



Determinants of child labor and school attendance: the role of household unobservables

P. Deb
F. C. Rosati

December 2002

Determinants of child labor and school attendance: the role of household unobservables

P. Deb*

F. C. Rosati**

**Working Paper
December 2002**

Understanding Children's Work (UCW) Project
University of Rome "Tor Vergata"
Faculty of Economics
V. Columbia 2
00133 Rome Tor Vergata

Tel: +39 06.7259.5618

Fax: +39 06.2020.687

Email: info@ucw-project.org

As part of broader efforts toward durable solutions to child labor, the International Labour Organization (ILO), the United Nations Children's Fund (UNICEF), and the World Bank initiated the interagency Understanding Children's Work (UCW) project in December 2000. The project is guided by the Oslo Agenda for Action, which laid out the priorities for the international community in the fight against child labor. Through a variety of data collection, research, and assessment activities, the UCW project is broadly directed toward improving understanding of child labor, its causes and effects, how it can be measured, and effective policies for addressing it. For further information, see the project website at www.ucw-project.org.

This paper is part of the research carried out within UCW (Understanding Children's Work), a joint ILO, World Bank and UNICEF project. The views expressed here are those of the authors' and should not be attributed to the ILO, the World Bank, UNICEF or any of these agencies' member countries.

* Indiana University-Purdue University Indianapolis, Hunter College of New York

** UCW-Project and University of Rome "Tor Vergata"

Determinants of child labor and school attendance: the role of household unobservables

**Working Paper
December 2002**

ABSTRACT

We develop a semi-parametric latent class random effects multinomial logit model to distinguish between observed and unobserved household characteristics as determinants of child labor, school attendance and idleness. We find that much of the substitution between activities as a response to changes in covariates is between attending school and being idle, with work being rather resistant. Unobserved household heterogeneity is substantial and swamps observed income and wealth heterogeneity. A characterization of households into latent types reveals very different intrinsic propensities towards the three children's activities and that households with a high propensity to send their children to school are poorer and have less educated parents compared to households in the other classes.

Determinants of child labor and school attendance: the role of household unobservables

Working Paper
December 2002

CONTENTS

1.	Introduction.....	1
2.	The model	2
	2.1 Latent Class Multinomial Logit Model	4
	2.2 Characterizing unobserved heterogeneity	5
	2.3 Computational issues.....	6
3.	Data.....	7
4.	Results.....	8
	4.1 Parameter estimates	9
	4.2 Characteristics of household-level unobserved heterogeneity	9
	4.3 Characteristics of households by posterior class assignment	10
5.	Conclusions.....	11
	References	12
	Appendix A: Statistical tables	14
	Appendix B: Figures.....	21

“My father, ... , had the great respect for education that is often present in those who are uneducated”. (Nelson Mandela, Long Walk to Freedom, p. 6)

1. INTRODUCTION

1. The theoretical literature on child labor has stressed the role of poverty as one of the main determinants of parents’ decision to send their children to work rather than to school (see, for example, Basu, 1999). The empirical results, however, are not so clear cut (Rosati and Tzannatos, 2000; Cigno and Rosati, 2001). More recently, the discussion in the literature has been extended to distinguish between income, assets and the availability of credit but once again the empirical results are ambiguous (Balland and Robinson, 2000; Ranjan, 2001).

2. To the extent that unobserved characteristics of the household determine unobserved components of income and access to credit markets, neglecting such heterogeneity may explain the ambiguous empirical results with respect to the effects of income on child labor supply. But these are not the only sources of unobserved household-level heterogeneity. Costs of and returns to education, and returns to current work are imperfectly observed, if measured at all. Variables such as whether land is cultivated, age composition and location of the household are often used as proxies for returns to child labor. Transportation and distance variables are used as proxies for the cost of education. Returns to education are even more difficult to observe, as one should take in to account the expectation of the parents about the sector in which the child is likely to find employment as an adult, the quality of the child’s education, etc.

3. In this paper, we develop a method that explicitly models household-level heterogeneity and allows us to distinguish between unobserved and observed household heterogeneity. We are thus able to quantify the relative importance of observed household heterogeneity, especially as it relates to differences in income, assets and wealth, and unobserved household heterogeneity that is likely to include important components of costs of education and returns to education and work. Previous research on activity of children has typically ignored household-level heterogeneity. An exception is Jensen and Nielsen (1997) in which fixed and random effects binomial logit models are estimated. While linear models that ignore the unobserved heterogeneity yield unbiased estimates (although not efficient estimates), in nonlinear models, ignoring such unobserved heterogeneity may lead to biased parameter estimates (Heckman and Singer, 1984).

4. We also extend the standard conceptual framework to include the possibility of children being idle, i.e., neither working nor attending school. Much of the literature on determinants of child labor does not distinguish between non-work alternatives, often treating school attendance as the only alternative to work (Jensen and Nielsen, 1997; Ray, 2000; Ravallion and Wodon, 2000). Most survey data show, however, that a substantial fraction of children neither attend school nor participate in work outside the home. In some cases, these children may be engaged in substantial household chores, including taking care of younger children. But in other cases, these children are idle because reasonable work opportunities do not exist and, at the same time, parents do not send them to school either because of a lack of resources or a high relative price of education. We explicitly consider this additional possibility because these children may be substantively different from those who attend school as well as those who work. Ignoring the difference may lead active policy to have unintended

consequences. For example, if school is incorrectly thought of as the only alternative to work, a policy that reduces child work (by reducing returns to work) may simply increase the pool of idle children rather than increasing school attendance, especially if schooling costs are high or returns from schooling are low.

5. The econometric framework we develop is a multinomial logit model with a household-level random intercept. We assume that household-level heterogeneity can be described by a finite number of latent classes or “types” so that the random intercept is drawn from a discrete distribution. This latent class multinomial logit model (LCMNL) is semiparametric in the sense that the discrete density of the random intercept serves as an approximation to any probability density (Lindsay, 1995). An alternative approach would be to specify a parametric density for the random intercept and use integration methods to calculate the response probabilities. But an incorrect specification for this distribution will lead to biased parameter estimates. Our approach liberates us from the difficult task of choosing the correct density. Furthermore, although the discrete representation of the density of the random group effect may be framed as an approximation to some underlying continuous density, the discrete formulation is, itself, a natural and intuitively attractive representation of heterogeneity (Heckman, 2001). An additional desirable feature of the LCMNL is that one can classify each household into a particular class using Bayesian posterior analysis after classical maximum likelihood estimation. Once classified, latent classes or types may be related to group characteristics.

6. A conceptual framework and the LCMNL model is developed in the following section. Data are described in Section 3 and results of the empirical analysis are described in Section 4. We conclude in Section 5.

2. THE MODEL

7. Our empirical model is based on a conceptual framework in which parents allocate the available time of their children to different activities¹. The framework is in the spirit of the new household economics and extends that class of models to explicitly consider children’s labor supply. In particular, following the analysis and classification of Behrman (1997), our model belongs to the class of wealth models with equal concern. We assume that parents control the time of their children when they are young. Children’s time can be used for work and/or for schooling. Work adds to current household consumption, while education increases their future income. We also assume that parents control all the income that accrues to the household, both from adult and child work².

8. Parental decisions are typically framed in a two-period overlapping generations model. During adulthood, individuals earn their income by working, and generate and look after their offspring. Adult’s incomes depend on the stock of human capital accumulated during childhood. Children’s consumption is entirely determined by the transfers they receive from the parents. Given their preferences, parents take into consideration the relative cost of present to future consumption and the amount of resources available in deciding how to allocate their children’s time. This relative cost increases with the costs of education and the returns to child labor and decreases with

¹ The model will not be fully developed here, but just briefly described as its main implications have already been discussed in details (see Rosati and Tzannatos, 2000, and the literature cited therein).

² The unitary model has been criticized and sometime rejected in empirical analysis. However, as shown in Browning et al. (1994) rejection of the unitary model has not implied the rejection of the collective model.

the returns to human capital accumulation. Optimal behavior within this framework is typically modeled as leading to two corner solutions (a child works only or study only) and to an internal solution (a child both studies and works). However, a third corner solution is possible, where children neither go to school nor work, if current children's leisure has a positive value or if there are fixed costs associated with work or schooling.

9. In general, the probability of a child working can be expected to decrease with the household income (net of children's contribution) if capital markets are imperfect or negative bequests are not allowed. On the other hand, higher returns to education, lower education costs and returns to child labor are likely to decrease children's labor supply. With a few exceptions (for example, when children work for a wage) such variables are, at best, imperfectly observed. Typically, distance from school and/or school availability in the village/district are used to proxy for (indirect) education costs. Household composition, availability of land, presence of small children are proxy for the return to a use of time different from education. Age, sex and other individual characteristics are also likely to influence the children's labor supply for well known reasons. The variables used as proxies for the relative cost of education have, hence, both an household and an individual dimension. Returns to work, for example, depends both on the individual ability and on the household availability of labor and of other factor of production.

10. With this underlying conceptual framework in mind, assume that parents in household $j = 1, 2, \dots, J$ choose among activities $k = 0, 1, 2$ (school, work, idle) for child $i = 1, 2, \dots, N_j$ ($\sum_j N_j = N$) on the basis of a random indirect utility function

$$y_{ijk}^* = \alpha_{jk} + x_{ij}\beta_k + z_j\gamma_k + \varepsilon_{ijk} \quad (2.1)$$

x_{ij} is a vector of individual-specific covariates and z_j is a vector of household-specific covariates. α_{jk} is the household-specific intercept and represents the intrinsic propensity (based on variables unobserved by the researcher) of household j for activity k . Assume that the ε_{ijk} are i.i.d. Weibull errors and are orthogonal to the distribution of α_{jk} . Parents will choose activity k over alternative k' if $y_{ijk}^* > y_{ijk'}^*$. Let Y_{ij} ($k = 0, 1, 2$) be an indicator variable denoting the actual choice. Then,

$$\Pr(Y_{ij} = k | \alpha_{jk}, x_{ij}, z_j) = \frac{\exp(\alpha_{jk} + x_{ij}\beta_k + z_j\gamma_k)}{\sum_{k'=0}^K \exp(\alpha_{jk'} + x_{ij}\beta_{k'} + z_j\gamma_{k'})}, \quad (2.2)$$

which is a multinomial logit specification. The standard normalization for the multinomial logit model, which we also adopt, is given by $\alpha_{j0} = \beta_{j0} = \gamma_{j0} = 0$. The joint probability of children's activity in household j is given by

$$\prod_{i=1}^{N_j} \Pr(Y_{ij} = k_i | \alpha_{jk}, x_{ij}, z_j) = \prod_{i=1}^{N_j} \left[\frac{\exp(\alpha_{jk} + x_{ij}\beta_k + z_j\gamma_k)}{\sum_{k'=0}^K \exp(\alpha_{jk'} + x_{ij}\beta_{k'} + z_j\gamma_{k'})} \right]. \quad (2.3)$$

11. Assume that the household-level intercept α_{jk} is a realization from a probability density f . Then the contribution of the j^h household to the log likelihood is given by

$$l_j = \ln \left\{ \int_{-\infty}^{\infty} \prod_{i=1}^{N_j} \left[\frac{\exp(\alpha_{jh} + x_{ij}\beta_h + z_j\gamma_h)}{\sum_{h'=0}^K \exp(\alpha_{jh'} + x_{ij}\beta_{h'} + z_j\gamma_{h'})} \right] f(\alpha_{jh}) d\alpha_{jh} \right\}. \quad (2.4)$$

12. This is a random effects MNL. But note that the integral given in (2.4) does not have a closed form solution for most parametric mixing densities.

2.1 Latent Class Multinomial Logit Model

13. In the latent class multinomial logit (LCMNL) model, the probability density f is assumed to have a discrete support. Specifically, each element of the vector α_{jk} has S points of support with values $[\alpha_{11}\alpha_{12} \dots \dots \alpha_{1S}]$, $[\alpha_{21}\alpha_{22} \dots \dots \alpha_{2S}]$, \dots , $[\alpha_{K1}\alpha_{K2} \dots \dots \alpha_{KS}]$, and associated probabilities $\pi_1, \pi_2, \dots, \pi_S$ where $0 < \pi_1, \pi_2, \dots, \pi_S < 1$ and $\sum_{s=1}^S \pi_s = 1$. Then the contribution of the j^{th} group to the log likelihood is given by

$$l_j = \ln \left\{ \sum_{s=1}^S \pi_s \prod_{i=1}^{N_j} \left[\frac{\exp(\alpha_{hs} + x_{ij}\beta_h + z_j\gamma_h)}{\sum_{h'=0}^K \exp(\alpha_{h's} + x_{ij}\beta_{h'} + z_j\gamma_{h'})} \right] \right\}. \quad (2.5)$$

14. The sample log likelihood is given by

$$l = \sum_{j=1}^J l_j. \quad (2.6)$$

15. It is maximized using the Broyden-Fletcher-Goldfarb-Shanno quasi-Newton constrained maximization algorithm implemented in SAS/IML (SAS Institute, 1997).

16. The standard errors of the parameter estimates are calculated using the robust “sandwich” formulation of the covariance matrix. Because parameter estimates in multinomial logit models are difficult to interpret directly, we report marginal effects of interest. For continuous covariates, the marginal effects are derivatives of the choice probabilities calculated at the mean values of the covariates. For binary covariates, the marginal effects are changes in choice probabilities associated with the discrete changes in the covariates. Standard errors of the marginal effects are constructed using a Monte Carlo technique. First, 500 Monte Carlo replicates of the model parameters are drawn from a multivariate normal distribution with mean given by the point estimates of the parameters and covariance matrix given by the robust sandwich estimate. Next, marginal effects are calculated for each of the 500 parameter vectors. Finally, the standard deviations of the sample of marginal effects are reported as estimates of the standard errors.

17. Selecting a model with an appropriate number of support points is essential. Although a sequential comparison of models with different values of S constitute nested hypotheses, the likelihood ratio test does not have the standard χ^2 distribution because the hypothesis is on the boundary of the parameter space and thus violates the standard regularity conditions for maximum likelihood (Deb and Trivedi, 1997). Model selection criteria based on penalized likelihoods have desirable properties for selecting S and are valid even in the presence of model misspecification (Sin and

White, 1996). We use the Akaike Information Criterion, $AIC = -\ln L + 2K$, and Bayesian Information Criterion, $BIC = -2 \ln L + K \ln(N)$, where $\ln L$ is the maximized log likelihood, K is the number of parameters in the model and N is the sample size. Models with smaller values of AIC and BIC are preferred.

18. Our model is econometrically novel because there is no general methodology for the estimation of random effects models in the context of discrete, count and duration data. There is, however, literature on the estimation of the random effects binomial probit and logit models. In the maximum likelihood estimation of this model, numerical integration (Butler and Moffitt, 1982) or stochastic integration (Keane, 1993) methods are typically used to integrate over the normally distributed random intercept in order to calculate the value of the objective function. Deb (2001) develops a latent class random effects probit model for preventive medical care. Pudney, et al. (1998) estimate a latent class random effects logit model in an analysis of farm tenures. Random effects models in the conditional logit framework have been developed by Jain, et al. (1994) and Kim, et al. (1995). McFadden and Train (2000) discusses random utility formulations, estimation and testing of multinomial logit models with parametric random effects.

2.2 Characterizing unobserved heterogeneity

19. Post-estimation, one can calculate various moments of the distribution of the variance of the intrinsic household-level preference α_k . We report

$$\begin{aligned} E(\alpha_k) &= \sum_{s=1}^S \pi_s \alpha_{k_s}, \\ Var(\alpha_k) &= \sum_{s=1}^S \pi_s (\alpha_{k_s} - E(\alpha_k))^2, \\ Corr(\alpha_k, \alpha_{k'}) &= \frac{\sum_{s=1}^S \pi_s (\alpha_{k_s} - E(\alpha_k))(\alpha_{k'_s} - E(\alpha_{k'}))}{\sqrt{Var(\alpha_k)Var(\alpha_{k'})}}, \end{aligned} \quad (2.7)$$

for $k, k' = 1, 2, \dots, K$. The variance of the household-level unobserved heterogeneity is compared to variances of observed household-level heterogeneity. The correlations describe whether intrinsic household propensities for one activity over another are correlated.

20. In the latent class interpretation of the random intercept, each point of support and associated probability describes a latent class or a type of household. The posterior probability that a particular household belongs to a particular class can be calculated as

$$Pr[j \in \text{class } c] = \frac{\pi_c \prod_{i=1}^{N_j} \left[\frac{\exp(\alpha_{k_s} + \beta_{ij} \beta_k + \gamma_j \gamma_k)}{\sum_{k'=0}^K \exp(\alpha_{k'_s} + \beta_{ij} \beta_{k'} + \gamma_j \gamma_{k'})} \right]}{\sum_{s=1}^S \pi_s \prod_{i=1}^{N_j} \left[\frac{\exp(\alpha_{k_s} + \beta_{ij} \beta_k + \gamma_j \gamma_k)}{\sum_{k'=0}^K \exp(\alpha_{k'_s} + \beta_{ij} \beta_{k'} + \gamma_j \gamma_{k'})} \right]}; \quad c = 1, 2, \dots, S. \quad (2.8)$$

21. These posterior probabilities are used to classify each household into a latent class in order to study the properties of the classes of households further (see Deb and Trivedi, 2002, for an example).

2.3 Computational issues

22. Two computational issues arise in the estimation of LCMNL. The first of these is a general issue in the estimation of latent class models. The second is a general issue in the estimation of random effects discrete choice models.

23. Even when the parameters are identified, estimation of latent class models is not always straightforward. Their likelihood functions can have multiple local maxima so it is important to ensure that the algorithm converges to the global maximum. Moreover, if a model with too many points of support is chosen, one or more points of support may be degenerate, i.e., the π_s associated with those densities may be zero. In such cases, the solution to the maximum likelihood problem lies on the boundary of the parameter space. This can cause estimation algorithms to fail, especially if unconstrained maximization algorithms are used. Such cases are strong indication that a model with fewer components adequately describes the data. Therefore, a small-to-large model selection approach is recommended, i.e., the number of points of support in the discrete density should be increased one at a time starting with a model with only two points of support.

24. The performance of the maximum likelihood estimators of random effects models for binary and multinomial responses given by (2.3) may not be satisfactory for large group sizes, N_j , since the log likelihood involves the integration or summation over a term involving the product of probabilities for all group members. In the context of the random effects probit model, Borjas and Sueyoshi (1994) point out that with 500 observations per group, and assuming a generous likelihood contribution per observation, the product would be well below standard computer precision. They speculate that group sizes over 50 may create significant instabilities if the model has low predictive power. Based on Monte Carlo experiments, they find that such computational problems lead to quite in-accurate statistical inference on the parameters of the model. Although the group sizes in our data are considerably smaller, we cannot rule out the possibility of underflows.

25. We use a method developed by Lee (2000) to alleviate this computational problem. The likelihood function (2.5) is evaluated as

$$l_j = \ln \left\{ \sum_{s=1}^S \exp(h_{js}) \right\}, \quad (2.9)$$

Where

$$h_{js} = \ln(\pi_s) + \sum_{i=1}^{N_j} \ln \left\{ \left[\frac{\exp(\alpha_{ks} + x_{ij}\beta_k + z_j\gamma_k)}{\sum_{k'=0}^K \exp(\alpha_{k's} + x_{ij}\beta_{k'} + z_j\gamma_{k'})} \right] \right\}, \quad (2.10)$$

for all $s = 1, 2, \dots, S$ and $j = 1, 2, \dots, J$. Denote $p_j = \max\{h_{js} : s = 1, 2, \dots, S\}$.

Then

$$l_j = p_j + \ln \left\{ \sum_{s=1}^S \exp(h_{js} - p_j) \right\}, \quad j = 1, 2, \dots, J. \quad (2.11)$$

We have found this method to be quite accurate and fast.

3. DATA

26. We examine the importance of household-level observed and unobserved characteristics using data from two large household surveys. The first sample consists of data from the Core Welfare Indicators Questionnaire (CWIQ) Survey conducted in Ghana in 1997. The second sample consists of data from the Human Development of India Survey (HDIS) conducted in rural India in 1994.

27. The CWIQ survey, which was carried out by the Ghana Statistical Service (GSS) in collaboration with the World Bank, is primarily designed to furnish policy makers with a set of indicators for monitoring poverty and the effects of development policies, programs and projects on living standards in the country. The CWIQ focuses on the collection of information to measure access to, utilization of, and satisfaction with key social and economic services. A total of 14,514 households were successfully interviewed. Almost 23 percent of all children between the ages of 6 and 15 live in households where no parent is present. This raises substantial theoretical and empirical issues because one expects households in which parents of the children are not present to have different decision-making structures and behave differently than households in which parents are present. An examination of such differences is clearly an important issue, but one we leave for future work. However, in order to cleanly model and interpret household-specific observable and unobservable effects, we eliminate children who do not live with at least one parent. Our sample consists of 13484 children between the ages of 6 and 15 in 6701 households with at least one parent present (henceforth we use the words child and children to refer to children between the ages of 6 and 15). Both parents are present in 73.3 percent of cases, the mother of the children is present alone in 23.3 percent of cases while the father is present alone in the remaining 3.4 percent of cases.

28. The HDIS, which was carried out by the National Council of Applied Economic Research (NCAER), is a multi-purpose, nationally representative sample survey of rural India. The sample consists of 34,398 households spread over 1,765 villages in 16 states. Two separate survey instruments were used, one to elicit the economic and income parameters from an adult male member, and the other to collect data on outcomes such as literacy, education, health, morbidity, nutrition, and demographic parameters from the adult female members of the household. In the HDIS sample, by definition, single parent households do not have complete data for our purposes. Our sample consists of 34211 children between the ages of 6 and 15 in 16371 households.

29. Table 1 shows the distribution of children within households. In Ghana, about 39 percent of households have only one child while in India the corresponding fraction is about 35 percent. Over 30 percent of households in either country have three or more children. These estimates highlight the importance of care-fully modeling household-level effects in any analysis of children's behavior or outcomes.

30. The dependent variable is defined using three mutually exclusive categories to identify children's activities: school, work and idle. In Ghana, 0.66 percent of children in our sample report working and attending school. In India, 0.71 percent of children report working and attending school. These frequencies are too small to analyze as a separate category. Consequently, we classify the activity of such children as working. Note however, that our results are robust to the exclusion of these observations from our sample. Table 2 shows that in Ghana, 78 percent of children are in school, less than 8 percent work and 14 percent are idle. In India school enrollment is about 64 per cent, while about 13 per cent of children work and 23 per cent are idle.

31. The set of explanatory variables is defined in Table 2. It includes individual characteristics such as age and gender (female). Resources available to the household

are proxied by a dummy variable for the household being poor, i.e., belonging to the lowest income quintile, and by appliances which measures the number of appliances in the household. We have chosen to use the dummy variable poor instead of a continuous measure of income for two reasons. First, it is well known that measures of income in developing countries, especially in the lower end of the income distribution, have significant measurement error. Our crude measure of income is not likely to have much measurement error. Second, continuous measures of income are endogenous because they include children's income. Our measure is likely to minimize endogeneity biases because it is unlikely that a child's income will change the value of poor for a household.

32. Returns to work are proxied by two variables that indicate whether the household owns land and livestock (livstk). We did not consider children's wages as only a few children in our sample work for a wage. Education of the parents (ed-mother, ed-father) is included in our models. In the sample from Ghana, education is measured in number of years of schooling. In the sample from India, education is an ordered variable with increments denoting substantive increases in education (e.g. from primary to lower secondary to higher secondary).³ Other household characteristics are the number of children (child) and the religion (hindu, muslim, christian) and social status (scst) in the case of India. Costs of primary and secondary education (primschl, secoschl) are proxied by the distance from primary and secondary schools in Ghana, and by dummy variables indicating the presence of primary and secondary schools in the village in the case of India. In the case of Ghana, we also include a dummy variable for urban location (urban) location of the household. In all our models, we also include a set of region fixed effects: nine regions in Ghana and fifteen states in India.

33. Table 2 also reports means of explanatory variables by category of activity. It shows that girls are more likely to be idle in either country, but there is little difference between girls and boys in terms of work and school in Ghana, while in India girls are also more likely to be working. Those children who are from households with the greatest number of children, most poorly educated parents, poorest in income and assets are apparently less likely to attend school. In addition, they are most likely to come from agricultural and rural households who live farthest from schools.

4. RESULTS

34. We have estimated LCMNL models with two through five points of support for the latent class densities. We have also estimated a standard MNL model, which does not allow for household-specific random intercepts, and may be interpreted as a degenerate latent class model with one point of support. As Table 3 shows, there is a dramatic improvement in the maximized log likelihood once household-specific random effects are introduced. The AIC and BIC, also reported in Table 3, both suggest that a density with four points of support adequately describes the distribution of the random intercepts⁴. Consequently, we conclude that there are four latent types of households and present further results from the model with four latent classes.

³ Although it would be preferable to treat education in the sample from India as a sequence of dummy variables, we chose not to do so to keep the model as parsimonious as possible.

⁴ In random effects models, there is an open question about whether one should use the number of groups (independent) or the number of individuals (not independent) as the sample size. We report BIC with N=number of families, although in our case, BIC supports the same model if N=number of children is used.

4.1 Parameter estimates

35. Tables 4a-b reports parameter estimates and marginal effects for the model with four latent classes for Ghana and India respectively. Being poor increases the probability of working and decreases the probability of attending school. The variable proxying for pure wealth effects, appliances, has the expected effect on the decisions concerning child labor and schooling, i.e., children in wealthier households are more likely to attend school and less likely to work. Land and livestock ownership have negative effects on the probability of attending school, but these effects are only statistically significant in the case of India. The lack of significance in the case of Ghana may be due to the fact that these variables are likely to have income and substitution effects. On one hand, ownership of land or livestock are likely to be associated with higher incomes; on the other hand they also proxy the marginal value of children's time in working activities. It is possible that income and substitution effects counterbalance each other, so that the estimated coefficients are not significant.

36. Girls are less likely to attend school and more likely to be idle. In Ghana, girls are no more likely to work than boys while in India, girls are also more likely to work. Older children are more likely to attend school and work and are less likely to be idle but in each case the effect is nonlinear. The presence of siblings reduces the probability of attending school and raises that of working and especially of being idle. Children with more educated parents are more likely to attend school and less likely to work or be idle. The further the school (especially primary school in the case of Ghana), the less likely children are to attend school and more likely to be idle, indicating that it represents a significant component of the cost of education. Interestingly, distance from school has little or no affect on the probability of working.

37. Overall, the marginal impacts of most covariates on being idle are statistically significant and large. Importantly, much of the substitution between activities as a response to changes in explanatory variables is between attending school and being idle. The effects of these exogenous covariates on work are substantially smaller. These results highlight the importance of treating idleness as a distinct category of activity and point to the possibility of unintended consequences when policies are based on a framework in which school and work are the only activity choices.

4.2 Characteristics of household-level unobserved heterogeneity

38. In Table 5 we report statistical characteristics of the random intercepts. Of special interest is the variance of the random intercept, $\text{Var}(\alpha_k)$, which measures unobserved heterogeneity as compared to the variance explained by a linear combination of covariates (using estimated coefficients as the weights), $\text{Var}(\widehat{Z\theta})$, a measure of observed heterogeneity. The results show that household-level unobserved heterogeneity is substantial. The unobserved household-level heterogeneity accounts for a minimum of 43 percent and a maximum of 117 percent of the variance due to the corresponding observed heterogeneity. If one focuses on the variance due to household income and wealth (poor, land, livstk, appliances), $\text{Var}(Z_1\widehat{\theta}_1)$, it is clear that household-level unobserved heterogeneity swamps observed income and wealth heterogeneity.

39. We also report the correlation between the random intercepts in the work and idle equations. They are positively correlated and large in magnitude indicating that households in which children are more likely to work relative to attending school are also households in which they are more likely to be idle.

40. In Tables 6a-b, we report the values of the support points of the distribution of random intercepts (α_{ks}) along with their associated probabilities (π_s). In addition, for each of the four points of support, we calculate the predicted probability of each activity, $\Pr(Y_{ij} = k)$, as the sample average over all individuals in the sample. Households in class 4 are most common. They account for about 56 percent of households in Ghana and 49 percent of households in India. These are “average” households in the sense that the probabilities of their children’s activities are close to the original sample probabilities. A small fraction of households, less than 2 percent of households in Ghana and just under 10 percent in India, belong to latent class 1 and have high intrinsic propensities towards child labor. Children in these households are more likely to work than children in any of the three other types of households. Note, however, that while the propensity for children to work in this class of households is extremely high relative to the two other activities in Ghana, the probability of school is also substantial in the case of India. In contrast, a relatively large number of households (over 30 percent in both countries) belong to class 2 who almost always send their children to school. Class 3 consists of households (around 7 to 12 percent) whose children are most likely to be idle, with school being the second most likely activity. We reported earlier that marginal changes in income, assets and other explanatory variables tend to have the largest impact on the likelihood of being idle and especially tend to cause substitution between attending school and being idle. Therefore, policy interventions and changes in external conditions are likely to produce the greatest changes in the behavior of households in class 3. On the other hand, children in households of class 1 are likely to have only small responses to marginal changes in external conditions.

4.3 Characteristics of households by posterior class assignment

41. The results described above suggest that policy should ideally be targeted towards particular types of households. Unfortunately, targeting on observables may be of limited value as we have shown that unobservables heavily influence the behavior of households. In order to improve targeting, it is important to improve the quality of data, especially as it relates to costs of and returns to education, credit constraints, etc. In the absence of richer data, however, our model allows the possibility of building a “risk profile” by examining the posterior class assignment of households. In order to do so, the posterior probability of belonging to each of the four classes was calculated for each household using the formula in (2.8), conditional on observed covariates and outcomes. Next, each household was classified into a unique class on the basis of the maximum posterior probability. Finally, sample averages were calculated for each explanatory variable stratified by household classification. Sample averages and 95 percent confidence intervals for each of these covariates by household-type are displayed in Figures 1a-b.

42. At first glance, these findings appear to contradict the random effects assumption: since the random intercept is assumed to be uncorrelated with the covariates, how can the covariate averages differ significantly across latent classes? But a closer look at the definition of the posterior probability (equation 2.8) resolves this apparent contradiction. The a priori assumption regarding the relationship between the random intercept and the covariates is conditional only on the covariates. The posterior relationship, however, is conditional on covariates and outcomes. In other words,

armed with only knowledge of explanatory variables, it is not possible to infer anything about the type of household. But once the outcome is known for each household member, this additional information makes it possible to infer features of the type of household.

43. Households in latent class 2, characterized by a high propensity to send their children to school are poorer compared to households in the other classes. For this large group of households (about 30 percent in both countries), the so-called poverty axiom is contradicted: they are poor yet they have a high propensity to send their children to school. We speculate that this is because the cost of education for children in the poorest households is less than for children in other households because their education expenses are heavily subsidized. Moreover, such children likely also have the fewest work opportunities. Of course, without better data on the costs of and returns to education, these possibilities cannot be explored further. Children in these households, most likely to attend school, also have the least educated parents on average. It is possible that parents' education proxies for household wealth and work opportunities for the children, but perhaps Mandela's observation (see the quote that precedes this paper) has merit!

5. CONCLUSIONS

44. We show that unobserved heterogeneity at the household-level is substantial compared to observed heterogeneity at the individual and household levels. Specifically, unobserved household heterogeneity is responsible for considerably greater variance of outcomes than observed income and wealth heterogeneity. The proxies for costs of and returns to education available in the data do not substantially reduce the effects of unobserved household-level heterogeneity. Our characterization of households into four latent classes reveals very different intrinsic propensities towards the three children's activities. Households with high propensities to send their children to school are poorer and have less educated parents compared to households in the other classes.

45. Changes in observed income, wealth, costs of and returns to education and other explanatory variables tend to cause substitution in children's activities between attending school and being idle. Child labor, however, appears to be rather resistant to marginal changes in explanatory variables.

46. These findings have three important implications. First, research and policy design should be reoriented to focus more attention on other household-level determinants of child labor besides income. To achieve this aim it might be necessary to modify survey instruments currently utilized to gather information on child labor. Secondly, the (partial) rejection of the poverty axioms suggests that it may be possible to reduce child labor without relying only on income growth. This offers support to the plans developed and/or under consideration by many governments and international agencies aiming to eradicate the worst forms of child labor. Finally, the phenomenon of children who neither work nor attend school warrant considerably greater attention in theoretical and empirical work on children's activities as well as in survey design. They are clearly a vulnerable group and may be worse off in a human capital sense than children who work.

REFERENCES

- Balland, J.M. and A. Robinson (2000), "Is Child Labor Inefficient?", *Journal of Political Economy*, 108, 663-679.
- Basu, K. (1999), "Child Labor: Cause, Consequence and Cure, with Remarks on International Labor Standards", *Journal of Economic Literature*, 37, 1083-1119.
- Behrman, J. (1997), "Intrahousehold Distribution and the Family", in Rosenzweig A. D. and O Stark, *Handbook of Population and Family Economics*, Elsevier.
- Borjas, G.J. and G.T. Sueyoshi (1994), "A Two-Stage Estimator for Probit Models with Structural Group Effects", *Journal of Econometrics*, 64, 165-182.
- Browning, B., F. Bourguignon F., P. Chiappori and V. Lechene (1994), "Income and Outcomes: a Structural Model of Intrahousehold Allocation", *Journal of Political Economy*, 102, 1067-1096.
- Butler, J.S. and R. Moffitt (1982), "A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model", *Econometrica*, 50, 76 1-764.
- Cigno A. and F. Rosati (2002), Child Labour, Education and Nutrition, *Pacific Economic Review*, 7, pp
- Deb, P. (2001), "A Discrete Random Effects Probit Model with Application to the Demand for Preventive Care", *Health Economics*, 10, 37 1-383.
- Deb, P., and P.K. Trivedi (1997), "Demand for Medical Care by the Elderly in the United States: A Finite Mixture Approach", *Journal of Applied Econometrics*, 12, 313-336.
- Deb, P., and P.K. Trivedi (2002), "The Structure of Demand for Health Care: Latent Class versus Two-part Models", *Journal of Health Economics*, 21, 601-625.
- Heckman, J. J. (2001), "Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture", *Journal of Political Economy*, 109, 673-748.
- Heckman, J. J. and B. Singer (1984), "A Method of Minimizing the Distributional Impact in Econometric Models for Duration Data", *Econometrica*, 52, 271-320.
- Jain, D.C., N.J. Vilcassim and P.K. Chintagunta (1994), "A Random-Coefficients Logit Brand-Choice Model Applied to Panel Data", *Journal of Business and Economic Statistics*, 12, 317-328.
- Jensen, P. and H.S. Nielsen (1997), "Child Labour or School Attendance? Evidence from Zambia", *Journal of Population Economics*, 10, 407-424.

Keane, M.P. (1993), "Simulation Estimation for Panel Data with Limited Dependent Variables", Chapter 20 in Maddala, G.S., C.R. Rao and H.D. Vinod (Eds.), *Handbook of Statistics*, Vol. 11, North Holland: Amsterdam.

Kim, B-D., R.C. Blattberg and P.E. Rossi (1995), "Modeling the Distribution of Price Sensitivity and Implications for Optimal Retail Pricing", *Journal of Business and Economic Statistics*, 13, 291-303.

Lee, L-F. (2000), "A Numerically Stable Quadrature Procedure for the One-Factor Random-Component Discrete Choice Model", *Journal of Econometrics*, 95, 117-129.

Lindsay, B. J. (1995), *Mixture Models: Theory, Geometry, and Applications*, NSF-CBMS Regional Conference Series in Probability and Statistics, Vol. 5, IMS-ASA.

McFadden, D. and K. Train (2000) "Mixed MNL Models for Discrete Response", *Journal of Applied Econometrics*, 15, 447-470.

Pudney, S., F.L. Galassi and F. Mealli (1998), "An Econometric Model of Farm Tenures in Fifteenth-Century Florence", *Economica*, 65, 535-556.

Ranjan, P. (2001), "Credit Constraints and the Phenomenon of Child Labor", *Journal of Development Economics*, 64, 8 1-102.

Ravallion, M. and Q. Wodon (2000), "Does Child Labour Displace Schooling? Evidence on Behavioural Responses to an Enrollment Subsidy", *The Economic Journal*, 110, C158-C175.

Ray, R. (2000), "Analysis of Child Labour in Peru and Pakistan: A Comparative Study", *Journal of Population Economics*, 13, 3-19.

Rosati, F. and Z. Tzannatos (2000), "Child Labor in Vietnam: An Economic Analysis", *World Bank Working Paper*.

SAS Institute (1997), *SAS/IML Software: Changes and Enhancements through Release 6.11*, SAS Institute Inc.

Sin, C-Y, and H. White (1996), "Information Criteria for Selecting Possibly Misspecified Parametric Models", *Journal of Econometrics*, 71, 207-25.

APPENDIX A: STATISTICAL TABLES

Table 1

no. of children	Ghana		India	
	no. of households	percent	no. of households	percent
1	2593	38.70	5656	34.55
2	2107	31.44	5492	33.55
3	1297	19.36	3661	22.36
4	485	7.24	1263	7.71
5	155	2.31	269	1.64
6	39	0.58	22	0.13
7	14	0.21	5	0.03
8	4	0.06	3	0.02
9	1	0.01		
10	4	0.06		
11	2	0.03		

Table 2

Variable	Definition	Ghana			India		
		school=1	work=1	idle=1	school=1	work=1	idle=1
Sample size		10798	1057	1993	21895	4268	8048
percent		77.98	7.63	14.39	64.00	12.48	23.52
Female	= 1 if female age in	0.466	0.462	0.528	0.411	0.629	0.537
Age	years number of children	10.099	11.482	9.801	10.385	12.804	9.117
Child	Number of children	4.245	4.870	4.616	3.770	4.0529	4.145
Ed-mother	education of mother	6.333	1.117	3.163	1.497	1.091	1.072
Ed-father	education of father	8.095	1.570	4.146	2.270	1.481	1.485
Primschl	distance to primary school ¹	2.398	3.501	2.875	0.553	0.510	0.493
Secoschl	distance to secondary school ²	5.513	6.296	5.920	0.542	0.579	0.637
Poor	= 1 if household income in lowest quintile ³	0.168	0.360	0.309	0.519	0.632	0.679
Land	= 1 if household owns land	0.415	0.621	0.419	0.682	0.626	0.632
Livstk	= 1 if household owns livestock	0.373	0.655	0.447	0.706	0.701	0.688
Appliance	Number of appliances in household	2.201	0.929	1.315	1.077	0.557	0.443
Urban	= 1 if urban	0.337	0.103	0.238			
Hindu	= 1 if Hindu				0.824	0.818	0.800
Muslim	= 1 if Muslim				0.110	0.137	0.169
Christian	= 1 if Christian				0.028	0.007	0.009
scst	= 1 if scheduled caste or tribe				0.328	0.440	0.483

NOTES:

1. Distance to primary school is measured in 10 minute increments in the sample from Ghana. In the sample from India, distance is a binary indicator equal to 1 if a school is not in the village.
2. Distance to secondary school is measured in 10 minute increments in the sample from Ghana. In the sample from India, distance is a binary indicator equal to 1 if a school is not in the village.
3. In the sample from India, income quintiles are defined over urban and rural populations, although the sample consists of only rural households. Hence the fraction poor is much greater than the expected 25%.
4. Nine region categories are defined for the sample from Ghana. Fifteen state categories are defined for the sample from India.

Table 3

classes	Ghana				India			
	log likelihood	K	AIC	BIC	log likelihood	K	AIC	BIC
1	-7273.85	44	14635.71	14935.35	-22723.23	64	45574.46	46067.47
2	-6911.10	47	13916.20	14236.27	-21758.92	67	43651.84	44167.96
3	-6788.45	50	13676.91	14017.41	-21503.21	70	43146.42	43685.65
4	-6707.51	53	13521.03 ^w	13881.96 ^w	-21419.24	73	42984.48 ^w	43546.82 ^w
5	-6707.47	56	13526.94	13908.30	-21419.12	76	42990.25	43575.70

Notes:

1. * model preferred by information criterion.

Table 4: Ghana

	Parameters			Marginal effects	
	work	idle	school	work	idle
female	0.290 ^w (0.112)	0.437 ^w (0.069)	-3.501 ^w (0.585)	0.082 (0.064)	3.419 ^w (0.578)
age	0.276 (0.206)	-1.794 ^w (0.138)	13.753 ^w (1.087)	0.294 ^w (0.144)	-14.047 ^w (1.075)
age2	0.006 (0.010)	0.086 ^w (0.007)	-0.669 ^w (0.051)	-0.006 (0.006)	0.675 ^w (0.050)
kids	0.067 (0.035)	0.050 ^w (0.020)	-0.410 ^w (0.160)	0.024 (0.016)	0.386 ^w (0.155)
ed-mother	-0.109 ^w (0.024)	-0.046 ^w (0.008)	0.392 ^w (0.062)	-0.042 ^w (0.013)	-0.349 ^w (0.060)
ed-father	-0.077 ^w (0.016)	-0.042 ^w (0.007)	0.353 ^w (0.056)	-0.029 ^w (0.013)	-0.324 ^w (0.054)
primschl	0.120 ^w (0.041)	0.144 ^w (0.030)	-1.149 ^w (0.234)	0.038 (0.024)	1.112 ^w (0.230)
secoschl	0.008 (0.063)	0.007 (0.029)	-0.057 (0.231)	0.003 (0.030)	0.054 (0.223)
poor	0.611 ^w (0.157)	0.385 ^w (0.102)	-3.405 ^w (0.964)	0.261 ^w (0.113)	3.144 ^w (0.935)
land	0.005 (0.193)	0.061 (0.100)	-0.472 (0.826)	-0.004 (0.094)	0.476 (0.798)
livstk	-0.236 (0.163)	0.010 (0.100)	-0.000 (0.854)	-0.101 (0.096)	0.101 (0.825)
appliances	-0.272 ^w (0.072)	-0.180 ^w (0.034)	1.482 ^w (0.288)	-0.099 ^w (0.034)	-1.383 ^w (0.281)
urban	-1.205 ^w (0.302)	-0.284 ^w (0.127)	2.468 ^w (0.947)	-0.428 ^w (0.195)	-2.041 ^w (0.912)

Notes:

1. * statistically significant at the 5 percent level.

2. Marginal effects and associated standard errors are reported in percentage points.

Table 4b:India

	Parameters		Marginal effects		
	work	idle	school	work	idle
female	1.734 ^w (0.067)	1.159 ^w (0.060)	-15.343 ^w (0.929)	6.377 ^w (0.713)	8.966 ^w (0.566)
age	-0.132 (0.526)	-1.972 ^w (0.165)	16.761 ^w (2.826)	1.575 (1.958)	-18.335 ^w (1.187)
age2	0.035 (0.023)	0.086 ^w (0.008)	-0.830 ^w (0.129)	0.058 (0.091)	0.772 ^w (0.053)
child	0.114 ^w (0.023)	0.150 ^w (0.022)	-1.622 ^w (0.231)	0.339 ^w (0.100)	1.282 ^w (0.202)
ed-mother	-0.834 ^w (0.089)	-0.657 ^w (0.075)	8.202 ^w (0.950)	-2.959 ^w (0.453)	-5.243 ^w (0.785)
ed-father	-0.580 ^w (0.041)	-0.603 ^w (0.034)	6.912 ^w (0.389)	-1.899 ^w (0.224)	-5.013 ^w (0.316)
primschl	0.221 (0.122)	0.435 ^w (0.106)	-4.333 ^w (1.106)	0.497 (0.467)	3.836 ^w (0.950)
secoschl	0.472 ^w (0.126)	0.597 ^w (0.118)	-6.505 ^w (1.179)	1.428 ^w (0.462)	5.076 ^w (1.040)
poor	0.238 ^w (0.075)	0.085 (0.071)	-1.494 ^w (0.731)	0.956 ^w (0.346)	0.538 (0.592)
land	-0.210 ^w (0.085)	-0.270 ^w (0.070)	2.930 ^w (0.772)	-0.629 (0.351)	-2.301 ^w (0.663)
livstk	-0.069 (0.097)	-0.184 ^w (0.081)	1.750 (0.980)	-0.103 (0.391)	-1.647 ^w (0.753)
appliances	-0.555 ^w (0.041)	-0.568 ^w (0.038)	6.545 ^w (0.472)	-1.826 ^w (0.263)	-4.720 ^w (0.379)
hindu	-0.078 (0.202)	0.520 (0.327)	-4.049 (3.149)	-0.912 (0.771)	4.961 (3.078)
muslim	0.670 ^w (0.240)	1.369 ^w (0.350)	-13.556 ^w (3.286)	1.455 (0.930)	12.101 ^w (3.031)
christian	-0.761 (0.463)	0.836 (0.451)	-4.396 (4.414)	-4.268 ^w (1.790)	8.663 ^w (3.935)
scst	0.448 ^w (0.081)	0.568 ^w (0.069)	-6.184 ^w (0.758)	1.353 ^w (0.338)	4.831 ^w (0.632)

Notes:

- * statistically significant at the 5 percent level.
- Marginal effects and associated standard errors are reported in percentage points.

Table 5

	Ghana		India	
	Work	Idle	Work	Idle
$E(a_k)$	-4.196	6.601	-5.202	8.790
$Var(a_k)$	5.990	2.3223	2.676	3.171
$Corr(a_k)$	0.872		0.794	
$Var(Z\hat{\theta})^1$	14.055	1.982	7.272	5.725
$Var(Z_1\hat{\theta}_1)^2$	0.824	0.195	0.415	0.506

NOTES:

1. Z denotes the full set of covariates and θ the associated estimated parameter vector.
2. Z1 denotes covariates associated with household wealth (poor, land, livstk, appliances) and $\hat{\theta}_1$ the associated estimated parameter sub-vector

Table 6a: Ghana

	latent class 1	latent class 2	latent class 3	latent class 4
α_1	7.070 ^W (1.970)	-7.065 ^W (1.019)	-1.859 (1.733)	-3.093 ^W (1.089)
α_2	10.210 ^W (1.810)	4.885 ^W (0.699)	10.418 ^W (0.832)	7.071 ^W (0.677)
π	0.017 ^W (0.002)	0.341 ^W (0.053)	0.067 ^W (0.018)	0.575 (.)
<i>Pr(school)</i>	0.068	0.954	0.261	0.757
<i>Pr(work)</i>	0.843	0.010	0.055	0.095
<i>Pr(idle)</i>	0.088	0.035	0.684	0.148

Table 6b: India

	latent class 1	latent class 2	latent class 3	latent class 4
α_1	-1.625 (2.478)	-7.168 ^W (2.643)	-3.519 (2.698)	-5.224 ^W (2.580)
α_2	9.608 ^W (0.707)	6.203 ^W (0.679)	12.386 ^W (0.844)	9.142 ^W (0.701)
π	0.095 ^W (0.028)	0.281 ^W (0.071)	0.132 ^W (0.043)	0.492 (.)
<i>Pr(school)</i>	0.353	0.914	0.260	0.642
<i>Pr(work)</i>	0.470	0.042	0.110	0.111
<i>Pr(idle)</i>	0.177	0.044	0.629	0.247

Notes:

* statistically significant at the 5 percent level.

APPENDIX B: FIGURES

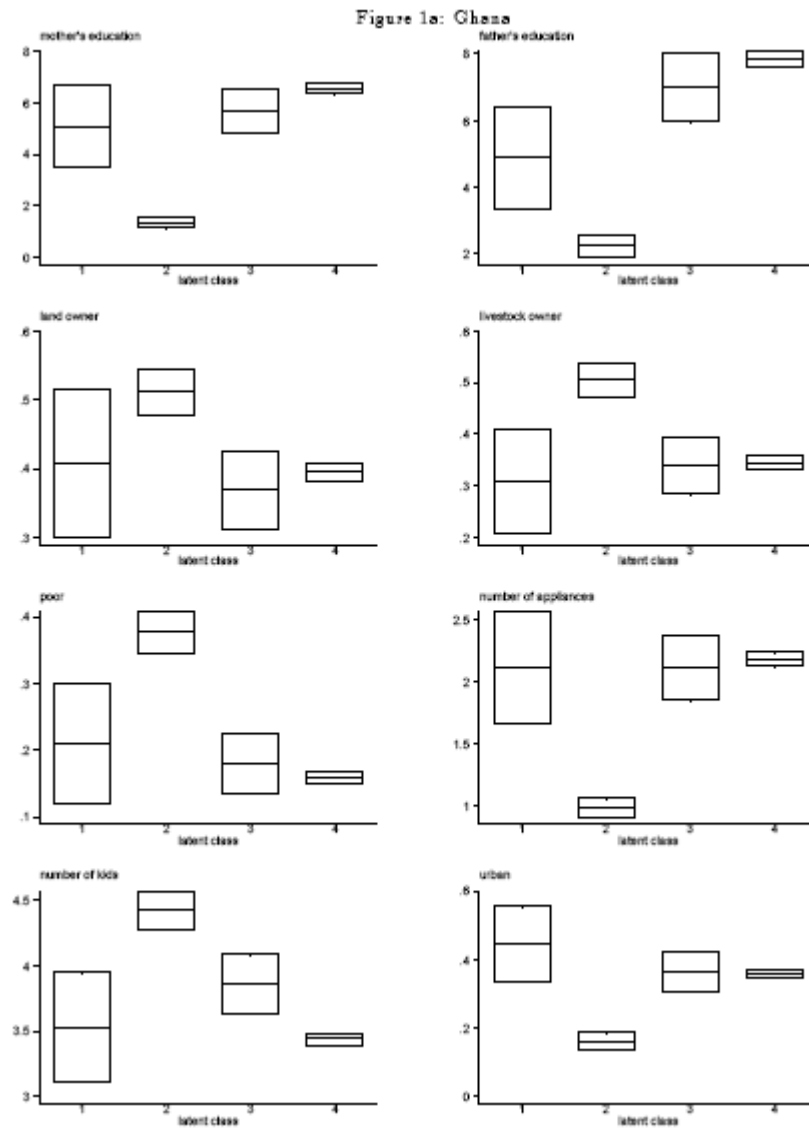


Figure 1b: India

