



Understanding Children's Work
An Inter-Agency Research Cooperation Project

Understanding Children's Work Project Working Paper Series, February 2004

Estimating the effect of fertility decisions on child labor and schooling

P. Deb
F. C. Rosati

February 2004

Estimating the effect of fertility decisions on child labor and schooling

P. Deb*

F. C. Rosati**

**Working Paper
February 2004**

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.

* Department of economics, Hunter college, City University of New York

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

Estimating the effect of fertility decision on child labor and schooling

**Working Paper
February 2004**

ABSTRACT

We use an econometric model of fertility and children's activities to examine the causal effects of fertility on a child's activities taking the endogeneity of fertility into account. Our specification is nonlinear and simultaneous and uses latent factors to allow for unobserved influences on fertility to affect a child's activities. We apply maximum simulated likelihood (MSL) techniques to estimate the parameters of our models. We find that the effect of fertility has a large downward bias in naive models. The effect of fertility on the probability of attending school is twice as large once its endogeneity is taken into account. The effect of fertility on the probability of work changes sign and becomes statistically significant.

Estimating the effect of fertility decisions on child labor and schooling

Working Paper
February 2004

CONTENTS

1.	Introduction.....	1
2.	Conceptual Framework.....	3
3.	Econometric Model.....	4
3.1	Quasi-Monte Carlo simulation.....	7
4.	Data.....	8
4.1	Instruments.....	9
5.	Results.....	10
6.	Conclusion.....	12
	References.....	13
	Annexes.....	15

1. INTRODUCTION

1. Household composition and the allocation of time of its members are jointly determined. Consequently, the effect of fertility on labor supply cannot be properly determined without taking the endogeneity of the number of children in the household into account. In the context of adult labor supply, this issue has been recognized and modeled in a variety of ways. See, for example, Angrist (2001), Angrist and Evans (1998), Francesconi (2002), Gong and van Soest (2002), Hotz and Miller (1988), Lundberg and Rose (2002) and Moffit (1984). The results generally show strong evidence of the endogeneity of fertility.

2. In the empirical literature on child labor, however, the possible endogeneity of fertility has not been addressed. For example, Deb and Rosati (2002), Jensen and Nielsen (1997), Maitra and Ray (2002), Ravallion and Wodon (2000) and Ray (2000), all include variables for the number of children in models of child labor and schooling but these variables are treated as if they are exogenous. It is commonly believed that households with more children are also more likely to send them to work and less likely to send them to school, although evidence in the empirical literature is mixed. In any case, because fertility is endogenous to household decisions, it is difficult to give a casual interpretation to these stylized facts. For example, unobserved shifts in preferences or changes in costs of education might lead a household to have more children and to send them to work rather than to school.

3. In this paper, we examine the effect of fertility on child labor and schooling after taking into account the possible endogeneity of fertility. It is reasonable to believe that important sources of heterogeneity common to fertility and child activity equations are unobserved in the data. In addition to unobserved preferences, the data typically have poor proxies for income, wealth and returns to and costs of education and child labor. Therefore, econometric models of children's activity that treat the number of children in a household as exogenous would provide biased estimates of the effect of the number of children. While endogeneity is expected, the direction and magnitude of such effects is not clear. In the literature on adult labor supply, Angrist (2001) finds that the effect of children on female labor supply when endogeneity is accounted for is just over half as large as the effect when fertility is treated exogenously. Lundberg and Rose (2002), on the other hand, find evidence that the effect of children on male labor supply is substantially larger when endogeneity is taken into account.

4. We depart from the common practice of estimating our structural model using a linear, instrumental variables approach and instead develop a structural econometric model of fertility and children's activities that respects important distributional features of the data for the following reasons. First, note that a child's time can be allocated between schooling, labor and neither being in school nor working substantially. Most survey data show that a substantial fraction of children neither attend school nor participate in work outside the home and Deb and Rosati (2002) show that these children are substantively different from those who work and those who attend school. In some cases, these children may be engaged in substantial household chores, which may not get reported as work. 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. These suggest the importance of modeling a child's activities as a multinomial process and that linearization of children's activities requires that the multinomial variable be reduced to a binary variable, which has substantive and statistical consequences. Second, fertility takes only a small number of integer values, hence is a count process; linearizing it leads to substantially inefficient estimates with poor finite sample properties. Consequently, we model the number of children in the household using a Poisson process and a child's activities using a multinomial logit process.

5. Latent factors are incorporated into the equations for number of children and children's activities to allow for unobserved influences on fertility to affect children's activities. The latent factor structure has three main virtues over alternative ways of generating correlated errors. First, they can be used to combine sets of any marginal distributions into an appropriate joint distribution. Second, they have a natural interpretation as proxies for unobserved covariates since they enter into the equations in the same way as observed covariates. The factor loadings can therefore be interpreted in much the same way as coefficients on observed covariates can. Third, they can provide a parsimonious representation of error correlations in models with large numbers of equations. We apply maximum simulated likelihood (MSL) techniques to estimate the parameters of our models. Simulation is used to evaluate integral expressions in the likelihood function of the model as no closed form solutions exist. MSL using pseudo-random draws can become very computationally expensive in applications in which there are large numbers of observations and parameters, and a large number of draws is required to reduce simulation error to acceptable levels. Consequently, for computational efficiency, we adapt an acceleration technique that uses quasi-random draws based on Halton sequences (Bhat, 2001; Train, 2003).

6. Our approach has numerous generic connections with the empirical microeconomic literature which are illustrated by the following selected examples. In models related to fertility, Olsen and Farkas (1989) examine the effect of childbirth on the hazard of dropping out of school while Carrasco (2001) examines the effect of childbirth on labor force participation of women in a panel-data context. In other contexts, Bingley and Walker (2001) examine the effect of duration of husbands' unemployment on wives' discrete labor supply choices, Jensen (1990) examines the effect of contraceptive use on duration between births, and Pitt and Rosensweig (1990) study the effect of endogenous health status of infant children on their mothers' main daily activity. These models share many statistical features. First, both treatment and outcome processes are non-normal and nonlinear: multinomial, count, discrete or censored. Second, the treatment is endogenous in each model. Finally, investigators often have a good a priori reasons for choosing particular (and uncontroversial) marginal models for both treatments and outcomes. But, the transition from given marginal distributions to a joint model for treatment and outcome continues to be a methodological challenge because typically nonnormal multivariate distributions are involved. In some cases, the marginal models have no (or very restrictive) tractable multivariate counterparts (e.g., in models of counts and durations). In others, treatment and outcome are from different statistical families (e.g., treatment being a multinomial and the outcome being a hazard rate) and so analytically tractable multivariate distributions often do not exist.

7. In addition to importance for theoretical and empirical insight into the effects of fertility on children's activities, this work has importance for evaluating policies aimed at reducing child labor. Consider, for example, the introduction of a subsidy for school attendance. This will reduce the cost of attending school and also the cost of raising children. Therefore, while the direct objective of such a policy would be to increase school attendance, it might also have the undesirable effect of increasing fertility. If there is a significant effect of fertility on child labor and schooling, this policy would have an affect on schooling directly and indirectly through fertility. Furthermore, correct information on the effects of fertility on child labor are important also to assess the importance of cross sectorial policies. For example, policies aimed to modify fertility behavior such as those which improve children's survival probabilities will affect the supply of child labor and school attendance as unintended consequences.

8. In what follows, we briefly describe a conceptual framework. Then we describe the econometric model in Section 3. We describe the data in Section 4 and results in Section 5. Finally we conclude.

2. CONCEPTUAL FRAMEWORK

9. Our econometric framework is motivated by a model of a household's decisions concerning fertility and time-use of their children. Children's time can be allocated between schooling, labor and and neither being in school not working significantly (we label this state as being idle). Although much of the literature on determinants of child labor treat school attendance as the only alternative to work (Jensen and Nielsen, 1997; Ray, 2000; Ravallion and Wodon, 2000), most survey data show that a substantial fraction of children neither attend school nor participate in work outside the home. As described earlier, in some cases, these children may be engaged in substantial household chores, which may not get reported as work. 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.

10. The theoretical model we have in mind is an overlapping generations, unitary model with equal concern for children (according to the definition of Behrman, 1997). We assume that adults in the household decide how many children to have and how to allocate their time across different activities across two periods. In the first period, adults choose the number of children to have. During this period, children are born and constitute a net burden for the household. Their time cannot be used in any productive activity. In the next period, when children reach a minimum of physical and psychological maturity (school age for simplicity), their time can be used to generate current income or to accumulate human capital that will increase their level of consumption in the future. In this second period, parents decide the allocations of time-use of their children. The main theoretical reason for assuming that adults choose the number of children in one period and the activities of their children in the next is that child mortality makes it uncertain whether a child will survive from birth to school age¹. Hence parents decide on the

¹ Child mortality from birth to age 5 is substantial in developing countries.

allocation of the time-use of their children once they observe how many children have survived from infancy to childhood. This model has a dynamic-programming structure and can be solved accordingly. A detailed description of this model is given in Cigno and Rosati (2002).

11. Three comparative statics results are noteworthy. First, an increase in the cost of education decreases the number of children, decreases the probability of school attendance and increases the probability of work, the latter two effects measured holding the number of children constant. Second, an increase in the returns to education increases the number of children, increases the probability of school attendance and decreases the probability of work conditional on the number of children. Finally, an increase in income or wealth causes increases in the number of children and the likelihood of school attendance conditional on the number of children and decreases the likelihood on work. In each of these cases, the overall effects on school attendance and work are ambiguous because families with more children are less likely to send their children to school.

3. ECONOMETRIC MODEL

12. The theoretical structure described above indicates a recursive econometric system of equations with the number of children entering the equation for children's activities.² We assume that decisions regarding children's activity are made when they are school-age (at least 7 years old). We model the activity of only the youngest child in each family.

13. In the case of families in which all children are 7 years of age or older, the youngest of these children can be given the interpretation of the "marginal" child. In the case of families in which some children are less than 7 years old (but not all), we assume that the survival of the younger children is uncertain so the marginal child for purposes of household decision-making regarding children's activities is the youngest child at least 7 years of age. In addition, to the extent that the youngest child in the relevant age group may be least entrenched in their possibly undesirable activity (working or being idle), this child may be the desired target of social policy.³

14. The child's activity is categorized into three classes - school, work, and neither school nor work (which we describe as being idle) - and is modeled as a multinomial logit process conditional on the number of children. Note that 0.71 percent of children report working and attending school. This frequency is too small to analyze as a separate category so we classify the activity of such children as working. Specifically, let the

² A recursive structure is also required for logical coherency in models with discrete and count out-comes (see Maddala, 1986 and Blundell and Smith, 1994)

³ This choice is not without statistical consequences. One may wish to model the activities of each of the children in the relevant age-group, thus absolving the analyst of making a choice of which child to consider. Because the activities of all children within the family share the same unobserved heterogeneity, programming and estimation of the model is considerably more complicated than the model described below. We have programmed such a model but deemed it infeasible when estimation times exceeded a week of CPU (as compared to less than 10 hours for the model described below).

random utility y_{ij}^* of the activity of the youngest child in household $i = 1, 2, \dots, N$ in activity $j = 0, 1, 2$ be described by

$$y_{ij}^* = \beta_{0j} + x_i \beta_{1j} + c_i \gamma_j + u_{ij} \lambda_j + \varepsilon_{ij} \quad (3.1)$$

15. The vector x_i denotes observed individual and exogenous household-specific covariates and c_i denotes the number of children. The u_{ij} describe household- and alternative-specific unobserved heterogeneity, e.g., parents' preferences for childrens activities, unobserved components of income and wealth, costs of and returns to education. The errors ε_{ij} follow i.i.d. extreme value distributions. Alternative j is chosen over alternative j' if $y_{ij}^* > y_{ij'}^*$. Let y_i be an indicator variable denoting the actual choice. Then, conditional on u_{ij} ,

$$\Pr(y_i = j | u_{ij}) = \frac{\exp(\beta_{0j} + x_i \beta_{1j} + c_i \gamma_j + u_{ij} \lambda_j)}{\sum_{j'=0}^2 \exp(\beta_{0j'} + x_i \beta_{1j'} + c_i \gamma_{j'} + u_{ij} \lambda_{j'})}, j = 0, 1, 2. \quad (3.2)$$

16. The number of children in a family is assumed to follow a Poisson process with mean specified as

$$E(c_i) = z_i \alpha + \sum_{j=1}^2 u_{ij} \delta_j, \quad (3.3)$$

where z_i are observed household-specific covariates that determine the number of children.⁴ Note that the u_{ij} which describe household- and activity-specific unobserved heterogeneity also enter (3.3). If unobserved heterogeneity is dominated by unobserved components of relative education prices, income and wealth, our conceptual framework suggests that λ_j and δ_j will have opposite signs.

17. We only include households with at least one child in the 7-15 age group in our sample so the stochastic process for fertility is truncated at one, i.e., conditional on u_{ij} ,

$$\Pr(c_i = c | c_i \geq 1, u_{ij}) = \frac{\exp(-\exp(z_i \alpha + \sum_j u_{ij} \delta_j)) (z_i \alpha + \sum_j u_{ij} \delta_j)^c}{c! [1 - \exp(-\exp(z_i \alpha + \sum_j u_{ij} \delta_j))]} \quad (3.4)$$

18. Conditional on u_{ij} , the joint probability of observing a certain number of children in a family and the activities of the children is given by the product of the marginal probability of observing the number of children and the conditional probability of the type of activity, i.e.,

$$\Pr(y_i = j, c_i = c | c_i > 0 | u_{ij'}) = \Pr(y_i = j | u_{ij'}) * \Pr(c_i = c | c_i > 0 | u_{ij'}) \quad (3.5)$$

⁴ A discussion of identification follows the data description below.

$j' = 0, 1, 2$. If u_{ij} were known, the error terms of the two processes could be treated as independent. But they are not. Instead we assume that each u_{ij} is drawn from a standard normal distribution with joint density denoted by $f(\mathbf{u})$. Then u_{ij} can be integrated out of (3.5) to form a likelihood function that will yield consistent parameter estimates. Specifically, the contribution of the i^{th} household to the log likelihood is given by

$$l_j = \ln \int_{-\infty}^{\infty} \left[\frac{\exp(\beta_{0j} + x_i \beta_{1j} + c_i \gamma_j + \sum_j u_{ij} \lambda_j)}{\sum_{j'=0}^2 \exp(\beta_{0j'} + x_i \beta_{1j'} + c_i \gamma_{j'} + \sum_j u_{ij} \lambda_{j'})} \times \frac{\exp(-\exp(z_i \alpha + \sum_j u_{ij} \delta_j))(z_i \alpha + \sum_j u_{ij} \delta_j)^c}{c! [1 - \exp(-\exp(z_i \alpha + \sum_j u_{ij} \delta_j))]} \right] \mathbf{f}(\mathbf{u}) d\mathbf{u}. \quad (3.6)$$

19. The main problem with estimation of the parameters of (3.6) is the fact that the integral does not have a closed form solution. However, this difficulty can be addressed using maximum simulated likelihood (MSL) (Gourieroux and Monfort, 1996). The key insight required for MSL is that the integral expression in (3.6) may be written as the expectation expression,

$$E \left[\frac{\exp(\beta_{0j} + x_i \beta_{1j} + c_i \gamma_j + \sum_j u_{ij} \lambda_j)}{\sum_{j'=0}^2 \exp(\beta_{0j'} + x_i \beta_{1j'} + c_i \gamma_{j'} + \sum_j u_{ij} \lambda_{j'})} \times \frac{\exp(-\exp(z_i \alpha + \sum_j u_{ij} \delta_j))(z_i \alpha + \sum_j u_{ij} \delta_j)^c}{c! [1 - \exp(-\exp(z_i \alpha + \sum_j u_{ij} \delta_j))]} \right] \quad (3.7)$$

where $\tilde{u}_{ij,s}$ is the s^{th} draw (from a total of S draws) of a pseudo-random number from the standard normal density. The justification for this is based on a law of large numbers and a central limit theorem.

Maximization of the simulated likelihood then proceeds in the standard manner and usual asymptotic formulae apply. We calculate standard errors using the robust, sandwich formula for the covariance matrix of parameters because it also correctly incorporates simulation noise (McFadden and Train, 1997) while other formulae do not. Marginal effects of covariates on the outcomes are also calculated by simulation. We calculate marginal effects for dummy variables as discrete changes and for continuous variables using derivatives. We calculate marginal effects at the means of all other covariates. Note that each of these calculations requires averaging over simulated draws of the latent factors.

Standard errors of the marginal effects are calculated using a Monte Carlo technique using 500 replications.

20. A number of normalization restrictions are necessary before estimation algorithms can be implemented. First, we apply the standard normalization for the multinomial logit model: $\beta_{00} = \beta_{10} = \gamma_0 = \lambda_0 = 0$. Then without loss of generality, we assume $u_{i0} = 0$. In our empirical work, we normalize on school attendance so that the u_{ij} , $j = 1, 2$ are unobserved determinants of child labor and idleness relative to school attendance. Second, because the variances of the unobserved factors cannot be identified, a normalization is required on either λ_j or δ_j . Without loss of generality, we assume $\delta_j = 1$.

21. Once these normalization restrictions are imposed, the model is identified via non-linearity even if the same variables determine fertility and child activity. However, identification via nonlinearity is not desirable. Instead, we include a set of variables in z_i that are excluded from x_i . In other words, we use a set of instruments that are assumed to affect fertility but which are assumed not to affect the choice of the child's activity directly, conditional on the covariates in the child's activities equations. Details of these variables and their justification is described below, after the sample of data is described.

3.1 Quasi-Monte Carlo simulation

22. Standard normal pseudo-random number generators are available in standard statistics computer packages. In cases of multidimensional expectations such as ours, a large number of pseudo-random draws is required to achieve suitable levels of accuracy. Increasing the number of simulation draws to reduce the simulation error to acceptable levels is simple in principle but computationally costly. In numerical analysis, a literature has recently emerged that attempts to use intelligent, systematic draws rather than random draws to speed up convergence of the required expectations. The quasi-Monte Carlo method is similar to the Monte Carlo method but instead of using s pseudo-random points, it uses non-random points within the domain of integration. We use Halton sequences to generate quasi-Monte Carlo variates following Bhat (2001).

23. Halton sequences have two desirable properties vis-a-vis pseudo-random points. First, they are designed to give even coverage over the domain of the mixing distribution. With more evenly spread draws for each observation, the simulated probabilities vary less over observations, relative to those calculated with random draws. Second, with Halton sequences, the draws for one observation tend to fill in the spaces left empty by the previous observations. The simulated probabilities are, therefore, negatively correlated over observations. This negative correlation reduces the variance in the simulated likelihood function. Consequently, under suitable regularity conditions, the integration error using pseudo-random sequences is in the order of N^{-1} as compared to pseudo-random sequences where the convergence rate is $N^{-1/2}$ (Bhat, 2001).

24. Bhat (2001) and Train (2003) report dramatic improvement in simulation errors from the use of Halton-sequence based draws relative to the usual pseudo-random draws. They both find that simulation errors using 100 Halton-based draws are smaller than

those using 1000 pseudo-random draws. Our limited experience in the context of the model considered here suggests less dramatic, yet substantial, improvement.

25. It is useful to describe a Halton sequence with an example (see Train, 2003, for an extensive and lucid description). Consider the prime number 2. Its Halton sequence is constructed as follows. Divide the unit interval (0,1) into 2 parts. The dividing point 1/2 becomes the first element of the Halton sequence. Next divide each part into two more parts. The dividing points, 1/4 and 3/4 become the next two elements of the sequence. Divide each of the four parts into two parts each, and continue. Halton sequences on non-prime numbers are not unique because the Halton sequence for a non-prime number divides the unit space in the same way as each of the prime numbers that constitute the non-prime.

26. In our model, we have two latent factors u_{i1} and u_{i2} that need to be integrated out. We begin by generating two Halton sequences based on the primes 2 and 3:

$$\begin{aligned}\xi_{i1} &= \{1/2 \ 1/4 \ 3/4 \ 1/8 \ 3/8 \ \dots\} \\ \xi_{i2} &= \{1/3 \ 2/3 \ 1/9 \ 2/9 \ 4/9 \ \dots\}\end{aligned}$$

27. The length of each sequence is determined by the number of observations N and the numbers of simulation draws S which we have chosen to be 2000. The early elements of Halton sequences with different primes have a tendency to be correlated with each other. Consequently, we begin by generating Halton sequences of length $N \times S + 20$ and discard the first twenty elements of each sequence. The required normally distributed quasi-random draws for l_{i1} and l_{i2} are generated by applying the inverse of the normal cumulative distribution function to the Halton sequences, i.e.

$$\begin{aligned}l_{i1} &= \Phi^{-1}(\xi_{i1}) \\ l_{i2} &= \Phi^{-1}(\xi_{i2}).\end{aligned}$$

28. The first group of s elements is assigned to the first observation in the sample, the next s elements to the second observation, and so on.

4. DATA

29. Our sample of households is drawn from the Human Development of India Survey (HDIS) conducted in India in 1994. The HDIS, which was carried out by the National Council of Applied Economic Research, is a multi-purpose, nationally representative sample survey of rural India. The sampling frame consists of 34,398 households living in 1,765 villages in 16 states. We only consider households with at least one child between the ages of 7 and 15 because child labor and schooling are both relevant in that age-group. We model the activity of only the youngest child (between the ages of 7 and 15) in

each family. Children are assumed to spend their time in one of three activities, attending school, working, or being idle. The number of children in the family is calculated from survey responses. It is defined as the number of children over 6 years of age alive at the time of the survey. Our sample consists of 12830 observations. Summary statistics for the variables used in the analysis are presented in Table 1.

30. The explanatory variables common to fertility and children's activity equations include the following household-level variables. Two variables proxy for the amount of resources available to the household: a dummy variable for the household being poor, (belonging to the lowest income quintile) and the number of appliances in the household. We have chosen to use the dummy variable for poverty status, 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 poverty status for most households.

31. Two additional variables may be thought of as additional proxies for household resources, although they also proxy for returns to work. These are whether the household owns land and whether the household owns livestock. Costs of education are proxied by three variables: two dummy variables indicating the presence of primary and secondary schools in the village and the village-average monetary cost of school. In addition, we include variables for the education of each parent; dummy variables for religion and a dummy variable for social status (designated by whether the household belonged to a 'scheduled caste' or 'scheduled tribe'. In order to control for common heterogeneity due to state policies and culture, we include a full set of state dummy variables. Finally, we include two child-specific variables, age and gender of the child, in the multinomial logit equations for the child's activities which are excluded from the fertility equation.

4.1 Instruments

32. Our model is nonlinear and structural, so in principle it is identified via functional form. However, as in most such models, it may be preferable to also have exclusion restrictions. We use a dummy variable for the gender of the first born child, ages of the parents and the village-level mortality rate as instruments in the exclusion-restriction sense, i.e., we assume that these variables directly affect the number of children in the family but do not directly affect the chosen activity of the youngest (eligible) child. The gender of the oldest child controls specifically for the well-known preference for male children in India. The ages of the parents control for the possibility that the observed number of children may not reflect completed fertility, or that there may be cohort effects in fertility. The village-level mortality rate is used as a proxy for the effects of expected child mortality on fertility decisions. Our informal checks suggest that these

instruments are valid in a statistical sense but arguments against their use can be made.⁵ In fact, Rosensweig and Wolpin (2000) present data suggesting that the sex composition of children has significant effects on child-rearing costs so they cannot identify female labor supply. Similar logic might apply to child labor as well. Lacking foolproof instruments and natural experiments, we instead check the robustness of our results to the choice of instruments by estimating our structural model with a variety of subsets of these potential instruments. Stability of estimates in such an exercise is the intuitive basis of the formal overidentifying restrictions tests in the linear context. The results of these models are described below following discussion of the main results.

5. RESULTS

33. As described earlier, the parameters of the fertility and child activities equations are jointly estimated using MSL. We begin by describing parameter estimates from the fertility equation. Then we describe estimates from the child activities equations. Finally, we describe the effects of changing instrument sets.

34. Parameter estimates of the truncated Poisson model for fertility are reported in Table 2. We find that households in which the oldest child is male have a fewer children. This supports the claim of gender bias in India. The ages of the parents have positive effects on the number of children, supporting the view that fertility may be incomplete for younger parents or that there may be cohort effects. The negative coefficient on mortality indicates that although a reduction in the mortality rate may reduce the number of births, it increases the number of surviving children in the household.

35. Households with more educated mothers have fewer children. Contrary to a priori expectation, we find that households with educated fathers have more children. We speculate the reason is because education of the father acts as a proxy for income. It would be consistent with our findings that relatively richer households and those who own livestock have more children. Note that these results are coherent with earlier findings on fertility using the same data (Atella and Rosati, 2000). Costs of education proxied by dummy variables for the existence of primary and secondary schools and the village-average monetary cost of education are not significant, perhaps because they are crude aggregate proxies and do not measure actual costs of education within households well. Our proxies for wealth, dummy variables for land and livestock ownership, are not statistically significant either. On the other hand, religion and social class have large and statistically significant effects on fertility.

36. Parameter estimates and marginal effects of the multinomial logit model for a child's activities are reported in Table 3. In order to evaluate the statistical and substantive impact of the endogeneity of fertility, we also estimated a multinomial logit model for a child's activities in which fertility was treated as if it was exogenous. These

⁵ There are no formal tests for validity of overidentification restrictions in the context of nonlinear models such as ours.

parameter estimates and marginal effects are reported in Table 4. The effect of fertility is substantially different once its endogeneity is taken into account. The coefficient on number of children in the equation for *idle* is statistically significant and of roughly the same magnitude in both cases, but the coefficient on *work*, which is negative and statistically insignificant in the exogenous case, is positive and statistically significant when endogeneity is taken into account. The factor loading parameter, λ , is not statistically significant in the *idle* equation, but is negative and highly significant in the *work* equation. The negative factor loading is indicative of a negative correlation between the unobserved components of the choice of work (relative to schooling) and the unobserved components of fertility choice. This is consistent with our a priori belief that the most likely sources of common unobserved heterogeneity between fertility and choice of the child's activities are household-level costs of education, returns to education and unobserved components of wealth, taken in the context of our conceptual framework. This finding is consistent with the results of Lundberg and Rose (2002) who find a negative correlation between the errors of fertility and female labor supply equations.

37. The marginal effects show that the youngest child is about 2 percentage points less likely to attend school, 1 percentage point more likely to work and 1 percentage point more likely to be idle in household with more children. Although these effect sizes may appear small at first glance, they are substantial. Consider, for example, that the presence of a primary school increases the probability of schooling by about 4 percentage points, decreases the probability of being idle by 4 percentage points and has no effect on child labor. Land ownership increases school attendance by 2.4 percentage points and decreases the probability of being idle by less than 2 percentage points. The effect of fertility on the likelihood of school attendance is twice as large when endogeneity of fertility is taken into account as compared to the effect obtained when exogeneity is assumed. When exogeneity of fertility is assumed, we observe that a child in a larger household is less likely to work, although this effect is only marginally significant. Once endogeneity of fertility is taken into account, we find that a child in a larger household is significantly more likely to work and the effect size is more than twice as large as in the exogenous case. The effect of fertility on the likelihood that the child is idle appears robust to whether fertility is assumed exogenous or not. Consequently, accounting for endogeneity does make a substantial qualitative and quantitative difference with the potential for very different policy recommendations.

38. Accounting for endogeneity of fertility has little effect on the effects of other exogenous covariates on a child's activities. Moreover, the effects have the expected signs and are of plausible magnitudes. Girls are more likely than boys to work and to be idle. Note that this may simply be a consequence of the fact that the work of girls may be more likely to be not treated as such. Children of better educated parents, and especially of better educated mothers, are more likely to attend school. The presence of a school, either primary or secondary, in the village does increase school attendance and reduces the probability that a child works or is idle. The same is true for the village-level cost of education. Ownership of both land and livestock increases the probability of attending school, but leaves almost unchanged that of working. Families with land and/or livestock are less likely to have idle children. The number of appliances appears, on the other hand, to reduce also the probability that a child works.

39. 5.1. Robustness to Changes in Instruments

40. We have described results from a specification which has four instruments - a dummy variable for the gender of the first born child, ages of the mother and father and village-level mortality. We have also estimated four alternative models using a variety of subsets of these instruments. The marginal effects of fertility on a the likelihood of a child's activity are reported in Table 5. A quick glance shows that the estimates are quite robust. Specifically, the largest deviation of the alternative estimates compared to the main specification is about 8%. The largest change in estimates across all specifications is about 10%. To the extent that the estimates are stable, they provide confidence in the plausibility of the results and the validity of the identification of the econometric model.

6. CONCLUSION

41. We use a nonlinear, simultaneous equations econometric model of fertility and children's activities to examine the causal effects of fertility on a child's activities taking the endogeneity of fertility into account. Our specification uses latent factors to allow for unobserved influences on fertility to affect a child's activities. Our a priori expectation was that unobserved heterogeneity included significant components of income and wealth, costs of and returns to education but that the direction and magnitudes of such effects was ambiguous.

42. We find that the effect of fertility is substantially different once its endogeneity is taken into account. Although the effects of fertility on the probability of being idle remains the same whether fertility is assumed to be exogenous or its endogeneity modeled, the probability of attending school is twice as large when endogeneity is modeled. The effect of fertility on the probability of work changes sign and becomes statistically significant when endogeneity of fertility is taken into account. Thus, we conclude that accounting for endogeneity is important and can point to very different policy recommendations.

REFERENCES

- Angrist, J. D. (2001), "Estimation of Limited-Dependent Variable Models with Dummy Endogenous Regressors: Simple Strategies for Empirical Practice", *Journal of Business and Economic Statistics*, 19, 2-16.
- Bhat, C. R. (2001) "Quasi-random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model", *Transportation Research: Part B*, 35, 677-693.
- Blundell, R. and R. J. Smith (1994), "Coherency and Estimation in Simultaneous Models with Censored or Qualitative Dependent Variables", *Journal of Econometrics*, 64, 355-373.
- Brownstone, D. and K. Train (1999), "Forecasting New Product Penetration with Flexible Substitution Patterns", *Journal of Econometrics*, 89, 109-129.
- Carrasco, R. (2001), "Binary Choice with Binary Endogenous Regressors in Panel Data: Estimating the Effect of Fertility on Female Labor Participation", *Journal of Business and Economic Statistics*, 19, 385-394.
- Cigno, A. and F. C. Rosati (2002), "Child Labour, Education and Nutrition in Rural India", *Pacific Economic Review*, 7, 65-83.
- Francesconi, M. (2002), "A Joint Dynamic Model of Fertility and Work of Married Women", *Journal of Labor Economics*, 20, 336-380.
- Gong, X. and A. van Soest (2002), "Family Structure and Female Labor Supply in Mexico City", *Journal of Human Resources*, XXXVII, 163-191.
- Gouriéroux, C. and A. Monfort (1996), *Simulation Based Econometrics Methods*, New York: Oxford University Press.
- Heckman, J. J., and E. Vytlačil (2001), "Policy-Relevant Treatment Effects", *American Economic Review Papers and Proceedings*, 91(2), 107-111.
- Hotz, V. J. and R. A. Miller (1988), "An Empirical Analysis of Life Cycle Fertility and Female Labor Supply", *Econometrica*, 56, 91-118.

Jensen, P. and H.S. Nielsen (1997), "Child Labour or School Attendance? Evidence from Zambia", *Journal of Population Economics*, 10, 407-424.

Lundberg, S. and E. Rose (2002), "The Effects of Sons and Daughters On Men's Labor Supply and Wages", *The Review of Economics and Statistics*, 84, 251-268.

Maddala, G. S. (1986), *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge: Econometric Society Monographs.

Maitra, P. and R. Ray (2002), "The Joint Estimation of Child Participation in Schooling and Employment: Comparative Evidence from Three Continents", *Oxford Development Studies*, 30, 41-62.

Moffitt, R. (1984), "Profiles of Fertility, Labour Supply and Wages of Married Women: A Complete Life-Cycle Model", *Review of Economic Studies*, 51, 263-278.

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.

Rosensweig, M. R. and K. I. Wolpin (2000), "Natural "Natural Experiments" in Economics", *Journal of Economic Literature*, 38, 827-874.

Train, K. (2003). *Discrete Choice Methods with Simulation*, New York: Cambridge University Press.

ANNEXES

Table 1. Summary statistics

Variable	Definition	Mean	Std. dev.
school	=1 if child is in school	0.703	0.457
work	=1 if child is working	0.086	0.280
idle	=1 if child is neither in school nor working	0.211	0.408
number of children	number of children	3.374	1.336
child is female	=1 if female	0.449	0.497
child's age	age in years/10	0.977	0.252
mother's education	education of mother	1.374	0.816
father's education	education of father	1.993	1.321
primary school in village	= 1 if primary school is in the village	0.457	0.498
secondary school in village	=1 if secondary school is in the village	0.441	0.497
cost of education	village mean of education cost in Rs./1000	0.660	0.577
household is poor	=1 if household income is in lowest quintile	0.559	0.497
household owns land	=1 if household owns land	0.669	0.471
household owns livestock	=1 if household owns livestock	0.693	0.461
number of appliances	number of appliances in household	0.912	1.168
household is hindu	= 1 if hindu	0.836	0.370
household is muslim household	=1 if muslim	0.106	0.307
household is Christian	=1 if christian	0.024	0.153
household is in low social class	= 1 if scheduled caste or tribe	0.370	0.483
first child is male	=1 if first child was male	0.669	0.470
mother's age	age of mother in years/10	3.663	0.881
father's age	age of father in years/10	4.136	1.130
mortality rate in village	village mean of child mortality rate	0.082	0.029

Notes:

Income quintiles are defined over urban and rural populations, although the sample consists of only rural households.

Education costs include expenditures on exams, fees, books, stationary and uniforms.

Table 2. Truncated Poisson Model for Number of Children

	parameter
first child is male	-0.060* (0.013)
mother's age	0.085* (0.008)
father's age	0.048* (0.007)
mortality rate in village	-0.627+ (0.371)
mother's education	-0.069* (0.009)
father's education	0.010* (0.005)
primary school in village	-0.031 (0.020)
secondary school in village	-0.018 (0.020)
cost of education	0.017 (0.011)
household is poor	-0.099* (0.012)
household owns land	0.002 (0.013)
household owns livestock	0.047* (0.013)
number of appliances	-0.000 (0.005)
household is hindu	-0.094* (0.036)
household is muslim	0.140* (0.040)
household is christian -	0.265* (0.055)
household is in low social class	0.041* (0.012)

The model is estimated jointly with multinomial logit equations for a child's activities. Asymptotic standard errors are in parentheses.

w The estimate is statistically significant at the 5 percent level.

+ The estimate is statistically significant at the 10 percent level.

Table 3. Multinomial Logit Model for a Child's Activity

	parameters			marginal effects	
	work	idle	school	work	idle
Number of children	0.208 ^w (0.083)	0.110 ^w (0.033)	-0.019 ^w (0.004)	0.008 ^w (0.004)	0.011 ^w (0.004)
Child is female	1.263 ^w (0.103)	0.719 ^w (0.051)	-0.126 ^w (0.008)	0.054 ^w (0.005)	0.072 ^w (0.006)
Child's age	-0.920 (1.552)	-9.720 ^w (1.019)	1.041 ^w (0.137)	0.016 (0.065)	-1.056 ^w (0.118)
Child's age ²	2.519 ^w (0.730)	4.161 ^w (0.490)	-0.526 ^w (0.064)	0.086 ^w (0.031)	0.440 ^w (0.057)
Mother's education	-0.759 ^w (0.119)	-0.687 ^w (0.080)	0.100 ^w (0.009)	-0.029 ^w (0.005)	-0.071 ^w (0.008)
Father's education	-0.419 ^w (0.053)	-0.461 ^w (0.029)	0.064 ^w (0.004)	-0.016 ^w (0.002)	-0.048 ^w (0.003)
Primary school	-0.087 (0.159)	-0.353 ^w (0.089)	0.039 ^w (0.012)	-0.002 (0.007)	-0.038 ^w (0.010)
Secondary school	-0.295 ⁺ (0.161)	-0.527 ^w (0.092)	0.065 ^w (0.012)	-0.010 (0.007)	-0.055 ^w (0.010)
Cost of education	0.185 ^w (0.082)	0.201 ^w (0.053)	-0.028 ^w (0.007)	0.007 ⁺ (0.004)	0.021 ^w (0.006)
Household is poor	0.005 (0.095)	0.105 ⁺ (0.060)	-0.011 (0.008)	-0.000 (0.004)	0.011 ⁺ (0.006)
Household owns land	-0.159 (0.100)	-0.170 ^w (0.060)	0.024 ^w (0.008)	-0.006 (0.004)	-0.018 ^w (0.007)
Household owns livestock	0.025 (0.100)	-0.135 ^w (0.060)	0.013 ⁺ (0.008)	0.002 (0.004)	-0.015 ^w (0.006)
Number of appliance	-0.414 ^w (0.054)	-0.406 ^w (0.035)	0.058 ^w (0.005)	-0.016 ^w (0.003)	-0.042 ^w (0.004)
Household is hindu	-0.518 ⁺ (0.276)	0.171 (0.208)	0.006 (0.025)	-0.027 ⁺ (0.015)	0.021 (0.020)
Household is muslim	-0.027 (0.314)	0.900 ^w (0.221)	-0.119 ^w (0.038)	-0.008 (0.012)	0.127 ^w (0.038)
Household is Christian	-1.373 ^w (0.529)	0.069 (0.366)	0.025 (0.043)	-0.037 ^w (0.009)	0.013 (0.042)
Household is in low social class	0.274 ^w (0.087)	0.338 ^w (0.055)	-0.047 ^w (0.007)	0.010 ^w (0.004)	0.036 ^w (0.007)
λ	-1.256 ^w (0.407)	-0.096 (0.164)			

The model is estimated jointly with a truncated Poisson equation for fertility. Asymptotic standard errors are in parentheses.

^w The estimate is statistically significant at the 5 percent level.

⁺ The estimate is statistically significant at the 10 percent level.

Table 4. Multinomial Logit Model for Child's Activities Assuming Exogeneity of Fertility

number of children	-0.026 (0.031)	0.093 ^w (0.020)	-0.009 ^w (0.003)	-0.003 ⁺ (0.002)	0.012 ^w (0.002)
child is female	1.162 ^w (0.075)	0.720 ^w (0.051)	-0.134 ^w (0.007)	0.062 ^w (0.005)	0.072 ^w (0.006)
child's age	-0.675 (1.382)	-9.721 ^w (1.013)	1.106 ^w (0.141)	0.116 (0.092)	-1.221 ^w (0.120)
child's age ²	2.207 ^w (0.623)	4.171 ^w (0.487)	-0.565 ^w (0.065)	0.072 ⁺ (0.038)	0.493 ^w (0.061)
mother's education	-0.720 ^w (0.102)	-0.692 ^w (0.080)	0.110 ^w (0.011)	-0.034 ^w (0.006)	-0.076 ^w (0.010)
father's education	-0.383 ^w (0.043)	-0.461 ^w (0.029)	0.069 ^w (0.004)	-0.017 ^w (0.003)	-0.052 ^w (0.004)
primary school	-0.103 (0.142)	-0.348 ^w (0.089)	0.043 ^w (0.013)	-0.001 (0.009)	-0.042 ^w (0.011)
secondary school	-0.295 ^w (0.144)	-0.523 ^w (0.092)	0.072 ^w (0.013)	-0.010 (0.009)	-0.061 ^w (0.011)
cost of education	0.171 ^w (0.072)	0.201 ^w (0.053)	-0.030 ^w (0.007)	0.008 (0.005)	0.023 ^w (0.006)
household is poor	-0.054 (0.083)	0.099 ⁺ (0.059)	-0.008 (0.008)	-0.005 (0.005)	0.013 ⁺ (0.007)
household owns land	-0.155 ⁺ (0.090)	-0.172 ^w (0.060)	0.026 ^w (0.008)	-0.007 (0.006)	-0.019 ^w (0.008)
household owns livestock	0.048 (0.089)	-0.129 ^w (0.059)	0.012 (0.009)	0.005 (0.006)	-0.017 ^w (0.007)
number of appliances	-0.382 ^w (0.045)	-0.406 ^w (0.035)	0.063 ^w (0.005)	-0.018 ^w (0.003)	-0.045 ^w (0.004)
household is hindu	-0.410 ⁺ (0.248)	0.121 (0.205)	0.006 (0.027)	-0.028 ⁺ (0.016)	0.022 (0.026)
household is muslim	0.196 (0.276)	0.854 ^w (0.218)	-0.104 ^w (0.029)	-0.001 (0.017)	0.105 ^w (0.027)
household is christian	-1.273 ^w (0.479)	-0.011 (0.364)	0.061 (0.050)	-0.081 ^w (0.029)	0.019 (0.047)
household is in low social class	0.280 ^w (0.078)	0.337 ^w (0.054)	-0.050 ^w (0.008)	0.012 ^w (0.005)	0.038 ^w (0.008)

Asymptotic standard errors are in parentheses.

^w The estimate is statistically significant at the 5 percent level. ⁺ The estimate is statistically significant at the 10 percent level.

Table 5. Marginal Effects of Number of Children with Alternative Instruments

Instruments	School	Work	Idle
first child is male, mother's age, father's age, village mortality rate	-0.0193 ^w (0.0045)	0.0085 ^w (0.0036)	0.0108 ^w (0.0036)
first child is male, village mortality rate	-0.0189 ^w (0.0051)	0.0087 ^w (0.0048)	0.0102 ^w (0.0038)
first child is male	-0.0187 ^w (0.0053)	0.0086 ^w (0.0048)	0.0101 ^w (0.0040)
village mortality rate	-0.0185 ^w (0.0053)	0.0086 ^w (0.0048)	0.0099 ^w (0.0040)
mother's age, father's age	-0.0193 ^w (0.0046)	0.0084 ^w (0.0038)	0.0109 ^w (0.0036)

Notes:

Asymptotic standard errors are in parentheses.

W The estimate is statistically significant at the 5 percent level. + The estimate is statistically significant at the 10 percent level.