al. 2004). Concerning the impact of child labor on health outcomes, the literature is less developed, and causality relationships have been explored even less. To our knowledge, the few exceptions are represented by O'Donnel, Rosati and van Doorslaer (2005) (henceforth ORD) and Beegle et al. (2004), both using a panel data of Vietnamese households. Beegle et al. (2004) do not find negative consequences of child labor on health for those children who are in school, while ORD find a small negative effect of child labor on health outcomes five years later.

On an initial look at the correlation between child labor intensity and health indicators, it appears that working children are associated with a better health status than non-working children (UCW, statistics portal). Positive correlation between working and health status, however, does not necessarily mean that children who work are better fed and treated with better medication as it could be that working children have a different perception of their health status than non-working children. In fact, the health indicators used are usually self-assessed and thus they can suffer from serious measurement problems.

With this paper we are being innovative with respect to the existing literature by shedding light on how child labor and schooling participation jointly affect household decisions with regard to health expenditure for children. In our view this is the only appropriate way to study such choices. Modeling separately the relationship between child work and school attendance or health status may produce biased estimates, given that an important part of the behavioral model is left out.

We then derive a theoretical implication to test the relation between child labor decisions, school attendance and access to health care services, which is considered to be a proxy for the (unobserved) health status. Our model departs from the theoretical setting developed in Rosati and Rossi (2003), where only child labor and school outcomes are modeled. In fact, a working child may accumulate less human capital (resulting in her lower future income), however she would generate additional income by contributing to the production function of the household. As medical expenditure increases the marginal productivity of child labor, higher medical expenditure may be correlated to higher intensity of child labor.

The empirical analysis is carried out using a cross-section of data on Ghana for the year 1999. Our estimates deal with both unobserved heterogeneity and simultaneity bias by means of a three equation simultaneous Probit model that is estimated using both simulated maximum likelihood (SML) techniques and a new algorithm developed by Lazard-Holly and Holly (2003) that allows the obtaining of a closed form solution of the likelihood function that is then estimated by FIML technique.

[2] Excellent reviews of the recent literature on hazardous work can be found in Basu and Tzannatos (2003), while O'Donnel, Rosati and Van Doorslaer (2005) is a good reference for studies on non-hazardous work.

The layout of the paper is as follows: Section 2 contains the theoretical model; in Section 3 we present the data set employed for the empirical analysis; in Section 4 we illustrate the empirical model, the main econometric problems encountered and the relative solutions adopted; in Section *5* the empirical results are discussed. Finally, in the last section we draw the conclusions and report some policy implications of our findings.

2. - Theoretical framework

Preserving a good health status depends on several variables. In the specific case of children, we could imagine that a working child may receive less preventive care and less monitoring of health conditions compared to those enrolled, for example, in a school program. At the same time, child labor is often a non-trivial source of household income and it may be intertwined with the intensity of health care access, especially in those systems where health care is not free. According to ORD findings, not only is non-hazardous work non detrimental to health in the short run, but it can also have a positive effect on nutritional status of children and their health.

The rationale of this conclusion goes back to Pitt et al. (1990), who introduced the idea of a welfare maximizing family unit in which a large amount of resources is devoted to maintaining the nutritional and health status of its productive members. Consequently, notwithstanding the impact of possible health hazards confronted in the workplace, non-working children might be expected to experience lower nutritional status and greater morbidity than their working siblings. Empirical support for this hypothesis was later found by Immink and Payongayong (1999) in a study in rural Guatemala. Their results show that while participation of school-age children in farm production was not associated with a reduction in their own growth and development, younger siblings not participating in farm production experienced growth deficits. Similar results are found also by Ralston (1997) in a study on calorie intakes in rural Indonesia, where the intra-household calorie allocation is related to children's labor contributions and lower calorie intakes are associated with higher levels of morbidity. Clearly, a decent nutritional status is not the only way to preserve a good health status: access to health care services may be just as important, or even more so.[3]

We model child labor, school attendance and medical expenditure in a two-period model where parents are altruistic, as is commonly used in the literature (see, among others, Rosati and Rossi, 2003). The number of children is taken as given and for simplicity of exposition is normalized to 1.

We assume that human capital accumulation is the only way to transfer resources for children's future consumption. Human capital is accumulated by sending children to school. Hours spent at work reduce time available for study and to enjoy leisure time, tire the child and reduce her learning productivity. Under these hypotheses, we can assume the existence of a parent inter-temporal utility function in which each single child health status is a choice variable that is positively related to the parent utility level.

Each child devotes her time in three possible activities: leisure, school and work. Extending on the previous literature, we model leisure as a choice of the family by overcoming the assumption that child labor trivially mirrors schooling. Thus, every child's time devoted to school (h_s) displaces time that can be spent on the labor market (h_c) or free time (l). Normalizing the total amount of hours to one, we have that:

$$h_s=1-h_c-l \qquad (1)$$

[3] We can easily think of situations where an inadequate nutritional status may be detected only after a visit to a physician.

Human capital accumulation is a function of the time spent at school according to the following (see Baland and Robinson (2000), Pouliot (2006) and Fan (2004) for similar approaches):

$H=f(h_s)$, with $f'>0$ and $f''<0$ (2)

Children's time spent on work contributes to the production function of the household, Y, where the inputs are child labor, equal to time neither spent in school nor as leisure time ($h_c=1- h_s-l$, parents' worked hours (H) and children's medical status, proxied by medical treatment, M.

$Y=Y(h_c,H,M)$ with $Y_{hc}>0$, $Y_{hc,hc}<0$, $Y_{M>}0$ and $Y_{M,M}<0$; $Y_{hc,M}>0$ (3)

On the one hand, time spent at school would deter child labor and thus reduce current resources available to the household. On the other hand, it would increase future consumption through the channel of higher human capital accumulation.

Parents maximize a two-period utility function over medical expenditure, hours of school and saving (s) as follows:

$$\max_{M,s,hc} U(Y(h_c,H,M)-s-pM+v(l))+V(f(1-h_c-l)+s) \quad (4)$$

The utility of leisure is translated into monetary equivalent with the function v().

U is the parental utility and V is their children's utility in the next period. M is the amount of medical services at price p, (h_c) is the number of hours worked in the market by the child, and H (exogenous) is the number of hours worked by parents and s is saving. For simplicity we set interest rate equal to zero.

In period one, household consumption is equal to:

$$c_1 = Y(h_c,H,M)-s-pM \quad (5)$$

As parents are altruistic, they also derive utility from consumption of their children as adults. Their future consumption consists of the resources saved and a function of the accumulated human capital, $f(h_s)$, such that:

$$c_2 = V(f(h_s)+s) = V(f(1-h_c-l)+s) \quad (6)$$

From first order conditions with respect to M, h_c, l, and s, we obtain, respectively:

$$U'\big(Y(h_c,H,M)-pM-s+v(l)\big)\big(Y_M-p\big)=0 \quad (7)$$

$$U'\big(Y(h_c,H,M)-pM-s+v(l)\big)Y_{hc} = V'\big(f(1-h_c-l)+s\big)f'(1-h_c-l) \quad (8)$$

$$U'\big(Y(h_c,H,M)-pM-s+v(l)\big)v'(l) = V'\big(f(1-h_c-l)+s\big)\mathrm{f}'() \quad (9)$$

$$U'\big(Y(h_c,H,M)-pM-s+v(l)\big) = V'\big(f(1-h_c-l)+s\big) \quad (10)$$

Where the subscript indicates the partial first derivative with respect to that variable. From equation (11) we can derive that at the optimum the marginal benefit of an additional visit should be equal to its price:

$$Y_M\left(h*_c(p), H, M*(p)\right) = p \tag{11}$$

where the upper script * indicates the optimal level of the control variables. Dividing (8) over (10) the usual efficiency condition of child labor applies (see Baland and Robinson, 2000). Thus child labor is set to its efficiency level such that the marginal productivity of an additional hour of child labor equals that of foregone schooling:

$$Y_{hc}\left(h*_c(p), H, M*(p)\right) = f'\left(1 - h*_c(p) - l*(p)\right) \tag{12}$$

Dividing (8) over (9) we also get that:

$$v'(l*(p)) = f'\left(1 - h*_c(p) - l*(p)\right) \tag{13}$$

Using implicit differentiation of the optimal condition in (12) and (13), we obtain the following

$$\frac{dh_c}{dp} Y_{h_c,h_c} + Y_{M,hc}\frac{dM}{dp} = -f''\left(1 - h_c - l\right)\left(\frac{dh_c}{dp} + \frac{dl}{dp}\right) \tag{14}$$

$$v''(l)\frac{dl}{dp} = -f''\left(1 - h_c - l\right)\left(\frac{dh_c}{dp} + \frac{dl}{dp}\right) \tag{15}$$

Or, equivalently,

$$f'''()\frac{dh_c}{dp} = \left(-f''\left(1 - h_c - l\right) - v''(l)\right)\frac{dl}{dp}$$

From which, substituting the expression for dl/dp into equation 14, we can derive that:

$$\frac{dM}{dh_c} Y_{hc,M} = \left(-Y_{hc,hc} - f''\left(1 - h_c - l\right)\right) + \frac{f''^2}{v'' + f''} \tag{16}$$

From equation (16) we can derive that dM/dh_c is positive if

$$-Y_{hc,hc} - f''\left(1 - h_c - l\right) > \frac{-f''^2}{v'' + f''}$$

Which is verified for any concave or linear function v.

As a consequence, medical expenditure and hours of child labor are positively correlated. Let us give the intuition behind this result. Suppose that the price of medical services raises, as a result the demand for medical expenditure will decrease. Less money spent on medicine or visits to the doctor will, in turn, decrease the marginal productivity of child labor, as an additional input of labor will add less productivity due to a lower level of medical investment in the production function. This would increase the opportunity cost of

child labor, with respect to school, which becomes a more attractive investment than child labor. To reach again the equilibrium, as stated in equation (12), marginal productivity of child labor must equate that of time spent at school, which can be verified by decreasing both time spent at work and leisure time. Therefore, factors deterring medical expenditure, such as their price, cause time spent at school to increase and as a consequence, time spent on the labor market will be reduced.

Our empirical analysis tests this implication by estimating if and to what extent medical expenditure might enhance child labor by testing the theoretical implication found.

3. - The database used

The data used in this analysis come from the fourth round of the Ghana Living Standard Survey (GLSS) obtained from the Ghana Statistical Service (October 2000). The amount of information collected with this survey allows the provision of information on patterns of household consumption and expenditure disaggregated at greater levels, to provide in-depth information on the structure and composition of the wages and conditions of work of the labor force in the country and to identify vulnerable groups for government assistance in fields such as education and health.

Out of the selected 6,000 households 5,999 were successfully interviewed. One household was further dropped during the data cleaning exercise because it had very few records for many of the sections in the questionnaire. This gave 5,998 households representing 99.7% coverage. Overall, 25,694 *eligible household members* (un-weighted) were covered in the survey. The survey was spread over a 12-month period (from 1 April 1998 to 25 March 1999) in order to ensure continuous recording of household consumption and expenditure and changes occurring therein. The year was divided into 10 cycles of 36 days each.

For the purpose of this survey the list of the 1984 population censual Enumeration Areas (EAs) with population and household information was used as the sample frame. The primary sampling units were the 1984 EAs and the secondary units were the households in the EAs. An EA is a demarcated geographic area consisting of a locality or group of localities with a precise boundary description. On average EAs have about 200 households. This frame may be considered rather inadequate, but is the best available in the system. Indeed, this frame was used for the earlier rounds of the GLSS.

In order to increase precision and reliability of the estimates (the 1984 frame may be considered rather inadequate) the technique of stratification was employed in the sample design using geographical factors, ecological zones and location of residence as the main controls. Specifically, the EAs were first stratified according to the three ecological zones (Coastal, Forest and Savannah), and then within each zone further stratification was done based on the size of the locality into rural or urban.

A two-stage sample was selected for the survey. In the first stage, 300 EAs were selected using systematic sampling with probability proportional to size method (PPS) where the size measure is the 1984 number of households in the EA. This was achieved by ordering the list of EAs with their sizes according to the strata. The size column was then cumulated, and with a random start and a fixed interval the sample EAs were selected. It was observed that some of the selected EAs had grown in size over time and therefore needed segmentation. In this connection, such EAs were divided into approximately equal parts, each segment constituting about 200 households. Only one segment was then randomly selected for listing of the households. At the second stage, a fixed number of 20 households were systematically selected from each selected EA to give a total of 6,000 households.

3.1 Definition and descriptive statistics of the variables

The full dataset included 25,649 eligible household members, of which 7,326 are children. They have been identified as members of the survey whose age ranges between 7 and 16 years. Among them, 4,945 children (67.5%) live in rural areas and the remaining 2,381 (32.5%) live in urban areas. Due to availability of data at community level only in rural areas, our sample will be limited to those children living in those areas. Finally, after dropping a number of cases (216 – about 4%) for which information on some variables (capacity of reading and writing in English and Ghanaian and religion of head of household) are missing, our final dataset is composed of 4,729 observations.

Table 1 reports the descriptive statistics referring to this dataset. The variables have been grouped into "dependent" and "independent" and, among these, in variables referring to children, to households and the communities to which they belong.

Concerning the dependent variables, we can see that only 6.1% of children had a visit to any form of health care services in the previous two weeks. Access to health care services may include physicians' visits as well as dentists, nurses, traditional healers, medical assistants, etc. It is important to stress that access to health care services excludes visits due to accidents, which could be correlated to child labor intensity by affecting the analysis robustness. Due to the way the health variable is constructed, it represents a good proxy for health investment as it does not include visits for treatments after accidents.

A child is considered to work if, over the past 12 months, she declared to have been involved either in paid (in cash or in kind) or unpaid work, as is usually the case in the child labor literature. Housekeeping activities have not been included in this definition of child labor, following the ILO definition of child labor. A total of 22.8% of children living in rural areas and aged between 7 and 16 years were involved in child labor activity in 1999. At the same time, 81.5% of children have declared to be officially enrolled in a school program.

On average the self-reported health status of children is quite satisfactory, given that more than 80% have declared a "good health status".[4] About 18% of children report that they have suffered bad health status in the previous month. For most of them (87%) the spell of the bad health condition is limited to 7 days, with an average of about 4 days. In terms of education more than 80% declare to be able to read and write in English and about 50% to read and 20% to write in Ghanaian. Most of them (78%) are Christians.

At household level, we can see that the average level of education is quite low (either illiterate or with primary school education), with the heads of household slightly more educated than the mothers/wives. In about 70% of cases the wife lives in the household and in 97% of cases the head of household has a paid job. Household size is quite large (about 7 individuals on average). Concerning household income, we report the per capita household income.[5]

---

[4] - The survey asks if during the past 2 weeks the interviewed suffered from either injury or illness. Those responding "neither" have been categorized as being in "good health status". Those who answered as suffering from "illness" have been categorized as being in "bad health status".

[5] - In a previous stage of our research, we have also constructed another variable measuring the "household income net of the child income", computed summing up all sources of labor and non labor income accruing in a household and earned solely by adults (aged 17 and over). The idea behind the construction of such a measure was to help in identifying those situations in which child labor can be revealed as being crucial for the household budget. Unfortunately, this new variable had a strong correlation with the previous one, mainly due to the fact that in both measures we can only proxy the true "child income". This is for two main reasons: (a) although the survey asks for a monetization of the salary received "in kind", the quantification is based on subjective estimates and therefore the quality of the data reported is very poor; and (b) in all those cases where children have declared experiences with "unpaid work in family enterprise" it has been difficult to assign a value to the implicit wage/saving accruing to the household that can avoid hiring a non-household

Approximately all households (93%) own or operate a piece of land, a common feature to rural areas where almost all households are involved in agricultural activities.

The presence of health care services at community level is quite rare. Family planning workers and trained midwives are present only in about 20% of the communities surveyed. Traditional healers are present in about 70% of communities. Other forms of health care services such as ambulatory health care with physicians, pharmacies and hospitals are indeed very rare (below 2%).

The presence of a primary school is common in about 88% of local communities, while this percentage drops to 65% in the case of junior schools. For some communities the distance to these schools is quite high. Lack of school infrastructures and school costs are seen as the main problem for enrolling children in school programs in many communities. As we will see better in the next paragraphs, we will use these variables as exclusion restrictions for a better identification of our parameters.

4. The econometric model

The theoretical model presented in the previous section suggests using a three equation simultaneous Probit model in which the two equations for the probability of choosing to work and to attend school are reduced form equations, while the third equation is a structural equation for the propensity to use any form of medical treatment, conditional on the decision to work and/or to attend school and to a number of other exogenous covariates. Thus, this last equation contains two endogenous dummy variables and a set of other covariates.

We assume y_1^* to be an endogenous variable representing the propensity of a child to attend school, y_2^* an endogenous variable representing the propensity of the same child to work, and y_3^* the propensity he/she has to visit any form of health care service. This three-equation model simply relates school attendance and work decisions to health service access.

The structural form of our model could, then, be rewritten in the following way:

$$\begin{aligned} y_1^* &= \mathbf{x_1'}\boldsymbol{\beta}_1^0 + l_1^0 \\ y_2^* &= \mathbf{x_2'}\boldsymbol{\beta}_2^0 + l_2^0 \\ y_3^* &= \alpha_{31}^0 y_1^* + \alpha_{32}^0 y_2^* + \mathbf{x_3'}\boldsymbol{\beta}_3^0 + l_3^0 \end{aligned} \tag{17}$$

where the variables l_i represent latent variables. In fact, to the extent that some individual characteristics are unobserved, they are absorbed into the latent variables l_i. Among individual characteristics, unobserved components of individual and family health history, attitudes towards health risks, lifestyle choices etc. that influence individual perceptions of health events are likely to be important. It is also reasonable to expect the l_i to be correlated with each other and this is a possible source of simultaneous equations bias if not taken into account in estimation.

The Log-likelihood function for a sample of N independent observations will then be equal to:

$$L = \Sigma_i \, w_i \log\Phi_3(\mu_i; \Omega) \tag{18}$$

where $\Phi_3(\mu_i; \Omega)$ is a standard trivariate normal CDF, with the vector of means equal to:

---

member. As a result, only in a very few cases (about 40) have we been able to identify significant differences between the two measures of income.

$$\mu_i = (K_{i1}\beta_1'X_{i1}\ ,\ K_{i2}\beta_2'X_{i2}\ ,\ K_{i3}\beta_3'X_{i3}), \qquad K_{ik} = (2y_{ik} - 1), \text{ for each } j,\ k = 1,\ldots,3$$

the matrix Ω has elements Ω_{jk} , where $\Omega_{jj} = 1$ for $j = 1, \ldots, 3$; $\Omega_{21} = \Omega_{12} = K_{i1}K_{i2}\rho_{21}$, $\Omega_{31} = \Omega_{13} = K_{i3}K_{i1}\rho_{31}$, $\Omega_{32} = \Omega_{23} = K_{i3}K_{i2}\rho_{32}$, and w_i is an optional weight. Until recently, the estimation of models such as (18) could be carried out only by the standard technique of Simulated Maximum Likelihood (SML). In fact, this technique allows us to model the correlation structure of the error term, and the endogeneity structure arising from the estimation of a structural model. However, no closed form solution for the likelihood function was available. Standard procedures such as the Geweke-Hajivassiliou-Keane (GHK) simulator evaluate the M-dimensional normal integrals in the likelihood function.[6] For each observation, a likelihood contribution is calculated for each replication, and the simulated likelihood contribution is the average of the values derived from all the replications. The simulated likelihood function for the sample as a whole is then maximized using standard methods.

Under standard conditions, the SML estimator is consistent as the number of observations and the number of draws tends to infinity and it is asymptotically equivalent to the true maximum likelihood estimator as the ratio of the square root of the sample size to the number of draws tends to zero. Thus, all things being equal, the more draws, the better. Estimation is numerically intensive and may be very slow if the data set is large, if the number of draws is large, or (especially) if the number of equations is large. Models for which the variance-covariance matrix of the cross-equation error terms (ρ_s) is close to not being positive definite are likely to be difficult to maximize. Furthermore, results may differ depending on the sort order of the data, because the sort order affects which values of the random variable(s) get allocated to which observation.[7]

As an alternative, recently Lazard-Holly and Holly (2003) and Huguenin (2004) have derived closed form solution for the Log-likelihood function of models such as in (17), which could be maximized by means of LIML and/or FIML, using BHHH algorithm. To this end, they have derived the first-order derivatives with respect to the parameters of Log-likelihood estimation procedures. Obviously, use of the analytical method suggested by Lazard-Holly and Holly (2003) is much quicker and more accurate than the simulation procedures available in the literature. Moreover, it reduces the uncertainty on the final estimates due to the simulation procedure.

Although issues of model identification arise due to the introduction of endogenous variables (y_1 and y_2) as regressors in the third equation (y_3), in this case the identification of the causal parameters through non-linear functional forms is feasible in principle. However, for more robust identification we introduce nontrivial exclusion restrictions. In this respect, we use information at community level on the availability of health care services (see Section 5.2 for more details).

5. Empirical results

The main goal of this empirical analysis is to check if working children have more favorable access to health care services compared with non-working children. At the same time, we check if being enrolled in a school program does increase the probability of visiting health care services. We condition these results on a large set of covariate that

[6] See for example the command MVProbit in Stata

[7] This potential problem is reduced by increasing the number of random draws that is used (that in turn increases the computational time).

should take into account child specific preference/characteristics, household characteristics and community level characteristics.

As already discussed in the fourth paragraph, we have estimated the same model using two econometric techniques. First we have estimated a multivariate Probit model using SML techniques by means of the Geweke-Hajivassiliou-Keane (GHK) simulator to evaluate the 3-dimensional normal integrals in the likelihood function. We have then estimated the same model using the novel methodology developed by Lazard-Holly and Holly (2003). In what follows we will only discuss the results from this last methodology, and will limit the comparison with the SML only to the model fit and to the parameters of "child labour" (y_1) and "school" (y_2) in the visit equation (the third equation in (17)). For computational simplicity, we have estimated the marginal effects and elasticities at the sample average of the other covariates. Clearly, the magnitudes of the marginal effects would change if evaluated at other points, but the qualitative conclusions should not.

5.1 Model fit

Table 2 reports the pseudo log-likelihood values for the different models estimated. In order to solve the problem arising with the sort order of the data when using the SML technique, we have produced estimates for different numbers of random draws (from 50 to 1,500). According to these results the log likelihood is maximized around values of draws of 1,500. In any case, the model seems to be quite robust to random draws choice and this may be considered a good indication of the correct specification of our structural model. The log likelihood for the closed form solution specification is slightly lower, but using a likelihood ratio test it is hard to find any statistical significance for a difference among the two. The closeness of these sets of results provides support in favor of the correct model specification and result robustness.

Concerning the parameter values of the endogenous variables, all specifications provide a statistical significant parameter for the "child labor" dummy. Similar results do not hold for the "school" dummy parameter. In terms of magnitude the "child labor" parameter estimated using the FIML technique is a little bit lower (0.78) compared to the parameters estimated using the SML technique (from 0.79 to 0.88). In contrast, the results for the "school" parameters change with the number of draws (from 0.22 to 0.49).

5.2 The determinants of health care access

The results in terms of parameter estimates are reported in Table 3 (FIML technique using closed form solution), while Table 4 reports the marginal effects[8] associated with our regressors, as derived from the FIML estimates.[9] In what follows we focus on the variables that in our opinion are most relevant for our study. All results, although not always significant, have the correct sign from the theoretical point of view.

Concerning the probability of working it is interesting to note that among the variables capturing the well-being of a child, owning or operating a plot of land is the most influential variable to explain the working activities of a child. Owning land increases the probability of a child being involved in some labor activity by more than 20 percentage points while household income affects neither labor nor schooling decisions. This is in line with results obtained by Beegle et al. (2004) in which they find that landowning households

[8] Marginal effects are computed, as usual, for each observation as $dE(y_i|x_j)/dx_j$ for the dependent variable *i* with respect to the continuous independent variable *j*, and as $E(y_i|y_j=1)-E(y_i|y_j=0)$ for the dependent variable *i* with respect to the dummy dependent variable *j*, and averaged over all observations.

[9] Results based on the SML technique are available upon request.

could have a greater demand for child labor, given that household land ownership is both a proxy for wealth, but in an agricultural setting it is also correlated with the demand for child labor (see Bhalotra and Heady, 2001). Household income does, instead, increase health-associated outputs. The inability of household income to explain schooling and working decisions, although the theoretical predictions would suggest a positive relation,[10] could be partly due to problems existing in the correct measurement of the variable. Finally, there doesn't seem to be any gender discrimination in the probability of being involved in labor activities.

An older child with some degree of apprenticeship has a higher chance of being a working child. Conversely, being a Christian or Muslim, as well as belonging to a large household, reduces the probability of working.

The probability of being enrolled in a school program is highly and positively influenced by parental educational attainments, religious practice and gender (males are more favored). The presence of an elderly person in the household reduces this probability of school enrolment by almost ten percentage points. This could be explained by the fact that children may be used for care-giving within the household. In the same way, having some apprenticeship reduces the probability of being enrolled in a school program – this is symmetrical to what was obtained for the child labor equation. As said before, household income is not significant in affecting the probability of being a student. At the same time, child income as a separate explanatory variable has never appeared to be significant in any specification tested. Lack of school infrastructure and school costs strongly reduce the probability of being enrolled in a school program. This result also confirms those obtained by Beegle et al. (2004).

Finally, the probability of going to any health care service provider is highly influenced by the health status and by the seriousness of the illness (measured as the number of days of bad health conditions). At the same time, belonging to wealthier households increases the probability of visiting health care providers. Age, sex and religious practices do not influence this probability at all. On the other hand, the presence of some providers[11] (a supply side effect) significantly affects the probability of going to the health provider for a medical visit. In particular, the presence of traditional healers has a positive effect, while that of family planning workers has a negative impact.

However, the most important result is the positive and significant parameter for being a working child in the visit equation. This result confirms our theoretical prediction, namely that a working child has a higher probability than her non-working siblings of receiving health care treatment by five percentage points. At the same time, being enrolled in a school program does not seem to have any significant effect on the probability of accessing health care services. It is worth stressing that by relaxing the endogeneity assumption (thus relying on the simpler multivariate Probit model) the significance of the child labor parameter is lost.

5.3 The exclusion restrictions

[10] Beegle et al. (2004) observe a significant and negative effect of wealth on child labour, although they recognize that the magnitude is very low.

[11] Our dataset includes a much larger array of health care providers compared to those listed in our empirical model. However, as already discussed in Section 3.1, in many cases the availability of physicians, hospitals, pharmacists, etc. is extremely rare across local communities. This partly explains why we have not been able to find any statistical significance for most of these providers in our model. In order to preserve parsimony, we have then decided to limit our "provider" covariates to those for which at least 20% of local communities report its presence.

As mentioned in Section 4, we use exclusion restrictions to aid identification of the causal parameters. We have identified two potential sets of variables for exclusion, included in the schooling equation and the labor equation, which are excluded from the health expenditure equation. The first set is represented by skill variables such as writing and reading ability, which should help children to stay away from child work. We expect these variables to enter with a negative sign in the child labor equation, but not to have direct effects on the numbers of visits to doctors. The second set of variables is represented by school supply variables (number and type of school available) and cost of transportation variables to go to school at community level. We expect that these variables affect the choice to go to school (with positive and negative sign, respectively) but, again, not the number of visits. Note also that our model is theoretically identified by functional form only, a variant we have also estimated (but not reported in the tables presented). The results shown in Tables 3 and 4 confirm these theoretical predictions.

6. Conclusions

In this paper we develop a theoretical set-up guiding our empirical strategy to estimate the causal relationship linking child labor, schooling and access to health care in LDCs. According to our predictions we expect higher levels of child labor to be correlated with more frequent access to health care services. Higher medical expenditure increases the marginal productivity of labor by making school investments, whose return is unchanged, less attractive. We have tested these predictions drawing from a Living Standard Measurement Survey dataset on Ghana for the year 1999. Our empirical results, using a structural three equation simultaneous Probit model, confirm the theoretical implication.

From a policy standpoint, our results reveal an important result. In fact, policies aiming at reducing child labor may induce a reduced amount of health care treatment among several children, if compensations to the family do not compensate for the foregone productivity lost. In that respect, policy makers should then put more effort in designing policies that could counterbalance such negative effects, for example by making returns to education higher or compensating households for the foregone productivity when children are not sent to work.

References

Baland, Jean-Marie, Robinson, James A. Is child labor inefficient? Journal of Political Economy 2000; 108 (4), 663-681.

Basu K. and Z. Tzannatos, The Global Child Labor Problem: What Do We Know and What Can We Do?, The World Bank Economic Review, 2003, vol. 17(2), pp. 147-173.

Beegle K., Dehejia R. and Gatti, R. Why should we care about child labor? The education, labor market, and health consequences of child labor. NBER WP 10980, 2004.

Bhalotra, S. and Heady, C. Child Farm Labor: The Wealth Paradox. Mimeo; 2001.

Fan, S. Relative Wage, Child Labor, and Human Capital. Oxford Economic Papers 2004; 56(4):687-700.

Ghana Living Standard Survey. Ghana Statistical Service; October 2000.

Huguenin, J. The Multivariate Normal Distribution and Multivariate Probit Analysis. PhD Thesis; 2004.

International Labour Organization. A Future Without Child Labour. Geneva:International Labour Office; 2002.

Immink, M.D.C. and Payongayong, E. Risk analysis of poor health and growth failure of children in the central highlands of Guatemala. Social Science and Medicine 1999; 48: 997-1009.

Lazard-Holly, H. and Holly, A. Computation of the Probability that a d--dimensional Normal Variable Belongs to a Polyhedral Cone with Arbitrary Vertex. Mimeo; 2003.

O'Donnel O., Rosati F.C., and van Doorslaer, E. Health effects of child work: Evidence from rural Vietnam. Understanding Journal of Population Economics 2005; 18, pp.437-467.

Pitt, M.M., Rosenzweig, M.R. and Hassan, N. Productivity, health and the intrahousehold distribution of food in the low-income countries. American Economic Review 1990; 80, 1139-1156.

Pouliot W. Introducing uncertainty into Baland and Robinson's model of child labour, Journal of Development Economics 2006; 272 (79), pp. 264-272.

Ralston, H. Health as an input to labor: Intrahousehold food distribution in rural Indonesia. Journal of Policy Modeling 1997; 19(5): 567-86.

Rosati, F.C. and M. Rossi, Children's Working Hours and School Enrollment. Evidence from Pakistan and Nicaragua, 2003, 17(2), pp.283-295

UCW, Statistics portal, www.ucw-project.org.

Table 1
Descriptive statistics based on a sample of 4,729 children living in rural areas

List of variables	Mean	Std. Dev	Min	Max
Dependent variables				
Visit to any health service (0=no, 1=yes)	0.061	0.240	0.0	1.0
Child laboring (0=no, 1=yes)	0.228	0.420	0.0	1.0
Child attending school (0=no, 1=yes)	0.815	0.389	0.0	1.0
Independent variables ***at child level***				
Gender (female=0 male=1)	0.513	0.500	0.0	1.0
Dummy for child aged 11-15	0.476	0.499	0.0	1.0
Dummy for child aged 16-17	0.066	0.248	0.0	1.0
Dummy Christian religion (0=no, 1=yes)	0.782	0.413	0.0	1.0
Dummy Muslim religion (0=no, 1=yes)	0.094	0.292	0.0	1.0
Dummy Traditional religion (0=no, 1=yes)	0.088	0.283	0.0	1.0
Child is an apprentice (0=no, 1=yes)	0.001	0.025	0.0	1.0
Child writes English (0=no, 1=yes)	0.815	0.388	0.0	1.0
Child writes Ghanaian (0=no, 1=yes)	0.202	0.401	0.0	1.0
Child does written calculation (0=no, 1=yes)	0.232	0.422	0.0	1.0
Child reads letter in English (0=no, 1=yes)	0.837	0.369	0.0	1.0
Child reads letter in Ghanaian (0=no, 1=yes)	0.484	0.500	0.0	1.0
Dummy for GOOD health status	0.817	0.387	0.0	1.0
Days of illness	0.883	2.447	0.0	31.0
at household level				
Age of head of household (HH)	48.8	12.7	18.0	99.0
Age of child's mother	33.9	14.5	1.0	98.0
Presence of an elderly person (0=no, 1=yes)	0.102	0.303	0.0	1.0
Education of HH (0=none, 11=tertiary)	1.427	1.412	1.0	11.0
Education of mother (0=none, 11=tertiary)	1.133	0.689	1.0	11.0
Household owns or operates land (0=no, 1=yes)	0.934	0.248	0.0	1.0
Household size	6.791	2.699	2.0	21.0
Log of per capita income	12.773	0.925	9.0	16.0
Head of household works (0=no, 1=yes)	0.979	0.142	0.0	1.0
Wife living in the household (0=no, 1=yes)	0.689	0.463	0.0	1.0
at community level				
Presence of primary school (0=no, 1=yes)	0.875	0.331	0.0	1.0
Presence of junior school (0=no, 1=yes)	0.652	0.476	0.0	1.0
Distance to primary school (Km)	0.483	1.666	0.0	15.0
Square of primary school distance / 100 (Km)	0.030	0.176	0.0	2.3
Distance to junior school / 10 (Km)	3.821	27.053	0.0	480.0
Square of junior school distance / 10000 (Km)	0.075	0.906	0.0	23.0
Presence of private school (0=no, 1=yes)	0.120	0.325	0.0	1.0
School cost as main problem (0=no, 1=yes)	0.111	0.314	0.0	1.0
School cost as second problem (0=no, 1=yes)	0.075	0.263	0.0	1.0
School cost as third problem (0=no, 1=yes)	0.129	0.335	0.0	1.0
School cost as fourth problem (0=no, 1=yes)	0.083	0.276	0.0	1.0
Lack of school infrastr. as 1st problem (0=no, 1=yes)	0.381	0.486	0.0	1.0
Lack of school infrastr. as 2nd problem (0=no, 1=yes)	0.104	0.305	0.0	1.0
Lack of school infrastr. as 3rd problem (0=no, 1=yes)	0.079	0.270	0.0	1.0
Lack of school infrastr. as 4th problem (0=no, 1=yes)	0.055	0.228	0.0	1.0
Presence of trained midwife (0=no, 1=yes)	0.227	0.419	0.0	1.0
Presence of family planning worker (0=no, 1=yes)	0.215	0.411	0.0	1.0
Presence of traditional healer (0=no, 1=yes)	0.687	0.464	0.0	1.0

Table 2
Model fit comparison

Model	Log-Pseudo Likelihood	Coefficient "child labor"	Coefficient "school"
SML with 50 draws	-4390.0	0.8502[(a)]	0.2288[(c)]
SML with 200 draws	-4389.4	0.8624[(a)]	0.3973[(c)]
SML with 500 draws	-4388.2	0.8887[(a)]	0.4381[(c)]
SML with 750 draws	-4388.2	0.8735[(a)]	0.4929[(c)]
SML with 1500 draws	-4388.6	0.7931[(a)]	0.3850[(c)]
FIML with closed form solution	-4391.8	0.7792[(b)]	0.2625[(c)]

[(a)] *Significant at 1% level*
[(b)] *Significant at 5% level*
[(c)] *Not significant*

Table 3
Three-Equations Simultaneous ML Probit Estimation[1]

Variable codes	CHILD LABOR Equation			SCHOOL Equation			VISIT Equation		
	Coeff.	t-stat	Prob.	Coeff.	t-stat	Prob.	Coeff.	t-stat	Prob.
at child level									
Gender (female=0 male=1)	0.043	0.970	0.330	0.204	4.360	0.000	-0.067	-0.760	0.448
Dummy for child aged 11-15	0.693	13.090	0.000	0.005	0.100	0.917	-0.078	-0.550	0.581
Dummy for child aged 16-17	1.076	11.590	0.000	-0.548	-5.970	0.000	0.022	0.350	0.728
Dummy Christian religion (0=no, 1=yes)	-0.507	-4.690	0.000	0.813	7.840	0.000	0.079	0.300	0.762
Dummy Muslim religion (0=no, 1=yes)	-0.320	-2.460	0.014	0.583	4.460	0.000	0.349	1.280	0.199
Dummy Traditional religion (0=no, 1=yes)	-0.266	-2.050	0.040	0.404	3.170	0.002	-0.248	-0.890	0.373
Child is an apprentice (0=no, 1=yes)	5.717	32.080	0.000	-1.411	-2.330	0.020			
Child writes English (0=no, 1=yes)	0.133	0.870	0.385						
Child writes Ghanaian (0=no, 1=yes)	-0.346	-2.650	0.008						
Child does written calculation (0=no, 1=yes)	-0.107	-1.740	0.082						
Child reads letter in English (0=no, 1=yes)	-0.283	-1.980	0.047						
Child reads letter in Ghanaian (0=no, 1=yes)	0.128	1.000	0.315						
Dummy for GOOD health status							0.055	4.020	0.000
Days of illness							-2.495	-13.200	0.000
at household level									
Age of head of household (HH)	-0.002	-1.000	0.317	-0.002	-0.870	0.386	-0.008	-2.070	0.038
Age of child's mother	0.002	1.020	0.307	0.008	2.820	0.005	-0.004	-1.210	0.226
Presence of elderly pers. (0=no, 1=yes)	0.046	0.600	0.547	-0.339	-4.510	0.000	-0.342	-2.140	0.032
Education of HH (0=none, 11=tertiary)	-0.024	-1.490	0.137	0.067	3.330	0.001			
Education of mother (0=none, 11=tertiary)	0.001	0.020	0.980	0.152	2.670	0.008			
Household owns or operates land (0=no, 1=yes)	0.873	6.000	0.000	0.049	0.440	0.658			
Household size	-0.027	-2.880	0.004	-0.028	-2.930	0.003			
Log of per capita income	0.013	0.680	0.493	0.020	1.240	0.216	0.120	2.480	0.013
Head of household works (0=no, 1=yes)	0.406	2.410	0.016				-0.296	-1.070	0.283
Wife living in the household (0=no, 1=yes)				0.049	0.650	0.514			
at community level									
Presence of primary school (0=no, 1=yes)				0.076	0.450	0.652			
Presence of junior school (0=no, 1=yes)				0.077	1.180	0.236			
Distance to primary school (Km)				-0.079	-1.260	0.209			
Square of primary school distance / 100 (Km)				0.705	1.550	0.120			
Distance to junior school / 10 (Km)				0.003	0.500	0.619			
Square of junior school distance / 10000 (Km)				-0.037	-0.230	0.819			
Presence of private school (0=no, 1=yes)				0.363	3.700	0.000			
School cost as main problem (0=no, 1=yes)				0.035	0.360	0.720			
School cost as second problem (0=no, 1=yes)				-0.196	-1.930	0.054			
School cost as third problem (0=no, 1=yes)				-0.138	-1.710	0.087			
School cost as fourth problem (0=no, 1=yes)				-0.147	-1.530	0.126			

Lack of school infrastr. as 1st problem (0=no, 1=yes)				-0.191	-3.060	0.002			
Lack of school infrastr. as 2nd problem (0=no, 1=yes)				0.023	0.240	0.812			
Lack of school infrastr. as 3rd problem (0=no, 1=yes)				-0.244	-2.420	0.015			
Lack of school infrastr. as 4th problem (0=no, 1=yes)				0.290	2.020	0.044			
Presence of trained midwife (0=no, 1=yes)							0.062	0.400	0.691
Presence of family planning worker (0=no, 1=yes)							-0.349	-2.250	0.024
Presence of traditional healer (0=no, 1=yes)							0.267	2.590	0.009
Endogenous variables									
childwork							0.894	2.920	0.003
school							0.304	0.720	0.471

(1) Parameters for constants, regions and ethnic groups have been omitted for space reasons. They are available upon request from the authors. Sample size is 4,729 observations. Standard errors have been computed using robust standard error formula.

Table 4
Marginal effects from 3-Equation Simultaneous ML Probit Estimation

	CHILD LABOR Equation	SCHOOL equation	VISIT equation
at child level			
Gender (female=0 male=1)	0.011	0.044	-0.004
Dummy for child aged 11-15	0.172	0.001	-0.096
Dummy for child aged 16-17	0.266	-0.118	0.027
Dummy Christian religion (0=no, 1=yes)	-0.123	0.176	-0.004
Dummy Muslim religion (0=no, 1=yes)	-0.084	0.132	0.02
Dummy Traditional religion (0=no, 1=yes)	-0.067	0.086	-0.017
Child is an apprentice (0=no, 1=yes)	1.438	-0.314	
Child writes English (0=no, 1=yes)	0.032		
Child writes Ghanaian (0=no, 1=yes)	-0.086		
Child does written calculation (0=no, 1=yes)	-0.025		
Child reads letter in English (0=no, 1=yes)	-0.070		
Child reads letter in Ghanaian (0=no, 1=yes)	0.029		
Days of illness			0.003
Dummy for GOOD health status			-0.163
At household level			
Age of head of household (HH)	-0.051	-0.042	-0.057
Age of child's mother	0.041	0.170	-0.023
Presence of elderly pers. (0=no, 1=yes)	0.01	-0.073	-0.022
Education of HH (0=none, 11=tertiary)	-0.062	0.152	
Education of mother (0=none, 11=tertiary)	0.009	0.323	
Household owns or operates land (0=no, 1=yes)	0.215	0.011	
Household size	-0.134	-0.122	
Log of per capita income	0.078	0.086	0.161
Head of household works (0=no, 1=yes)	0.099		-0.019
Wife lives in household (0=no, 1=yes)		0.011	
at community level			
Presence of primary school (0=no, 1=yes)		0.016	
Presence of junior school (0=no, 1=yes)		0.018	
Distance to primary school (Km)		-0.344	
Square of primary school distance / 100 (Km)		0.308	
Distance to junior school / 10 (Km)		0.286	
Square of junior school distance / 10000 (Km)		-0.157	
Presence of private school (0=no, 1=yes)		0.078	
Lack of school infrastr. as 1st problem (0=no, 1=yes)		0.009	
Lack of school infrastr. as 2nd problem (0=no, 1=yes)		-0.042	
Lack of school infrastr. as 3rd problem (0=no, 1=yes)		-0.031	
Lack of school infrastr. as 4th problem (0=no, 1=yes)		-0.028	
School cost as main problem (0=no, 1=yes)		-0.040	
School cost as second problem (0=no, 1=yes)		0.007	
School cost as third problem (0=no, 1=yes)		-0.053	
School cost as fourth problem (0=no, 1=yes)		0.062	
Presence of family planning worker (0=no, 1=yes)			0.004
Presence of traditional healer (0=no, 1=yes)			-0.024
Presence of trained midwife (0=no, 1=yes)			0.017
Endogenous variables			
Child working			0.051
Child attending school			0.029