

IZA DP No. 3531

School Attendance of Children and the Work of Mothers: A Joint Multilevel Model for India

Francesca Francavilla
Gianna Claudia Giannelli
Leonardo Grilli

June 2008

School Attendance of Children and the Work of Mothers: A Joint Multilevel Model for India

Francesca Francavilla

*University of Florence
and Policy Studies Institute*

Gianna Claudia Giannelli

*University of Florence,
CHILD and IZA*

Leonardo Grilli

University of Florence

Discussion Paper No. 3531

June 2008

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0

Fax: +49-228-3894-180

E-mail: iza@iza.org

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

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

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

ABSTRACT

School Attendance of Children and the Work of Mothers: A Joint Multilevel Model for India^{*}

This paper investigates the determinants of school attendance of children and their mother's working status when the mother decides how to allocate her time and that of her children. A multilevel random effects model is applied to study the mother's participation and the schooling status of her children in a joint framework. Using the second National Family Health Survey (NFHS-2) for India, we find that, controlling for many covariates among which wealth is the most powerful predictor, children of working mothers have a lower probability of attending school. This, together with the result that only illiterate and poor mothers with unskilled or unemployed partners have a high probability of working, points to the need for decent labour market opportunities for females. An implication of our findings is that any policy aiming both at enhancing women's empowerment through labour and increasing children's welfare should also target improvements in women's conditions in the labour market.

JEL Classification: J13, J22, O15, O18

Keywords: children's schooling, women's work, household allocation of time, random effects, India

Corresponding author:

Gianna Claudia Giannelli
University of Florence
Department of Economics
Via delle Pandette 9
50127 Firenze
Italy
E-mail: giannelli@unifi.it

^{*} We presented this paper in April 2008 at an IAB workshop in Nürnberg and in May at the third IZA/World Bank Conference on Employment and Development in Rabat. We thank participants for their useful comments. We are grateful to Hilary Salter of PSI for her editing suggestions. The usual disclaimer applies.

1 Introduction

Child schooling is universally acknowledged as one of the prerequisites of human development. Although there is less agreement on the role of labour for women, its empowering function is hardly denied. In fact, it is generally thought that women who contribute to household resources have a higher command of them, since earnings from their own work should represent an easier resource to control (see, for a study on developing countries, Desai and Jain, 1994).

Even if there is a huge literature on the effects of mothers' schooling on children's schooling (see, for recent examples, Behrman and Rosenzweig, 2002; Plug, 2004), so far, economists have tended to study mothers' work and children's schooling in separate settings. However, these aspects are recognized to be interdependent, as recently emphasized by some international institutions (for example UNICEF, 2007 and 2008). Evidence drawn from a model where children's schooling and mothers' work are treated as interdependent phenomena has revealed that, in India, children of working mothers have a lower probability of schooling and a higher probability of working (Francavilla and Giannelli, 2007). That model, however, did not take into account unobserved heterogeneity, that is, the residual correlation among the outcomes of mothers and the outcomes of their children. In this paper, we develop a model to derive a measure of this correlation, since a significant value of this parameter would support the hypothesis of the joint nature of the time allocation decisions of mothers. Moreover, the sign and the size of this correlation would answer our question on the direction and magnitude of the relation between children's schooling and their mothers' work.

In order to extend the model to this possibility, this paper employs a multilevel random effects model to investigate the determinants of children's schooling and mothers' work in a joint framework. Under the hypothesis that children's time is an extension of their mother's time, the two observed outcomes are the working status of the mother and the schooling status of each one of her children. We assume that working and schooling statuses are determined by the two underlying mother's utilities for working and for sending each child to school. The covariates are distinguished into child-level and mother-level. We specify a two-equation linear model for these utilities under the assumptions that children of the same mother share the same mother-level error such that the child equation becomes a *random effects probit*. Also, the mother equation has an error structure that allows for correlation between the mother and child equations. We aim at estimating two parameters: the residual correlation of the utilities of working and sending each child to school

(mother-child correlation), and the residual correlation of the utilities of schooling among siblings of the same mother (within class correlation).

Controlling for covariates, among which wealth is the most powerful predictor, we find that the mother-child correlation is significant and negative. That is, if mothers work, children may contribute to housework or to household income instead of attending school. Moreover, a significant and quite large within class correlation suggests that, all things being equal, mothers tend to choose the same state for each child. Indeed, some gender discrimination is present (males have a higher probability of studying), but the size of this correlation is such that preferences for treating all children equally seem to dominate, especially in poor families. As to the mother's employment decision, our results show that mothers have a higher probability of working when they are illiterate, poor and have unskilled or unemployed partners. This evidence diverges sharply from what is generally found for developed countries.

The paper is structured as follows. Section 2 sketches the theoretical model, while Section 3 outlines the econometric model. Section 4 describes the data and the variables, while Section 5 is dedicated to the presentation and discussion of the results. Section 6 concludes.

2 A theoretical model of maternal allocation of time of the child

Since our focus is on the mother-child relation, our model is based on the following assumptions on the household decision making process. It is general wisdom, confirmed by a large amount of evidence drawn from data on the use of time, that school age children spend more time with their mothers than with their fathers (see, for example, Yeung, Sandberg, Davis-Kean and Hofferth, 2001). Exploiting this stylised fact, although parental decisions about the education of children have been modelled in many different ways,¹ we assume that the mother has the main opportunity to manage her children's time.² We therefore develop a model in which the time of the child is assumed to be an extension of the time of their mother, who decides how to allocate it. The father's behaviour is exogenous to the model; he always works in the market and gives the family a labour income Y . The family consumes a bundle of home produced goods (X_d) and market (X_m) goods, and the mother (M) has to allocate her time between hours of domestic activities, $H_{M,d}$, market work, $H_{M,m}$, and leisure, L . The time of the child is an extension of their mother's time, and the mother decides how to allocate it between study and work.

¹ See Bourguignon and Chiappori (1992) on the collective versus unitary approach to the analysis of household allocation of time.

² Even if, especially in some developing countries, decisions on schooling may be taken exclusively by fathers, the fact that fathers usually spend most of their time outside home, gives some support to our hypothesis.

The mother maximises her utility function $U(X, L, S)$,³ where X is the sum of home produced goods (X_d) and market goods (X_m), and S is hours of schooling of her child. S necessarily includes going to school, but also includes hours of all those extra activities (including play time or time spent studying with their mother) necessary to develop the capabilities of her child. The inclusion of S as an argument of mother's utility can have a double interpretation. First, an altruistic interpretation, according to which, the mother derives utility from the fact that her child attends school. The second interpretation is egoistic, since the mother may guarantee herself future consumption by means of the support of her child,⁴ investing in their education. This implies the reasonable assumption that the child's returns on human capital are higher if they accumulate at least some education instead of only work experience. Under these hypotheses the mother's utility to be maximized is simply:

$$\begin{aligned} \text{Max } U &= U(X, L, S) & (1) \\ \frac{\partial U}{\partial X} &> 0, \frac{\partial U}{\partial L} > 0, \frac{\partial U}{\partial S} > 0 \end{aligned}$$

The utility maximization is subject to the consumption constraint:

$$X = X_d + X_m \quad (2)$$

The time constraint, with the time of the child (T_C) considered like an extension of their mother's time (T_M), is the following:

$$T = T_M + T_C \quad (3)$$

$$T_M = H_{M,d} + H_{M,m} + L \quad (4)$$

$$T_C = H_{C,d} + H_{C,m} + S \quad (5)$$

where the time of the child can be allocated to domestic production ($H_{C,d}$), market work ($H_{C,m}$) and study (S). The mother and the child produce a homogeneous domestic good X_d according to the following household production constraint:

³ If the mother were altruistic and cared about the child's utility (in other words, if the child's utility was an argument of the mother's utility function) the qualitative results would remain unchanged.

⁴ For the theory of family constitutions, according to which parents invest in their children's education in order to guarantee themselves consumption when they are old, see Cigno (2006) and Cigno and Rosati (2000). This theory is especially applicable to developing countries, where social security is often insufficient.

$$X_d = f(H_{M,d}) + g(H_{C,d}) \tag{6}$$

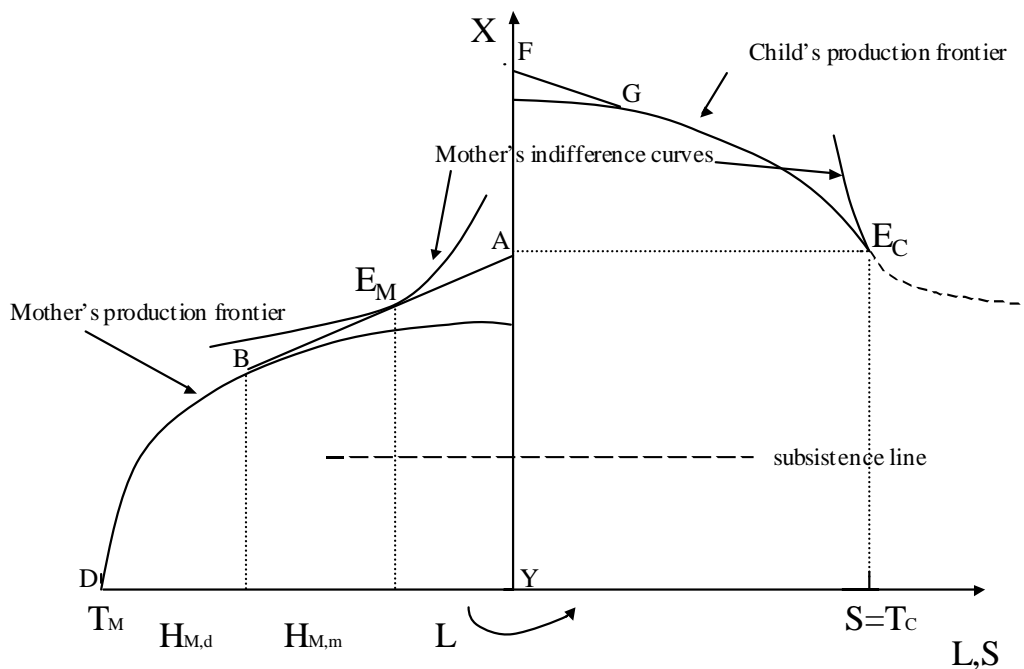
where $f(H_{M,d})$ and $g(H_{C,d})$ are the mother and the child household production function respectively and $f'(H_d) > g'(H_d)$, that is, the mother is more productive than her child. The budget constraint is:

$$X_m = W_M H_{Mm} + W_C H_{Cm} + Y \tag{7}$$

where W_M and W_C are mother's and child's wages respectively.

The graphical solution of this model is represented in Figure 1.

Figure 1 – Mother participates to the labour market and child studies all the time



The vertical X-axis divides T into two parts, T_M in the left panel with the household production function in the case of absence of the child in the household mother's production frontier, and T_C in the right panel with the child's production frontier. At the intersection of the X-axis with the L, the S-axis might be positive and equal to Y market goods (bought, for example, with the father's earnings). The dashed line intersects the subsistence consumption level of the family. All combinations of mother and child work could be represented by varying the level of wages, the form of the home production functions and the mother's preferences. In Figure 1 we show two of the possible combinations: one for the case in which the mother relies only on her time and the other for the case in which she considers her child's time as an extension of her time. In the first case, at the equilibrium point E_M , she works in the market, is engaged in home production and

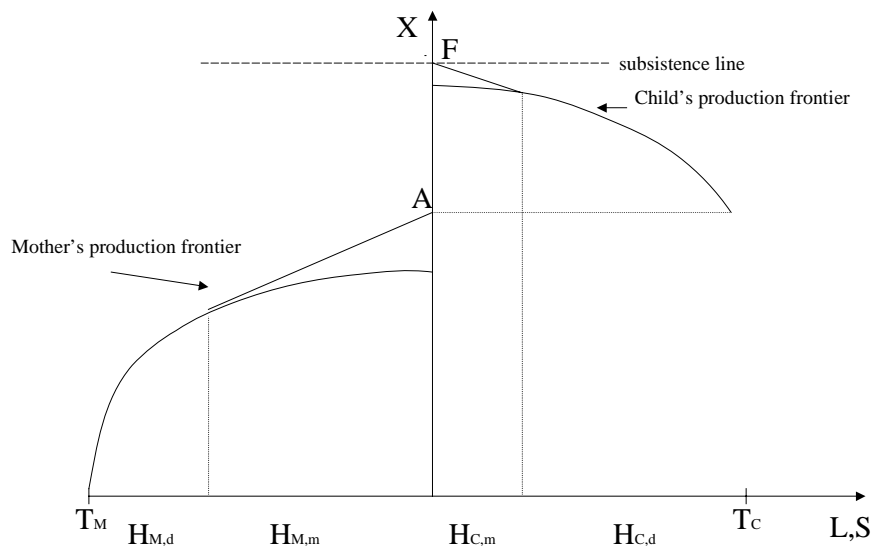
enjoys some leisure time (see the left panel). In the second case she can increase her utility at the equilibrium point E_C , where her child is engaged in full time education and $S=T_C$ (see the right panel).

Other cases could be represented. For example, if a child spends time on all kind of activities, including study, the child's equilibrium occurs at any point between F and G where $g'(H_{Cd})=W_C$. This allocation of the time of the child is jointly determined with their mother's equilibrium. The latter may occur either between A and B where the mother participates ($f'(H_{Md})=W_M$), or between B and D where the mother specializes in domestic work ($f'(H_{Md})>W_M$).

Alternatively, in a situation of poverty where all family members are forced to allocate all their time to work, the mother has no choice over the way she can use her time. She has to produce a given minimum amount of domestic goods for her and her family's survival. This activity will take all her time. If her time is not enough to achieve the household subsistence level, she will have to employ her child in domestic activities and/or send the child to the labour market.⁵

This special case in which the mother chooses to allocate the total time T (hers and her child) to work would take place at the frontier solutions A and F as shown in Figure 2.

Figure 2 - Subsistence: mother and child work all the time



⁵ This case is similar to the “Luxury axiom” (Basu and Van, 1998). This axiom states that: “A family will send the children to the labour market only if the family’s income from non-child labour sources drop very low.” (p. 416).

3 An econometric joint model for mother's work and children's schooling

We devise an econometric model reflecting the economic model outlined in the previous section. Let $j=1, \dots, J$ denote mothers and $i=1, \dots, n_j$ denote children aged 6-14 of mother j . The observed outcomes are the working status of the mother, $y_j^{(m)}$ (1=working, 0=otherwise), and the schooling status of each of her children aged 6-14, $y_{ij}^{(c)}$ (1=attending school, 0=otherwise). We assume that working and schooling statuses are determined by the *underlying utilities* of the mother:

$$\{y_j^{(m)} = 1\} \Leftrightarrow \{\tilde{y}_j^{(m)} > 0\} \quad \text{where } \tilde{y}_j^{(m)} \text{ is the utility of mother } j \text{ for working}$$

$$\{y_{ij}^{(c)} = 1\} \Leftrightarrow \{\tilde{y}_{ij}^{(c)} > 0\} \quad \text{where } \tilde{y}_{ij}^{(c)} \text{ is the utility of mother } j \text{ for sending her child } i \text{ to school}$$

The covariates determining the utilities are distinguished into child-level covariates \mathbf{z}_{ij} (child's age and gender) and mother-level covariates \mathbf{x}_j (every covariate that is constant for a mother, such as: mother's age and education, household structure, partner's occupation, household's wealth).

We assume that the joint model for the utilities has two linear equations:

$$\tilde{y}_j^{(m)} = \alpha^{(m)} + \boldsymbol{\beta}^{(m)} \mathbf{x}_j + u_j^{(m)} + e_j^{(m)} \quad (\text{mother equation}) \quad (8)$$

$$\tilde{y}_{ij}^{(c)} = \alpha^{(c)} + \boldsymbol{\beta}^{(c)} \mathbf{x}_j + \boldsymbol{\gamma}^{(c)} \mathbf{z}_{ij} + u_j^{(c)} + e_{ij}^{(c)} \quad (\text{child equation}) \quad (9)$$

with the following assumptions on the errors:

- 1) The u -errors $(u_j^{(m)}, u_j^{(c)})$ have a bivariate normal distribution with zero means and the following variances and covariance: $\text{Var}(u_j^{(m)}) = 1$, $\text{Var}(u_j^{(c)}) = \sigma_c^2$, $\text{Cov}(u_j^{(m)}, u_j^{(c)}) = \sigma_{mc}$. The error $u_j^{(m)}$ has a fixed variance to ensure identifiability. Note that the siblings share the same mother-level error $u_j^{(c)}$.
- 2) The residual e -errors $(e_j^{(m)}, e_{1j}^{(c)}, \dots, e_{n_j j}^{(c)})$ are independent and identically distributed with standard normal distribution, so $\text{Var}(e_j^{(m)}) = \text{Var}(e_{ij}^{(c)}) = 1$ and $\text{Cov}(e_j^{(m)}, e_{ij}^{(c)}) = \text{Cov}(e_{ij}^{(c)}, e_{i'j}^{(c)}) = 0$ ($i' \neq i$). The errors have a fixed variance to ensure identifiability. Note that the normal distribution of the e -errors corresponds to a *probit* model for the probabilities.
- 3) Every u -error is independent of any e -error.

The child equation is a random effects *probit*, since $u_j^{(c)}$ varies between mothers and $e_{ij}^{(c)}$ varies within mothers. Also the mother equation has an error structure with two terms, but it is *not* a random effects *probit* since both $u_j^{(m)}$ and $e_j^{(m)}$ vary between mothers: indeed, the mother equation could be written with a single error term $w_j^{(m)} = u_j^{(m)} + e_j^{(m)}$. Decomposing the error into two additive terms is just a trick to allow a correlation between the mother and child equations: in fact, the estimation methods for random effects models allow for correlated random effects and thus the introduction of the fictitious random effects $u_j^{(m)}$ is a simple way to fit correlated equations via standard software.⁶

Systems of random effects equations have been used to deal with endogenous covariates in multilevel settings (Cochrane and Guilkey, 1995; Degraff, Bilsborrow and Guilkey, 1997). In such cases the outcome of an equation appears as a covariate in another equation. Here we take a different approach: in fact, our theoretical model postulates that the mother jointly allocates her time and that of her children, so the working and schooling indicators have to be modelled jointly, but not regressed on each other. Therefore, our econometric model has a SUR structure (Seemingly Unrelated Regressions: e.g. Wooldridge, 2002), where the outcomes do not appear as covariates and the equations are correlated through the error terms.

To understand the properties of our econometric model it is essential to write down the model-implied residual variances and covariances of the utilities. The term “residual”, which means “after adjusting for the covariates”, is written explicitly in the following definitions, but it is left implicit later on. The *residual variances of the utilities* are

$$\text{Var}(\tilde{y}_j^{(m)} | \mathbf{x}_j) = \text{Var}(u_j^{(m)}) + \text{Var}(e_j^{(m)}) = 1 + 1 = 2$$

$$\text{Var}(\tilde{y}_{ij}^{(c)} | \mathbf{x}_j, \mathbf{z}_{ij}) = \text{Var}(u_j^{(c)}) + \text{Var}(e_{ij}^{(c)}) = \sigma_c^2 + 1$$

The *residual covariances/correlations of the utilities for any two siblings* are

$$\text{Cov}(\tilde{y}_{ij}^{(c)}, \tilde{y}_{i'j}^{(c)} | \mathbf{x}_j, \mathbf{z}_{ij}, \mathbf{z}_{i'j}) = \text{Cov}(u_j^{(c)}, u_j^{(c)}) = \text{Var}(u_j^{(c)}) = \sigma_c^2$$

⁶ A minor drawback due to the use of the fictitious random effects $u_j^{(m)}$ is the change in the scale of the mother equation since $\text{Var}(w_j^{(m)}) = \text{Var}(u_j^{(m)}) + \text{Var}(e_j^{(m)}) = 1 + 1 = 2$. Thus the mother equation is a *scaled probit*, i.e. a *probit* with a scale different from 1: in this case the scale factor is equal to $\sqrt{2}$, so the regression coefficients are $\sqrt{2}$ times the coefficients of an ordinary *probit*. Since a scaled *probit* is statistically equivalent to an ordinary *probit*, we divide the estimates by $\sqrt{2}$ to make them comparable to the results from an ordinary *probit*.

$$\text{Cor}\left(\tilde{y}_{ij}^{(c)}, \tilde{y}_{i'j}^{(c)} \mid \mathbf{x}_j, \mathbf{z}_{ij}, \mathbf{z}_{i'j}\right) = \frac{\sigma_c^2}{\sigma_c^2 + 1} \quad (10)$$

The residual covariances/correlations of the utilities for a mother with one of her children are

$$\begin{aligned} \text{Cov}\left(\tilde{y}_j^{(m)}, \tilde{y}_{ij}^{(c)} \mid \mathbf{x}_j, \mathbf{z}_{ij}\right) &= \text{Cov}\left(u_j^{(m)}, u_j^{(c)}\right) = \sigma_{mc} \\ \text{Cor}\left(\tilde{y}_j^{(m)}, \tilde{y}_{ij}^{(c)} \mid \mathbf{x}_j, \mathbf{z}_{ij}\right) &= \frac{\sigma_{mc}}{\sqrt{2(\sigma_c^2 + 1)}} \end{aligned} \quad (11)$$

Some relevant properties of the model are discussed in the following remarks:

- The interpretation of the variance-covariance parameters σ_c^2 and σ_{mc} is easier if they are transformed into correlations, namely the correlation of utilities among siblings (10) and the mother-child correlation (11).
- The mother-child correlation (11) can be positive or negative depending on σ_{mc} : a positive (negative) correlation means that mothers with a higher utility for working tend to have a higher (lower) utility for sending their children to school.
- The random effects of the child equation (9), $u_j^{(c)}$, summarize the effects of unobserved covariates at the mother level: the larger their variance σ_c^2 , the greater the role of unobserved covariates and thus the higher the correlation among siblings, also called Intraclass Correlation Coefficient (ICC), which is strictly positive unless $\sigma_c^2 = 0$.
- Apart from correlations (10) and (11), any other correlation among utilities is null (e.g. among two mothers, among two children of different mothers).
- Each slope has the usual interpretation in terms of change in the *probit* due to a unit increase in the corresponding covariate; however, the child equation (9) has random effects, so the slopes have a conditional meaning, i.e. they refer to the effect of the covariates *conditional* on the random effects $u_j^{(c)}$ (in general, conditional effects are larger than marginal effects).
- The mother-level covariates \mathbf{x}_j have different slopes for the effect on the utility for working, $\boldsymbol{\beta}^{(m)}$, and for sending children to school, $\boldsymbol{\beta}^{(c)}$.

To compute predicted probabilities we define a baseline mother and a baseline child by choosing a value for each covariate, $\mathbf{x}_j = \mathbf{x}_*$ and $\mathbf{z}_{ij} = \mathbf{z}_*$. The definition of the baseline mother also requires

the specification of the value of the u -errors (unobserved covariates), which we set to their mean, i.e. $u_j^{(m)} = 0$ and $u_j^{(c)} = 0$. Denoting with $\Phi(\cdot)$ the normal distribution function, the predicted probability that the baseline mother works is $\Phi(\alpha^{(m)} + \beta^{(m)}\mathbf{x}_*)$, while the predicted probability that the baseline child of the baseline mother attends school is $\Phi(\alpha^{(c)} + \beta^{(c)}\mathbf{x}_* + \gamma^{(c)}\mathbf{z}_*)$.

4 Data and variables

We draw our data from the National Family Health Survey (NFHS-2) 1998-1999, India (IIPS and ORC Macro, 2000).⁷ The NFHS-2 is a household survey with two different samples: a sample of around 92,500 households, who answered the *Household Questionnaire*, and a sample of around 90,300 married women in the age group 15–49 who are members of the household sample and who answered the *Woman's Questionnaire*. Data for the women comes from both questionnaires. The sample covers more than 99% of India's population living in all 26 Indian states. The sample size for each state was drawn separately for urban and rural areas proportionally to the size of the urban and rural populations of each state.

This survey is suitable to fulfil one of our objectives, that is, to analyse the schooling probability of children while controlling for the correlation that might arise from sharing the same home and, more importantly according to our theoretical model, the same mother. The data set has a multilevel structure, with children at level 1, mothers at level 2, households at level 3 and geographical areas at the upper level. This structure calls for multilevel, random effects models taking into account the correlations induced by the clustering.

Our analysis is based on two samples: a sample for urban areas made up of 14,181 mothers and their 26,269 children and a sample for rural areas of 33,137 mothers and their 65,726 children. We distinguish school-age children (6 to 14 years old) in two categories, *students* and *non-students*. The category of *students* includes not only full-time students, but also children who study and work either for the market (a very small proportion, 0.26% of all students) or for the family business (whose number cannot be computed from the survey, due to the structure of the questionnaire).⁸ On the other hand, the category of *non-students* includes those children who do not attend school at all, since they work full-time either outside (3%) or for the family business (3.3%) or they are *inactive*, namely children who are not in one of the preceding categories, but that can be assumed to do some work, most likely domestic (12% of the whole sample of children).

⁷ We thank MEASURE DHS for providing this survey for India.

⁸ At variance with most studies on child labour, we do not specify a category for work and study, in line with the focus of our paper, according to which this state is not entirely harmful if the children manage to reconcile study with work.

Mothers are classified either as *working* or *not working*. Precisely, a mother is defined to be working if she has worked in the last twelve months.⁹

The descriptive statistics reported in Table 1 show that in urban areas 91% of children study and 29% of mothers work, whereas in rural areas 80% of children study and 46% of mothers work.

The child-level covariates \mathbf{z}_{ij} for the child equation (9) are age and gender. The mother-level covariates \mathbf{x}_j for the mother equation (8) and for the child equation (9) include the number of members of the household (disaggregated by age group and distinguishing between siblings and other children), mother's education and age, partner's occupation, religion of the household head, household wealth,¹⁰ acres of land owned by the household, and five dummy variables for geographical areas. A careful control of household composition is important in India, where many families may cohabit in the same household.

A two-way table of the proportions of children attending school by age and number of siblings shows that the probability of studying decreases with age and also with the number of siblings present in the household. In the econometric model, we control for the interdependent effects of birth order and number of siblings by using the child's age and the number of siblings distinguished between the age groups 0-5 and 6-14.

The mother's education is described in Table 1 by a dummy variable, denoting mothers with at least one year of schooling (*Mother is literate*), and a quantitative variable for the number of years of schooling. Literate women amount to 69% in urban areas and 34% in rural areas.

5 Results

Fitting the joint model with the mother equation (8) and the child equation (9) is computationally heavy, so to select the best model specification we fitted the two equations separately.¹¹ Then we fitted the joint model with also the two equations simultaneously,¹² which also allows the

⁹ We prefer to use this variable, instead of one indicating if the woman is currently employed, since it is more suitable to study the woman's long-run attachment to the labour market in conjunction with the long-run decision of sending her children to school.

¹⁰ The indicator variables to be included in the wealth index are all household assets and utility services, including country-specific items. See Rutstein and Johnson (2004) for details.

¹¹ In searching for a flexible specification each numerical covariate was tested for quadratic effects, but for the sake of parsimony cubic effects were not considered.

¹² All the models are fitted using the maximum likelihood algorithms of Stata. The mother equation is a standard *probit*, so it can be fitted in a few seconds using the *probit* command. On the other hand, the child equation is a random intercept *probit* and it can be fitted in a few minutes using the *xtprobit* command. The computational burden of random effects models is due to the approximation of the integrals through adaptive Gaussian quadrature. Fitting the two equations simultaneously is not possible with standard Stata commands, but it is feasible with the user-written, freely downloadable *gllamm* command (Rabe-Hesketh, Skrondal and Pickles, 2004). Also *gllamm* is based on adaptive Gaussian quadrature, but it is computationally less efficient than *xtprobit*, so the task of fitting the joint model takes

estimation of the covariance σ_{mc} between the u -errors of the two equations and thus the mother-child correlation (11). The estimates for the two equations change only slightly when passing from separate to simultaneous fitting.

In this section we first show the results from the child equation, then the results from the mother equation and finally we discuss the findings of the mother-child correlation.

We compute the predicted probabilities that a child attends school and a mother works using the following definitions of baseline child and baseline mother (see Section 3):

- *Baseline child*: female aged 13.
- *Baseline mother*: aged 34, illiterate, with 2 children aged 6-14 and no child aged 0-5, whose partner is unemployed or unskilled, living in a household in the South with the following features: the household has a single family, the head is Hindu, the household has no acres of own land and the wealth is at the first quartile of the area (0.1902 for urban and -1.0224 for rural). Moreover, the baseline mother has a mean value on the unobserved covariates, namely $u_j^{(m)} = 0$ and $u_j^{(c)} = 0$.

several days. The `gllamm` command is designed to fit Generalized Linear Latent and Mixed Models (GLLAMM: Skrondal and Rabe-Hesketh, 2004), a broad class that encompasses also our bivariate two-level model.

Table 1 - Definition of the variables and descriptive statistics

<i>Mother-level variables</i>	<i>Urban (n=14181)</i>				<i>Rural (n=33137)</i>			
<i>Name</i>	<i>Mean</i>	<i>S. Dev.</i>	<i>Min</i>	<i>Max</i>	<i>Mean</i>	<i>S. Dev.</i>	<i>Min</i>	<i>Max</i>
Mother is working	0.287	0.453	0	1	0.464	0.499	0	1
Number of children aged 0-5	0.526	0.772	0	6	0.765	0.900	0	7
Number of children aged 6-14	1.857	0.929	1	7	1.993	0.992	1	7
Presence of children aged 0-5 of other mothers	0.117	0.322	0	1	0.156	0.362	0	1
Number of household members over 14	2.801	2.210	0	23	2.845	2.276	0	26
Mother is literate	0.690	0.462	0	1	0.338	0.473	0	1
Years of mother's education	6.127	5.223	0	21	2.144	3.493	0	22
Mother's age	34.054	5.959	16	49	33.160	6.506	15	49
Head of household is Muslim	0.163	0.369	0	1	0.116	0.320	0	1
Head of household is Christian	0.063	0.243	0	1	0.053	0.225	0	1
Head of household in scheduled caste or tribe	0.216	0.412	0	1	0.334	0.472	0	1
Household wealth index	0.837	0.916	-1.419	2.789	-0.436	0.734	-1.529	2.713
Acres of land owned by the household	1.520	11.767	0	99.990	0.949	7.903	0	99.990
Partner's job: clerical or professional	0.227	0.419	0	1	0.077	0.267	0	1
Partner's job: sales	0.182	0.386	0	1	0.065	0.246	0	1
Partner's job: skilled manual	0.299	0.458	0	1	0.162	0.369	0	1
North	0.281	0.449	0	1	0.219	0.414	0	1
Central	0.140	0.347	0	1	0.209	0.406	0	1
East	0.123	0.328	0	1	0.201	0.401	0	1
Northwest	0.100	0.300	0	1	0.140	0.347	0	1
West	0.178	0.383	0	1	0.079	0.269	0	1
<i>Child-level variables</i>	<i>Urban (n=26269)</i>				<i>Rural (n=65726)</i>			
<i>Name</i>	<i>Mean</i>	<i>Std. Dev.</i>	<i>Min</i>	<i>Max</i>	<i>Mean</i>	<i>Std. Dev.</i>	<i>Min</i>	<i>Max</i>
Child is attending school	0.908	0.288	0	1	0.795	0.404	0	1
Child's age	9.965	2.561	6	14	9.740	2.549	6	14
Child is male	0.520	0.500	0	1	0.518	0.500	0	1

5.1 The child equation: does the mother care about all her children in the same way?

The random effects *probit* model of equation (9) allows the detection of the factors affecting the schooling status of children and to quantify the residual correlation among siblings. The model has two levels (mother level and child level) even if the phenomenon has further levels above the mother, such as the household and the region levels. Including random effects for higher levels is conceptually simple, but computationally prohibitive. To check that neglecting higher levels is not harmful, we fitted the two-level model and computed robust standard errors with households as top-level clusters (Skrondal and Rabe-Hesketh, 2004). This is a way to assess how the standard errors are influenced by the correlation among children of different mothers living in the same household. Since the robust standard errors are only slightly bigger than the classical ones, the two-level specification seems to suffice.

Our economic model postulates that such correlation is due to sharing the same utility-maximizing mother. Indeed, in the child equation (9) all the children of a mother share the same mother-level error $u_j^{(c)}$. If the residual correlation between the utilities for any two children (10) is high, it means that, after controlling for the explanatory variables, mothers tend to treat their children in the same way and the probabilities of siblings to attend school are markedly pushed up or down by their mother's utility. Maximum likelihood estimates of the child equation (9) are presented in Table 2.

For each area, the first column reports the estimate of the slope, while the second column reports the predicted probability for a hypothetical subject differing from the baseline for a unit increase in the covariate under consideration. For example, the heading of the second column of the urban area informs us that in urban areas the baseline child of a baseline mother has a predicted probability of 94.6% of attending school, while the value corresponding to the covariate *Child is male* informs us that, if the baseline is modified by switching sex from female to male, then the predicted probability becomes 96.8%. For numerical covariates a unit increase from the baseline is considered and when a quadratic term is present the predicted probability is reported only in the row corresponding to the quadratic term: for example, 84.7% is the predicted probability of attending school obtained if *Child's age* is changed from 13 to 14 taking into account both the linear and the quadratic effect.

Table 2 – Estimates for the child equation (9): a random effects *probit* model of the probability that the child attends school

<i>Covariates</i>	<i>Urban</i>		<i>Rural</i>	
	<i>Estimate</i>	<i>Prob. (base= 94.6%)</i>	<i>Estimate</i>	<i>Prob. (base= 56.2%)</i>
Child's age: linear	1.199		0.989	
Child's age: quadratic	-0.066	84.7%	-0.055	36.7%
Child is male	0.247	96.8%	0.749	81.7%
Number of siblings aged 0-5: linear	-0.322		-0.274	
Number of siblings aged 0-5: quadratic	0.030	90.6%	0.027	46.4%
Number of siblings aged 6-14: linear	0.056 *		0.096	
Number of siblings aged 6-14: quadratic	-0.038	93.0%	-0.037	52.7%
Presence of children aged 0-5 of other mothers	-0.175 *	92.4%	-0.243	46.5%
Number of household members over 14	0.011 *	94.7%	0.041	57.8%
Mother's schooling: 5 years of education	0.486	99.2%	0.751	88.0%
Mother's schooling: 6 years of education	0.062	99.3%	0.054	89.1%
Mother's age	-0.012	94.5%	-0.018	55.5%
Head of household is: Muslim	-0.607	84.1%	-0.575	33.8%
Head of household is: Christian	0.153 *	96.1%	0.243	65.5%
Head of household in scheduled caste or tribe	-0.104 *	93.4%	-0.275	45.3%
Household wealth index: linear	1.057		0.958	
Household wealth index: quadratic	-0.107	99.4%	-0.260	91.7%
Acres of land owned by the household: linear	0.590		-0.046	
Acres of land owned by the household: quadratic	-0.006	98.6%	0.001	54.4%
Partner's job: clerical or professional	0.602	98.6%	0.578	76.8%
Partner's job: sales	0.246	96.8%	0.156	62.2%
Partner's job: skilled manual	0.145	96.0%	0.099	60.1%
Region: North	-0.300	90.4%	0.048 *	58.1%
Region: Central	-0.211	91.9%	0.065 *	58.7%
Region: East	-0.221	91.7%	-0.104	52.1%
Region: Northwest	0.337	97.4%	0.345	69.2%
Region: West	-0.145 *	92.8%	-0.047 *	54.3%
Constant	-2.575		-1.587	
σ_c	1.445		1.260	
Intraclass Correlation Coefficient	0.676		0.614	
Unobserved factors at 10th percentile ($u_j^{(c)} = -1.282\hat{\sigma}_c$)	-1.282	40.3%	-1.282	7.2%
Unobserved factors at 25th percentile ($u_j^{(c)} = -0.674\hat{\sigma}_c$)	-0.674	73.6%	-0.674	24.4%
Unobserved factors at 75th percentile ($u_j^{(c)} = +0.674\hat{\sigma}_c$)	0.674	99.5%	0.674	84.3%
Unobserved factors at 90th percentile ($u_j^{(c)} = +1.282\hat{\sigma}_c$)	1.282	100.0%	1.282	96.2%

*Not significant at the 5% level.

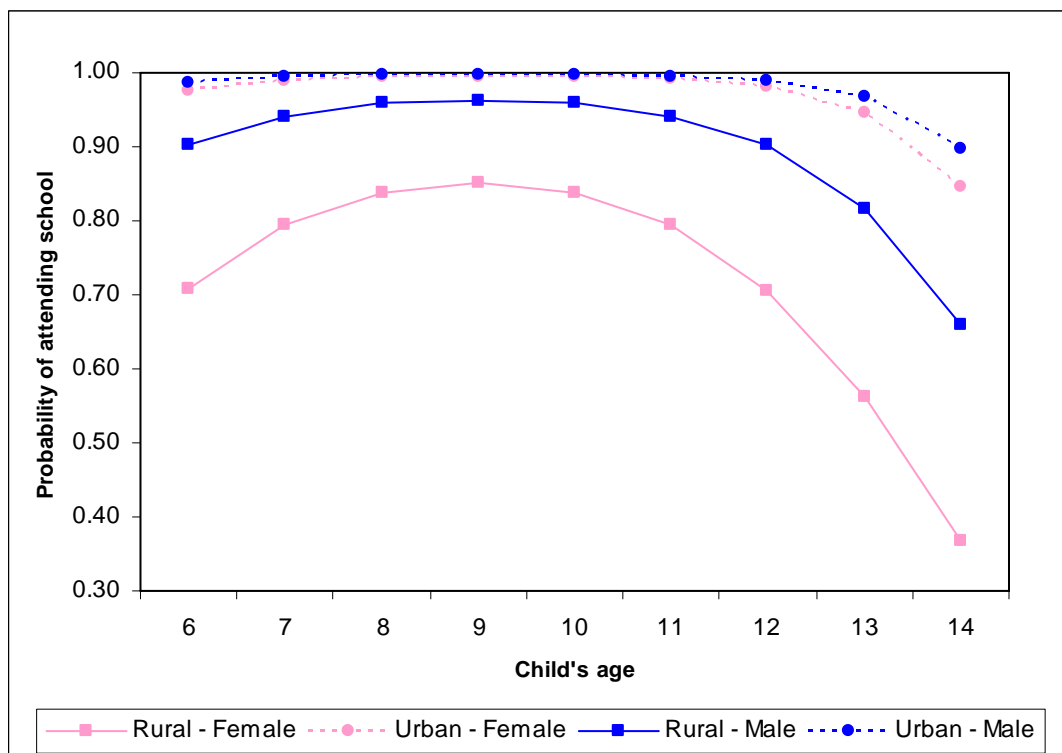
The effect of the mother’s education is modelled through a dummy variable (*Mother is literate*) and a numeric variable (*Years of mother’s education*): since switching the dummy while keeping the numeric at zero is meaningless, the predicted probability corresponding to *Mother is literate* is computed for a mother who is literate with 5 years of education. Then the predicted probability corresponding to *Years of mother’s education* is computed for a mother with one further year of education, i.e. literate with 6 years of education.

Starting with child-level covariates, we find that the child’s age has a significant quadratic effect and males have a higher probability of studying. To appreciate the role of age and gender it is important to see how they affect the predicted probability of attending school, keeping all the mother-level covariates at the baseline value, as in Figure 3.

The probabilities are very high and almost constant for ages 7 to 11. The lower values at age 6 are likely to be due to delayed entry or imperfections in age recording, while the decay starting at age 12 reflects school drop-out. The gender gap is modest in urban areas and relevant in rural areas, especially for ages 12 to 14.

The household structure has an important role for the probability of studying. Larger numbers of siblings aged 0 to 5 are associated with lower probabilities of attending school, even if the quadratic term implies a decreasing marginal effect of additional siblings.

Figure 3 – Predicted probability of attending school on child’s age, by area and gender



This effect is similar in urban and rural areas, with a slightly larger value in the former. On the contrary, the other aspects of household composition are markedly different in the two areas. The number of siblings aged 6 to 14, the presence of children of other mothers in the household and the number of all other members over 14 do not have significant effects in urban areas, whereas in rural areas these effects are significant. In particular, in rural areas, siblings aged 6 to 14 reduce the probability of studying by a small amount (but the quadratic term implies a decreasing marginal effect of additional siblings), the presence of children aged 0 to 5 of other mothers reduces the probability by the same amount of siblings of the same age, whereas the presence of other members in the household helps to improve the child's chances of going to school. These results suggest that in rural areas there is a sort of "pooling effect" of members of different families in the same household. For example, the number of all small coresident children reduces the probability of schooling of any school-age child present in the household, as if a small child absorbed time and income resources of the household as a whole irrespective of their own mother. In the same line of interpretation, any adult member may contribute by offering time and income to the household, thus securing better conditions for school-age children. This "pooling" effect is less likely in urban areas, where the provision of services for the family may be external to the household.

As for the mother's features, education has a crucial role, mostly in rural areas. An illiterate mother is detrimental for the schooling chances of her children and the higher the number of years of education of the mother, the higher the probability that her children attend school, thus confirming a well established result in the literature. The mother's age has a small negative effect, which we attribute to a cohort effect.

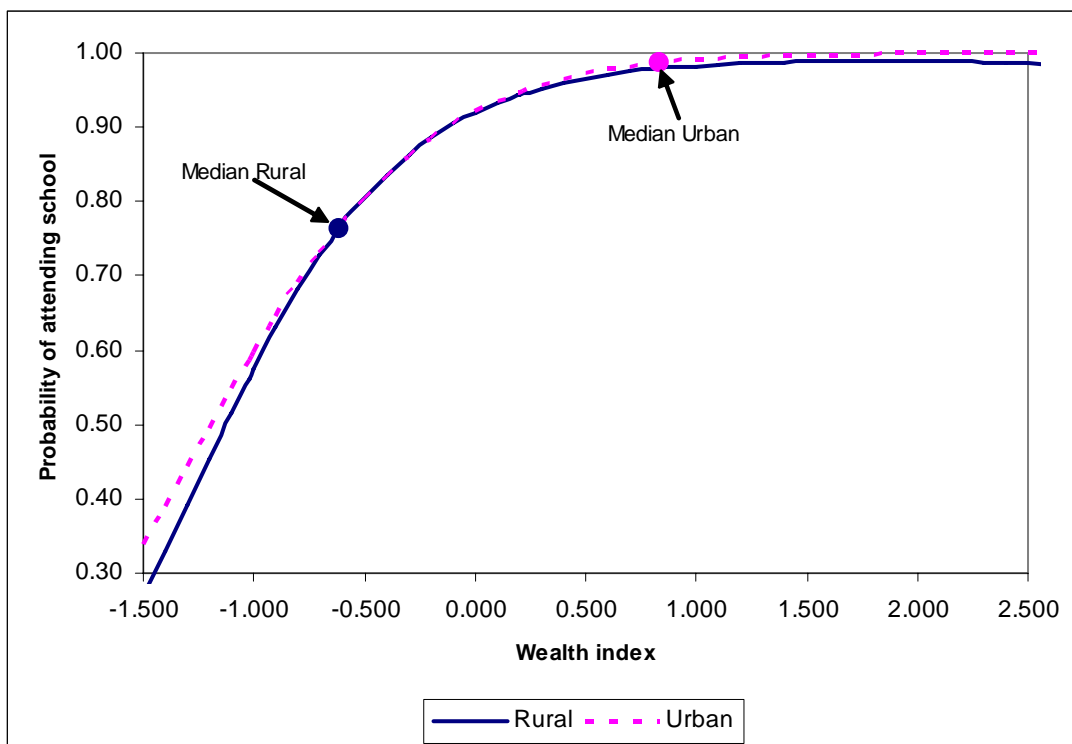
As to household characteristics, the religion of the head is relevant: compared to Hindu, the probability of attending school is lower for Muslim and higher for Christian. Also being in a scheduled caste or tribe proves to be negative for children's opportunities, especially in rural areas.

Household wealth is a very strong predictor that affects the probability of schooling in a quadratic way. In both areas the marginal effect on the probability is positive and decreasing, so a given difference in wealth is very important for poor families and negligible for rich families. Figure 4 shows the plot of the predicted probability of attending school against values of the wealth index in the observed range, when the other covariates are at baseline values. Children's schooling is strongly influenced by wealth with a plateau for values of the index above 1. The curves for urban and rural areas are very close: therefore, all things being equal (in particular wealth), the chance of attending school is similar in the two areas. However, the distribution of wealth is markedly different in the two areas, as pointed out by the median value highlighted in the picture:

this fact explains the large gap in the sample proportions of children attending school in urban and rural areas.

Acres of land owned by the household have an opposite effect in urban and rural areas. In urban areas they capture a pure wealth effect, that is, land ownership increases the child's probability of schooling. In rural areas the effect is peculiar: a few acres of land property decrease the probability of schooling, since children are expected to engage in the family agricultural activities, but as the number of acres of land owned rises, the effect tends to become a pure wealth effect, thus increasing the probability of studying (the fitted parabola has a minimum at 23 acres).

Figure 4 - Predicted probability of attending school on household wealth index, by area



Turning to the professional position of the partner of the mother, we find that partners in higher positions increase the probability of children attending school, since skilled workers, salesmen and, especially, clerical/professional workers have a significant and positive effect as compared to the unskilled workers or the unemployed.

Let us now turn to discuss unobserved heterogeneity. The random effects $u_j^{(c)}$ represent unobserved factors at the mother level. Their standard deviation σ_c is estimated to be significant and very high: 1.445 in urban areas and 1.260 in rural areas. Thus an increase of one in the value of the standard deviation of the unobserved factors at mother level is associated with an increase of 1.445 and 1.260, in urban and rural areas respectively, in the probability of sending children to

school. This effect is larger than any other observed covariate effect. It is instructive to consider some scenarios by computing the predicted probability of attending school for a few values of $u_j^{(c)}$: since the random effects have a normal distribution with zero mean and standard deviation σ_c , interesting values are $u_j^{(c)} = k\sigma_c$ for k taken at the 10th percentile (-1.282), at the 25th percentile (-0.674), at the 50th percentile (0), at the 75th percentile (+0.674), and at the 90th percentile (+1.282) of the distribution. The predicted probability for $u_j^{(c)} = 0$ is just the baseline reported in the headings of Table 2, while the predicted probabilities for the other values of k are reported in the last four rows of Table 2. If the mother has a high utility for sending her children to school due to a higher value of $u_j^{(c)}$, it is almost certain that her children actually attend school (100% in urban areas and 96.2% in rural areas, when the covariates are at baseline values). Conversely, if the mother has a low utility for sending her children to school due to a lower value of $u_j^{(c)}$, it is unlikely that they actually attend school (40.3% in urban areas and 7.2% in rural areas, when the covariates are at baseline values). Therefore, in this analysis unobserved heterogeneity plays a substantial role.

The standard deviation of the random effects can be converted into the ICC (10) among the mother's utilities of sending her children to school, yielding 0.68 for urban areas and 0.61 for rural areas. The size of the correlation confirms that the mothers' tendency to treat all children the same way dominates other observed effects, like discriminating among children according to their age or sex.

In order to check if the ICC is sensitive to the household's wealth, we fitted the model on two sub-samples defined by the bottom and top deciles of the wealth index, reporting the results in Table 3.¹³

Table 3 – ICC for the child equation (9): estimates in the full sample and in sub-samples of wealth deciles (sample size in parenthesis)

<i>Full Sample</i>		<i>Bottom wealth decile</i>		<i>Top wealth decile</i>	
<i>Urban</i> (n=26269)	<i>Rural</i> (n=65726)	<i>Urban</i> (n=4229)	<i>Rural</i> (n=10237)	<i>Urban</i> (n=2845)	<i>Rural</i> (n=9206)
0.68	0.61	0.66	0.61	NA ⁺	0.32*

*Not significant at the 5% level. ⁺Estimation algorithm did not converge

¹³ The sample sizes reported in the table are numbers of children, while the selection of the sub-samples is based on a mother-level covariate, so the decile sub-samples need not be one tenth of the full sample. Indeed, both sub-samples are larger than one tenth of the full sample, since the mothers in the poorest and richest households tend to have more children.

Interestingly, the ICCs of the bottom decile are nearly the same as in the full sample, whereas in the top decile they drop sharply. This might indicate that a higher wealth gives the mothers more freedom to choose the schooling status according to the child's observed characteristics. This is even more so in rural areas, where the ICC of the top decile shows the lowest value and is not significant (in urban areas it could not be estimated due to a very low variation in the outcome).

5.2 The mother equation: what determines the mother's decision to work?

Maximum likelihood estimates of equation (8) are presented in Table 4. The child-level covariates are obviously not usable. We tried all the mother-level covariates considered for the child equation, but we found fewer significant effects, so the final specification is simpler.

Table 4 - Estimates for the mother equation (8): the probability that the mother works

<i>Covariates</i>	<i>Urban</i>		<i>Rural</i>	
	<i>Estimate</i>	<i>Prob.</i> (base= 51.3%)	<i>Estimate</i>	<i>Prob.</i> (base= 81.5%)
Number of children aged 0-5	-0.150	45.3%	-0.107	78.6%
Number of household members over 14	-0.025	50.3%	-0.023	80.9%
Mother's schooling: 5 years of education	-0.541	43.7%	-0.331	71.8%
Mother's schooling: 6 years of education	0.070	46.5%	0.002 *	71.9%
Mother's age	0.020	52.1%	0.000 *	81.5%
Head of household is: Muslim	-0.149	45.4%	-0.246	74.3%
Head of household is: Christian	0.368	65.6%	0.495	91.8%
Head of household in scheduled caste or tribe	0.152	57.3%	0.305	88.6%
Household wealth index: linear	-0.535		-0.376	
Household wealth index: quadratic	0.090	35.3%	0.086	66.7%
Acres of land owned by the household: linear	-0.020 *		0.063	
Acres of land owned by the household: quadratic	0.000 *	50.5%	-0.001	83.2%
Partner's job: clerical or professional	-0.254	41.2%	-0.099	78.8%
Partner's job: sales	-0.415	35.1%	-0.234	74.7%
Partner's job: skilled manual	-0.328	38.4%	-0.137	77.7%
Region: North	-0.044 *	49.5%	-0.669	59.1%
Region: Central	-0.149	45.4%	-0.613	61.2%
Region: East	-0.441	34.1%	-1.072	43.1%
Region: Northwest	0.019 *	52.0%	-0.722	57.0%
Region: West	0.137	56.7%	0.266	87.8%
Constant	-0.640		0.424	

*Not significant at the 5% level

Own children aged 0 to 5 reduce the probability of working in both areas, but more in urban ones. Own children aged 6 to 14 and children of other mothers present in the household do not have significant effects, thus they are excluded from the final specification. The number of members aged 14 and over has a negative effect, as if their work would substitute that of mothers.

Literate mothers have a lower probability of working in rural areas, independently from the years of education. On the contrary, in urban areas the effect depends on the years of education: compared with illiterate mothers, the probability of working is lower for mothers with few years of education and higher for mothers with several years of education (for the baseline mother, which is illiterate, the predicted probability of working is 51.3%, that becomes 43.7% at 5 years of education, 46.5% at 6 years, 52.0% at 8 years, 57.6% at 10 years).

As shown also in other studies (see, for example, Giannelli and Francavilla, 2007), these results reflect the problem of poor job opportunities for women. The majority of occupations held by women are generally low paid and unskilled, so that only women in a severe state of necessity would accept them. In this light, it is easier to understand the negative association between work and literacy. On the other hand, in urban areas job opportunities for women are more likely to include higher quality jobs, so the positive role of education in improving women's autonomy through market labour is recovered.

Consistently with the standard model of women female participation in developed countries, age has a positive effect, even if it is significant only in urban areas. Moreover, differences due to religion and caste/tribe are notable, especially in rural areas.

As to wealth effects, the coefficients on a quadratic specification of the wealth index show that wealthier mothers have a lower probability of working. The marginal effects are decreasing (like for the probability of a child to study) and stronger in urban areas. Property of land has a significant effect only in rural areas, where the probability of working increases up to nearly 50 acres of land, and then declines.

The partner's professional position has a sound role, especially in urban areas. The position of salesmen seems to have the largest disincentive effect for women's work.

5.3 The correlation among child and mother equations: are maternal work and child schooling interdependent?

As for the relation between mother's work and children's schooling, previous evidence has shown that children of mothers who do not work have a higher probability of attending school than children of mothers who do work. Moreover, when mothers work, children are more likely to work

or to be inactive, this result pointing to a certain degree of complementarity between the work of mothers and children (see Francavilla and Giannelli, 2007). This outcome would represent a problem for policy, since the work of mothers, through its empowering function, should be associated with a better condition of children. Instead, the “perverse” effect of complementarity would suggest the need for policies directed at improving the working conditions of women. Thus, estimating this relation has relevant implications for policy choices.

Francavilla and Giannelli (2007) present a reference estimate of the probability of child work conditional on mother’s work showing a positive and significant coefficient. However, the estimated coefficient may be biased, since the covariate on mother’s work is likely to be endogenous. To circumvent this problem, the authors fitted a multinomial logit model of mother-child states, thus assuming, without testing, the joint nature of these decisions.

The model defined by the mother equation (8) and the child equation (9) explicitly recognizes the joint nature of the working and schooling decisions. Fitting the two equations simultaneously allows us to estimate the covariance σ_{mc} between the u -errors of the two equations and thus to test more properly the mother-child correlation (11).

In urban areas the covariance σ_{mc} is significant and estimated as -0.2805 , yielding a residual correlation of -0.11 between the utilities for working and sending children to school, after controlling for the observed covariates. The relationship is slightly stronger in rural areas, with a significant covariance of -0.3948 and a mother-child residual correlation of -0.18 . The presence of a significant mother-child correlation supports the interdependence hypothesis and the consequent choice of a joint model. Moreover, the mother-child correlation is negative (that is, if the mother works the child is less likely to attend school), thus confirming the previous evidence on the complementarity between maternal work and child labour (or between maternal work and child “inactivity”).

At first sight, the estimated mother-child correlation seems modest in both areas. However, such a correlation concerns latent variables, while its impact on observed responses is substantial. To clarify this point, note that the bivariate normal distribution of the u -errors implies $E(u_j^{(c)} | u_j^{(m)}) = \sigma_{mc} u_j^{(m)}$, so if a mother in a rural area ($\hat{\sigma}_{mc} = -0.395$) is at the third quartile of the unobserved factors determining the working status ($u_j^{(m)} = 0.674$), that is if this mother has a high propensity to work, then the mean value of the unobserved factors determining the schooling status of one of her children is $E(u_j^{(c)} | u_j^{(m)}) = -0.395 \times 0.674 = -0.266$. Such a shift makes the predicted probability for the baseline child decrease from 56.2% to 45.6%. Taking the 90th percentile

($u_j^{(m)}=1.282$) the predicted probability goes down to 36.3%. Specular changes of similar magnitude are obtained by considering mothers with low values of $u_j^{(m)}$.

Like for the ICC, in order to check if these results are sensitive to the household's wealth, we fitted the model on two sub-samples defined by the bottom and top deciles of the wealth index, reporting the results in Table 5.

Table 5 – Correlation between mother's work and child schooling: estimates in the full sample and in sub-samples of wealth deciles (sample size in parenthesis)

<i>Full Sample</i>		<i>Bottom wealth decile</i>		<i>Top wealth decile</i>	
<i>Urban</i> (<i>n=26269</i>)	<i>Rural</i> (<i>n=65726</i>)	<i>Urban</i> (<i>n=4229</i>)	<i>Rural</i> (<i>n=10237</i>)	<i>Urban</i> (<i>n=2845</i>)	<i>Rural</i> (<i>n=9206</i>)
-0.11	-0.18	-0.19	-0.14	NA ⁺	-0.01*

*Not significant at the 5% level. ⁺Estimation algorithm did not converge

We use these correlation coefficients (Table 5, col. 3 and 4) to calculate the probability of schooling of children living in households at the bottom deciles of the wealth scale. In these cases, if mothers have a high propensity to work (90th decile of unobserved factors) the probability of child schooling drops, from a baseline value of 56% in rural areas, to 40% and to 33% in rural and urban areas respectively (note, for rural areas, that the coefficient reported in col. 4 is higher in absolute value than the coefficient in col. 2, so the drop in the bottom decile sample is smaller than in the full sample. Instead, the drop in the schooling probability is larger for poorer households living in urban areas). At the top of the wealth scale, this is no more so, since the coefficient turns out to be very small and insignificant in rural areas, and not even derivable in urban areas, where all children are very likely to go to school.

6 Final remarks

Our theoretical model postulates that the mother jointly allocates her time and that of her children. Her optimal decision depends on the level of wages, on the form of the household production function and on her preferences. In principle, all combinations of mother's work and child's activity may occur. Our empirical model supplies evidence about the conditions under which each combination of mother's and child's statuses has a higher probability of being realized.

We find that when the mother has a high utility for working, her child has a higher probability of being involved in labour and domestic activities. After controlling for mother's, child's and

household's characteristics, a higher mother's preference for working is negatively correlated with her child's school attendance. This worrying outcome cannot be ignored and calls for additional investigation. Heterogeneity exists among wealth deciles and geographical areas. We find that, in urban areas poor mothers with a high propensity to work have a very low probability of sending their children to school, even lower than that of equally poor families in rural areas. Instead, at the top of the wealth scale, almost all children are very likely to go to school.

Moreover, the high residual correlation between the utilities for any two children suggests that mothers tend to treat their children in the same way. Also, the attitude of the mother to discriminate among her children differs by areas and wealth deciles. While the intra class correlation coefficients of the bottom deciles are nearly the same as those of the full sample, in the top deciles they drop sharply, indicating that richer mothers have a higher opportunity to choose the schooling status according to their children's observed characteristics.

As to the mother's employment decision, our results show that mothers have a higher probability of working when they are illiterate, poor and have unskilled or unemployed partners. This evidence diverges sharply from what is generally found for developed countries and brings concern for the condition of women in the labour market. The negative association between the employment of mothers and schooling of their children reinforces this concern. An implication of our findings is that any policy aiming both at enhancing women's empowerment through labour and increasing children's welfare should also target improvements in women's conditions in the labour market. However, since unobserved heterogeneity plays a substantial role in our results, further research with more detailed data on the environment in which the households live, such as the presence and quality of schools and local labour market conditions, is needed in order to recommend adequate policy measures.

References

- Basu, K., P. H. Van (1998). "The Economics of Child Labor", *The American Economic Review*, 88(3), pp. 412-427.
- Behrman, J. R., M. R. Rosenzweig (2002), "Does Increasing Women's Schooling Raise the Schooling of the Next Generation?", *The American Economic Review*, Vol. 92, No. 1., pp. 323-334.
- Cigno, A. (2006), "A Constitutional Theory of the Family", *Journal of Population Economics*, 19, pp. 259-283.
- Cigno, A., F.C. Rosati (2000), "Mutual Interest, Self-enforcing Constitutions and Apparent Generosity". In L. A. Ge´rard-Varet, S. C. Kolm, J. M. Ythier (Eds.), *The Economics of Reciprocity, Giving and Altruism*. London and New York: MacMillan and St Martin's Press.

- Cochrane, S.H., D.K. Guilkey (1995). "The Effects of Fertility Intentions and Access to Services on Contraceptive Use in Tunisia". *Economic Development and Cultural Change*, 43, pp. 779-804.
- Degraff, D.S., R.E. Bilborrow, D.K. Guilkey (1997), "Community-Level Determinants of Contraceptive Use in the Philippines: A Structural Analysis". *Demography*, 34, pp. 385-398.
- Desai, S., D. Jain (1994), "Maternal Employment and Changes in Family Dynamics: the Social Context of Women's Work in Rural India", *Population and Development Review*, 115-136.
- Francavilla, F., G.C. Giannelli G. C. (2007), "The Relation between Child Labour and Mother's Work: the case of India" IZA DP 3099, Bonn.
- Giannelli, G.C., F. Francavilla (2007), "Do family Planning Programmes Help Womens' Employment? The Case of Indian Mothers", IZA DP 2762, Bonn.
- IIPS, International Institute for Population Sciences, ORC Macro (2000). "National Family Health Survey (NFHS-2)", 1998-99 India. Mumbai: IIPS.
- Plug, E. (2004), "Estimating the Effect of Mother's Schooling on Children's Schooling Using a Sample of Adoptees", *The American Economic Review*, Vol. 94, No. 1., pp. 358-368.
- Rutstein, S.H., K. Johnson (2004), "DHS Comparative Reports. The DHS Wealth Index", No. 6, ORC Macro Calverton, Maryland USA.
- Rabe-Hesketh, S., A. Skrondal, , A. Pickles (2004), "Gllamm manual". U.C. Berkeley Division of Biostatistics, Working Paper Series 160.
- Skrondal, A., S. Rabe-Hesketh (2004), *Generalised Latent Variable Modelling*. CRC/Chapman Hall.
- Yeung, J., F.W. Sandberg, P. E. Davis-Kean, S.L. Hofferth (2001), "Children's Time with Fathers in Intact Families", *Journal of Marriage and the Family*, Vol. 63, No. 1., pp. 136-154.
- UNICEF, (2007), *The State of the World's Children. Women and Children*, UN New York.
- UNICEF, (2008), *The State of the World's Children. Child Survival*, UN New York.
- Wooldridge, J.M. (2002), *Econometric Analysis of Cross Section and Panel Data*, MIT Press, Cambridge, MA.