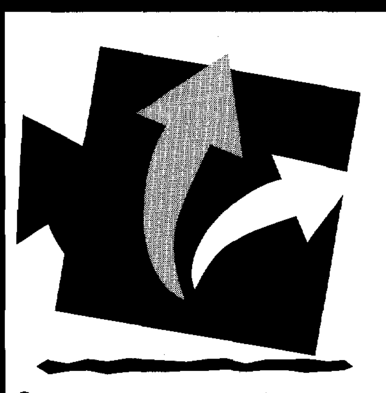
SP DISCUSSION PAPER NO. 9931

21334



Uncertainty About Children's Survival and Fertility: A Test Using Indian Microdata

Vincenzo Atella
Furio Camillo Rosati

December 1999

Social Protection
LABOR MARKETS, PENSIONS, SOCIAL ASSISTANCE
THE WORLD BANK

Uncertainty About Children's Survival and Fertility: A Test Using Indian Microdata

by

Vincenzo Atella
SEFEMEQ - Department of Economics
University of Rome II - Tor Vergata

Furio Camillo Rosati
CEIS – Department of Economics
University of Rome II - Tor Vergata

Corresponding author: Prof. Furio Camillo Rosati, Via San Nemesio 1, 00145 Roma Tel. 6.51602500 – Fax 6.2020500 - E-mail: f.rosati@agora.stm.it

Summary

Parents do not generally decide to send their children to work when a child reaches a critical age: this decision is largely made as part of a dynamic process, which starts at the time of household formation, and is shaped by expectations of the role of children as future assets. However, children's mortality is likely to influence the behaviour of households with respect to childbirth, not only through the average expected number of surviving children, but also through the uncertainty (or risk) about the realization of such expectations. Changes in the mean and variance of the survival rate can therefore affect parental choices. The majority of the current empirical analyses at the micro level have been based only on one dimension of mortality (its mean) and suggests a positive relationship between mortality and fertility. The use of only this dimension might not be sufficient to describe the relationship between fertility and mortality and subsequently, child labor. This paper by Atella and Rosati gives attention to these interlinkages and fills in a gap in the existing literature. With the use of individual microdata for India, which include effects of the so-called "community" variables, Atella and Rosati present some interesting policy implications. For example, an improvement in general sanitary conditions can lead to a reduction of the fertility rate via an increase in the mean survival rate. But, if an increase in the expected probability of survival is also associated with a reduction of its variance, there are two contrasting effects. On the one hand, the reduction of mortality will lead to a reduction of fertility. On the other hand, the reduction in the risk associated with each child will generate incentives to higher fertility. This is an important paper for broadening the analysis of decisions leading to child labor, which is the task of future research of the World Bank's Global Child Labor Program.

Abstract

In this paper we present a non altruistic model of demand for children in the presence of uncertainty about children's survival. Children are seen as assets, as they provide help during old age. If certain conditions are met, both the financial market and the family network are used to transfer resources to old age. Theoretical predictions relative to the change in the mean and variance of the survival rate are derived. The empirical analysis is based on data from the Human Development of India (HDI) survey. Different models for count data variables, such as Poisson, Hurdle and ZI models have been employed in the empirical analysis. The results highlight the importance of the uncertainty about children's survival in determining parental choices, thus showing that realized or expected children's death is not the only dimension that links fertility decision to children's mortality. The policy implications of such findings are briefly discussed.

1. Introduction

Children's mortality is likely to influence the behaviour of households with respect to childbirth not only through the average expected number of surviving children, but also through the uncertainty (or riskiness if we assume households know the probability distribution) about the realization of such expectations.

Such an issue has seldom received attention in the fertility literature and, to our knowledge, its empirical implications has never been tested. In a recent paper — Rosati (1996) — the implications for saving and fertility of a social security system has been analyzed within the framework of a non altruistic intergenerational model of family relationship characterized by the presence of uncertainty about the children's survival. Here we retain the same approach and extend the theoretical model to derive the effects on fertility of a change of the expected average survival rate and of its variance. This set-up allows us to derive some testable implications of the theory that will be used for the empirical analysis.

Our main objective, is to assess the importance of the expected survival rate and of the uncertainty about children's survival in fertility decisions.

The available empirical evidence, at both micro and macro level, seems to indicate the existence of a positive relationship between mortality and fertility. The direction of causality is, however, open to discussion (see Chowdhury (1988)), and it appears that joint determination is likely. Also, the size of the "replacement" effect is open to discussion. According to Olsen (1980), the dimension of such an effect obtained by dealing with the endogeneity of infant mortality appears to be smaller than previously thought.

The majority of the empirical analyses, also those at a micro level, have been based only on one dimension of mortality (mainly its average value or its realization). However, the use of only one dimension might not be sufficient to describe the relationship between fertility and mortality. The existence of more complex links between the two variables has also interesting policy implications that will be discussed in the paper.

Finally, in recent empirical studies on fertility (as well as on schooling, mortality, etc.) based on microdata, great attention has been paid to the so-called "community" variable. While their empirical relevance appears confirmed in many studies, the theoretical reasons for their introduction in the analysis have seldom been discussed. The present paper offers a theoretical foundation for the use of such variables.

The estimates have been carried out using techniques for count data models. In particular, Poisson, Negative Binomials, Hurdle and Zero Inflated (ZI) models have been employed and the selection among them has been based on a testing procedure. As we will see, the Hurdle-Poisson and the ZI Poisson models turn out to be the most appropriate for the data considered. Moreover, interesting differences emerge among the coefficients estimated in the two stages, indicating how relevant information on fertility behaviour can be obtained by the appropriate use of such models.

The paper is organized as follows. Sections 2 and 3 present the non altruistic model and derive some comparative static results. Section 4 reviews briefly the data-set employed for our empirical analysis and presents the econometric strategy used to derive the empirical results. Section 5 is devoted to present the empirical analysis. Finally, in section 6 we discuss the policy implications of these findings and draw some conclusions.

2. Uncertainty and Reproductive Decision

We consider a three period overlapping generation model. Individuals earn their income during the second period of life when they also take and realize their reproductive decisions. In the first and in the last period of life, consumption is financed, not necessarily fully, by inter-generational transfers. In particular we show that, under certain conditions, the middle-aged make transfers to both the young (children) and the old (parents). The former can be seen as a way of transferring resources to the future in order to secure consumption when old, while the latter can be seen as repayment to the parents of the "loan" received when young.

We also assume that a self-enforcing family constitution is in place (Cigno 1993), so that children will not voluntarily default on their repayments. Enforceability can be guaranteed, for example, if any member of any generation adopts a strategy according to which he/she will comply with the family "rules", i.e., repay his debt at the family interest rate when middle-aged, if and only if his parents repaid their own debt. Default from the children is, then, the threat that the parents have to face if they disobey the family rules. Note that such a threat is credible as it is in the threatener's interest to carry it out. The set of all such strategies is a Nash-equilibrium, because anybody's best response to the threat of punishment is to comply with the rules and, at the same time, to threaten to punish anyone who does not comply. This equilibrium is perfect in the sub-games[1] because it is clearly in a person's interest not to pay back a loan where that can be done without fear of punishment.

As children are seen by parents as an asset, reproductive decisions are determined as the result of the joint decision of the middle-aged concerning the number of assets to buy (children to raise) and the amount to invest in (lend to) each one. It is important to distinguish between the cost of having a child and the transfer made to him. The latter represents the amount of resources that the parent makes available for the consumption of the young and that the young pays back to the old when middle-aged, the former is an unrecoverable cost made up of parental time, health hazards and the like. Here we assume that there is a fixed, unrecoverable cost, c, that the parents will bear for each child they have.

If there were no uncertainty, given the non altruistic setting, in our model parents would have just one child and lend to him the whole amount of resources they are willing to transfer to the future. Nevertheless, even if the risk of a voluntary default is absent because of the self-enforcing family constitution, in order to be able to repay their "debt" to the parents, children have to become middle-aged and be able to work.

3. Fertility, Intergenerational Transfers and Mortality

In all the situations described above, putting all of one's resources in one child (asset), who has positive probability of not surviving to middle-age, might not be optimal. The optimal number of children obviously depends on the specific characteristics of the "uncertainty" considered.

Here we assume that each child has a probability of surviving to middle age equal to p, and of dying when young equal to $(1-p)$. Such probabilities are considered exogenous.

Let s denote the number of surviving children. This number is unknown at the moment of the decision on fertility and intergenerational transfers. Then:

[1] Assuming that each family member can observe the behavior of the others.

$E(s) = np$ and $Var(s) = np(1 - p)$

We assume that p is unknown to the individual. Let p be a random variable with mean π and variance v. Then, it follows by the law of iterated expectations that the number of surviving children s has mean.[2]

$E(s) = (E(s/p)) = n\pi$

and variance

$$Var(s) = Var(E(s/p)) + E(Var(s/p)) = n(n - 1)v + n\pi(1 - \pi) \qquad (1)$$

Each member of generation t receives income, y^t, when middle-aged. He can then use it for current consumption, C_2^t, and for transfers to his n^t children; b^t is the transfer per child. Part of the income, however, has to be used to repay the transfers received from his parents, $b^{t-1}\rho$. Old people do not get any income, but finance their old-age consumption, C_3^t, out of transfers from their children and from savings invested in the financial market. For each child the parents have to face an unrecoverable cost of c. In the financial market middle-aged individuals can buy a risk-free asset that earns a return given by an interest factor i. Children and bonds are, therefore, competing assets in the family portfolio: an individual chooses the portfolio composition and size that maximize his lifetime utility.

An individual maximizes the following expected utility function, with respect to n, x, and A:

$$V = U_2(C_2) + U_3(E(C_3), Var(C_3)) \qquad (2)$$

subject to

$C_2 = Y^t - A^t - b^{t-1}\rho - cn$,

$C_3 = xA^t\rho + (1 - x)A^t i$

and $x \geq 0$,

where A indicates the total amount of resources transferred from middle to old age and x is the share of the risky asset, i.e. of total transfers to children, in the portfolio. Notice that once n, A, and x are determined, transfers per child are implicitly defined by $xA = bn$.

---

[2] We thank the editor for suggesting us how to treat the uncertainty about the probability of survival.

The expectation of C_3 is given therefore by:

$$E(C_3) = E(b\rho s) + (1-x)Ai = b\rho n\pi + (1-x)Ai,$$

as the market offers a risk free asset.

The variance of C_3, making use of (1), is given by:

$$Var(C_3) = b^2\rho^2 Var(s) = b^2\rho^2\sigma$$

where $\sigma = n(n-1)v + n\pi(1-\pi)$.

As $b = \frac{Ax}{n}$, we have $Var(C_3) = A^2x^2\rho^2\frac{\sigma}{n^2}$.

Substituting back into eq. (2) for the first two moments and maximizing with respect to n, x, and A, we get the following first order conditions:

$$n: \frac{\partial Var(C_3)}{\partial n} U_3^{\sigma} - cU_2' = 0 \tag{3}$$

$$x: \frac{\partial Var(C_3)}{\partial x} U_3^{\sigma} + A(i-r)U_3' = 0 \tag{4}$$

$$A: \frac{\partial Var(C_3)}{\partial A} U_3^{\sigma} - U_2' + U_3'(xr + (1-x)i) = 0 \tag{5}$$

where U_3^{σ} indicates the derivative of the utility function with respect to the variance of future consumption. U_3^{σ} is assumed to be negative. r indicates the expected returns from intergenerational transfers, $r = \rho\pi$.

Eq. (3) has an interior solution only if $\frac{\partial Var(C_3)}{\partial n}$ is negative. If an increase in the number of children increases the variance of future consumption, individuals would not have any interest in diversifying their "portfolio", i.e. in having a number of children higher than one (if they decide to use the family network at all).

$\frac{\partial Var(C_3)}{\partial n}$ is negative if $v - \pi(1 - \pi) < 0$. In words, the model has an interior solution if the uncertainty of the individual about value of the probability of survival, p, is smaller than the expected variation in the survival rate. In what follows we assume that the above condition is met.

Eq. (3) gives the optimal number of children as a decreasing function of the unrecoverable cost of having a child, and as a positive function of the variance of future consumption, for a given level of A and x (i.e. for a given level of transfers to old age through intergenerational transfers). These are trade-off obtained holding the other endogenous variables constant. The results of the comparative statics will be discussed below. Observe that in order to have an interior solution, c, the unrecoverable cost of each child, must be positive. In fact as children are demanded as "insurance" against the uncertainty of future income, if there were no unrecoverable cost individuals would have the largest possible number of children (provided $\frac{\partial Var(C_3)}{\partial n} < 0$).

The first order conditions for x (eq. 4) can be rewritten as:

$$r - i = -\frac{U_3^{\sigma}\frac{\partial Var(C_3)}{\partial x}}{U_3^{'}A} \tag{6}$$

As $U_3^{\sigma} < 0$, the right-hand side of (6) is positive. An interior solution for x requires, therefore, the expected returns to intergenerational transfers to exceed the market interest rate in order to compensate for the risk associated with making transfers to children.[3]

Finally eq. (5) is the intertemporal arbitrage condition defined over the average returns from the portfolio selected. Such an arbitrage condition takes also in to consideration the increase in the variance of old age consumption due to an increase in total transfers to old age.

On the basis of eq. (3)-(5) the comparative statics effects were computed for changes in income, returns to intergenerational transfers and in the variance of the children's survival probability. A change in the returns to intergenerational transfers has obviously the same effects as a change in the expected survival probability.

In general the results from the comparative statics exercise are ambiguous. However, one of the crucial variables in determining the sign of the comparative statics is the degree of relative risk aversion. As the algebra is very cumbersome it is omitted from the paper and only the main results are discussed. In particular, if the degree of relative risk aversion is

[3] Note that (6) is a necessary but not sufficient condition for the existence of a family transfer system. Even if (6) holds, middle aged individuals could still consider leaving the family system, as the credit market offers an alternative way to finance old age consumption. Refer to Rosati (1996) for a discussion of this point.

relatively "low", then the effects of a change in the exogenous variables are likely to have the sign presented in Table 1.

An increase in income when middle-aged will tend to shift the portfolio of the family towards the risky asset, i.e. the inter-generational transfers. This implies an increase in the amount of transfers per child and/or an increase in the number of children. If individuals have a relative "low" degree of risk aversion, they will reduce the number of children and increase total transfers, A. When income in the second period of life increases, the individual is willing to transfer part of it to old age. If the relative degree of risk aversion is low, people will prefer to increase both the variance and the expected value of old age consumption. In fact, by reducing the number of children further, resources are freed for consumption, both present and future, and part of these resources is added to intertemporal transfers. Then the amount of transfers per child will unambiguously increase. If, on the contrary, individuals are "very" averse to risk, they will use the extra resource available, both to increase the mean of future consumption and to reduce its variance. In this case n will increase, while total transfer to old age A will decrease: consumption when old will increase, but less that in the previous case. The effect on the amount of transfers per child is ambiguous as both n and x increase.

An individual can react to a change in the variance of future consumption through two channels. On the one hand he can increase the demand for children (eq. 3), on the other he can reduce the share of the risky asset (i.e., of the intergenerational transfers) in its portfolio (eq. 4). If the degree of risk aversion is not "too" high, the individual will reduce the number of children and the share of the risky assets in the portfolio. The effect on total transfers to old age is ambiguous, but is likely to be positive. The reduction in n will leave more resources

Table 1
Comparative Static Results
When the degree of risk aversion is low (sufficient condition)

	n	X	A
y	-	+	+
v	-	+	?-
r	?	+	-

available for old-age consumption. On the other hand, the larger amount of resources invested in the market will reduce the expected return from postponed consumption, but also its variance.

If there are no financial markets or other instruments for transferring resources to old age, an increase of the variance of old age consumption will increase n. In the model utilized here, an increase of the variance of the survival rate will make intergenerational transfers less attractive as a way of transferring resources to old age relative to the other instruments. Hence the individual will react by having less children and moving towards the "market". If no "market" is available, then in order to reduce the riskiness of its old age consumption, the individual has no other choice but to devote part of his middle-age resources to buy more insurance, i.e., to increase the number of children.

An increase in the expected returns from intergenerational transfers will most likely decrease the amount transferred to old age (A). This does not imply, however, a reduction of old age consumption, as the return to intertemporal transfers are now higher. The share of the risky assets in the portfolio (intra-family transfers) will also rise, if risk aversion is "high". The individual will hence compensate the increase in the overall riskiness of his portfolio by increasing the number of children. If, on the other hand, the degree of risk aversion is "low", we might observe a reduction in the number of desired children. In this case, some of the increase in old age resources generated by the increase in x and ρ, will be transferred to middle age by means of the reduction in the fixed expenditure for children due to the smaller number of children desired.

4. Estimates of the Fertility Equations

The model presented in Section 3 could be easily extended to accommodate more explicitly for the role of other variables known both theoretically and empirically to influence the desired number of children. Male and female wage rates, sex preferences and the like. However, no new insights will be gained by explicitly introducing such variables into the theoretical model. Hence we will take such variables into account in the empirical analysis, without extending our theoretical model to include them.

We will therefore use in the estimates a reduced form fertility equation derived from eq. (3)-(5) that includes also a set of other explanatory variables like male and female wage rate, age, and other relevant variables.

4.1 The data set

All data are from the Human Development of India Survey carried out by the NCAER of New Delhi. This is a multi-model nationally representative sample survey conducted during 1994 in the rural areas of India. A two stage stratified and partially self weighting design was used to sample a total of 34,398 households spread over 1,765 villages and 195 districts in 16 states. Two separate survey instruments were devised, 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. The data are representative at the level of all India, according to States and population groups and according to selected population groups for the selected States.

Information on reproductive behavior are contained in a special part of the survey and constitutes the basis of the data set utilized for the estimates. Information about the husband and household level variables were added by merging the "reproductive information" file with the individual and household files.

4.2 Definition of the variables

The dependent variable is defined as the total number of live-births had by a woman of age 13 to 49.[4] The explanatory variables include the age of the wife, the age of the husband, various measures of income and wealth, dummies for the level of education of both wife and husband, dummies for the religion of the head of the household, the number of children dead up to the age of 5, male and female wage rates and the mean and variance of the children's survival rate. All these variables, with the exclusion of the last four are measured at the individual or, where relevant, at the household level and their definition does not pose any special problem (see Table 2 for summary statistics).

Rather than fitting a wage equation, we have decided to employ the average male and female wage rates observed at village level. Most of the people in the survey are farmers.

[4] In the fertility literature and in the official statistics the "period at risk" is considered from 15 to 44 years. The distribution of age in this sample of women is such that we have only 2.08% of women below 18 years and

Only a small fraction of them spend some time in market activities. The number of women involved in activities external to the household is even smaller. The relative wage at village level appears, in such a situation, to offer a better measure of the relative value of the time of man and women in the market.

To measure the mean and the variance of the survival rate expected by the household is rather demanding. Ideally we would like to have the information on the expectation held by the couples at the beginning of their reproductive activity and on the revision of such expectation during the fertile period. Obviously, this would require a large panel data set that is not available. Instead, we decided to exploit the cross-sectional variation in the data at village level and used it as proxy for the expectation of the variance and of the mean of the survival rate. In particular, we computed the mean and the variance of the survival rate (up to the age of 5) at village level from the individual data available for each village. If the cross-village differences have not changed widely over time, then our proxies should be a relatively good measure of inter-village differences in (expected) survival rates. As for some villages there are only few individual observations, we also used the mean and the variance computed at district level to check the robustness of our results.

4.3 Econometric methodology used for the estimation

The empirical literature on count data models is based on Poisson Models. According to Poisson models, the number of events in an interval (0,t) is a stochastic process with independent increments in non overlapping time periods. One of the main drawback of the Poisson model is that it postulate the equality between mean and variance (equidispersion). From the empirical point of view the equidispersion is rarely observed, thus rendering the Poisson model frequently unsatisfactory or inappropriate. According to Gurmu and Trivedi (1996) the "*excess of zero*" or the "*zero inflation*" problem from one side and the "*failure of conditional independence assumption*" from the other side are the main causes for experiencing under and over-dispersion.

It is for this reason that less restrictive models have been developed in order to overcome the lack of equality between mean and variance. Among them, Negative Binomial models are the most known and used in the recent empirical literature when overdispersion in

3.29% above 45 years. Utilizing a sample of women aged 15-41 does not affect the result in any significant way.

the data occurs. Hurdle models and "Zero Inflated" or "Zero Altered" models can represent a valid alternative to Negative Binomial models when under-dispersion is detected (Winkelmann, 1995).

Hurdle and zero inflated models are able to handle zeros in a way that allow decomposing the decision processes underlying (a) the choice to have children at all and (b) the choice about the number of children, once decided to have children. Having this distinction in mind, it is extremely useful to follow Shonkwiler and Shaw (1996) in their derivation of the Single-Hurdle model.

Let us define y_i as the number of births had by a woman. According to our previous classification, our population consists of women who had at least one child, and women who never had children (but we may or may not know if these women will some time in the future want to have children). Further, let D_i represent the latent decision variable to have a child; it follows that if $D_i \leq 0$, then it means that there are some impediments to having a child. Specifically, we adopt the discrete specification Prob(D_i = 0) = exp($-\theta_i$) where θ_i can be parametrized as

$$\theta_i = \exp(z_i'\gamma),$$

and γ is an unknown vector of parameters.

If the number of children is positive, then,

$$E(y_i) = \lambda_i = \exp(\Sigma \gamma_i z_i),$$

where z is the vector of relevant exogenous variables, γ is its corresponding vector of parameters.

The sequential (single-hurdle) model can be represented, following Winkelmann and Zimmermann (1995), as

$$P(Y=0) = f_1(0)$$

$$P(Y=y) = f_2(y)\cdot[1-f_1(0)] / [1-f_2(0)] = \Phi\cdot f_2(y), \qquad \text{for } y = 1, 2, \ldots..$$

where f_1 and f_2 are any probability distribution functions for nonnegative integers and where f_1 governs the hurdle part and f_2 the parent process.

It then follows that the likelihood function for the single-hurdle model with Poisson parent function specification is (Mullahy, 1986):

$$L_{SH} = \frac{\prod_{y=0} \exp(-\theta_i) \prod_{y>0} (1-\exp(-\theta_i)) \exp(-\lambda_i) \lambda_i^{y_i}}{[(1-\exp(-\lambda_i)) y_i!]}$$

This formulation is attractive because if $\theta_i = \lambda_i$, then the Single-Hurdle collapses to a simple Poisson specification.

For the Single-Hurdle model it follows that the expected values are given by:

$$E(y_i \mid y_i > 0) = \frac{\lambda_i}{1 - e^{-\lambda_i}}$$

and

$$E(y_i) = \frac{\lambda_i (1 - e^{-\theta_i})}{1 - e^{-\lambda_i}}$$

Finally, a different approach to modeling differently the two stage processes is the one proposed by Mullahy (1986) and by Lambert (1992). Zero Inflated count model can usually be represented in the following way:

$$\begin{cases} \pi + (1-\pi) h(y_i, \delta), & y_i = 0 \\ (1-\pi) h(y_i, \delta), & y_i = 1,2,3,\ldots \end{cases}$$

Zero Inflated Poisson Models and Zero Inflated Negative Binomial models can be obtained by simply replacing the $h(y_i, \delta)$ with the Poisson and Negative Binomial distribution.

5. Empirical results

5.1 The Econometric Design

Our strategy has been to move from the Poisson regression model to the less restrictive count data models discussed in the previous section. The Poisson regression model has been employed in order to have a reference estimate. We have also run the same Poisson models assuming the existence of misspecifications. Although consistency is guaranteed with the Poisson model, the asymptotic covariance matrix differs and this causes the standard errors to be different. In Table 3, for each model we have reported t-stat values using Pseudo Maximum Likelihood (PMLE), assuming a constant variance-mean ratio. Even if different from the MLE estimates, PMLE t-stat values do not change at all the significance of the parameters.

We then run three tests on the equality between mean and variance, to see if the Poisson model could be considered as the true data generating process for our count variable. The first two tests are from Lee (1986). Both (Lee Test1 and Lee Test2) tests the equality of variance and mean. Under H_0 (that mean and variance are equal) they are distributed as $N(0,1)$. Rejection indicates that the Poisson assumption is invalid.

The third test is from Wooldridge (1996), and is a regression-based test in which the dependent variable (standardized residuals -1) is regressed on the fitted values. The significance of the t-statistics leads to the rejection of the Poisson restriction, and a positive (negative) coefficient indicates over-dispersion (under-dispersion). The results of these tests are reported on the bottom of Table 3. According to our estimates, all tests reject the Poisson model and the Wooldridge test indicates the existence of under-dispersion.

On the basis of these results, the next step has been to estimate Hurdle and ZI models[5,6] with Poisson specification.

5.2 Results of the estimates

Overall, the results obtained are quite good and in line with the theory exposed in the previous paragraphs. All the results are presented in Tables 3 and 4.

We will limit our discussion to the Hurdle-Poisson models (Table 4) as they have the best statistical properties. In particular, we first discuss the results of the second stage and, then, highlight the differences with respect to the first stage.

The second stage of the Hurdle models represents the decision to have a certain number of children once decided to have children. An important role in this decision process is played by the age of the woman. As our dependent variable is the number of children ever born to a woman of age between 13 and 49 years old, the "period at risk" is an important variable to

[5] For computational reasons we used a Logit specification as first stage for both models.

[6] Estimation of these models has been possible thanks to the development of a specific procedure for count data written in TSP ver. 4.5. In fact, even though count data procedures are fairly available in many econometric packages, no one has yet a package that allows to estimate consistently all these models. The "COUNT" module implemented in TSP has been tested against available empirical results (using data from Gurmu and Trivedi (1996) and across different packages (GAUSS, Stata, Eviews)). The COUNT module and the relative data used in this paper are available upon request by the authors.

consider to explain the number of children had by each woman. In fact, we hardly would expect that a 18 year old girl can have the same number of children as a 49 year old woman.

We have approached this "period at risk" problem in two ways. The first utilizes both the linear and the quadratic terms of the age of the parents and the length of marriage duration as regressors The second approach includes only the logarithm of parent's age to reflect the hypothesis of "proportional risk". The results highlight the importance of the variables that proxies the "period at risk". The results from both specifications are very similar, with the exception that when we use the logarithm of age the coefficient of marriage duration becomes not significant. The remaining results are extremely robust to both specifications and they do not show any significant change[7]. Both the linear and the quadratic terms of the age of the parents are significant with the expected sign. However, the coefficients of woman's age are both better determined and, on average, ten times larger with respect to those of the age of the man. Marriage duration has also a significant impact on observed fertility, with a coefficient similar to that of the age of the man.

The education of the husband does not significantly influence fertility choices, while the woman's education has a significantly negative effect. This confirms well-known results. In our case, because relative wages were not significant (see below), woman's education might also capture the effects on fertility of the earning potential of the woman outside the household. The dummies for the religion of the household's head are also significant. Muslims and others tend to have a higher fertility with respect to Hindus (the reference group), while Christians have a smaller amount of offspring.

The effects of income on fertility have been analyzed using both wage rates and household income. In both cases we found difficult to identify a clear effect of these variables on fertility. As discussed above, we computed male and female wage rates as average at village level. The coefficients of these variables appeared, in the various specifications we employed, seldom different from zero and sometime with the "wrong" sign. We have, therefore, excluded them from our preferred estimates. Their introduction did not modify in any relevant way the estimated coefficients for the other variables, with the exception of female education. In fact, because of the small number of observations on man and, especially, woman working in the market, woman's education might better approximate

[7] The results with the logarithm of parent age is available upon request by the authors.

woman's earning potential with respect to actual observed wages.

The effect of the household income on fertility has resulted quite difficult to interpret. The total household income has a coefficient that is positive, but not significant. The same result, not shown here, is obtained using per capita or per adult equivalent income. In order to better understand the effect of income and wealth on fertility we have used the size of cultivated land and a measure of poverty that classifies households into four groups of relative poverty to proxy for income effects (and for unreported income). The results obtained with land size have shown a positive link with the number of births. This might indicate that this variable is a better approximation than reported income for the earning potential of the household. On the other hand, land size and births can be positively correlated because family members are inputs in the agricultural production activities of the household. The results obtained with poverty groups have shown that poorer parents tend to have a smaller amount of offspring with respect to less poor parents.

These somehow contradictory results about the effects of income on fertility seem to suggest the possible existence of non linearity. In order to capture these possible non linearity in a simple way, we have divided total household income into five groups (according to criteria provided by the survey) and used five different income variables in the regressions (*Y*1-*Y*5). As the results show, within each income group the number of births tends to increase with income. However as we move to higher income groups, the elasticity of births with respect to income becomes smaller.

To measure the effects of the expected survival rate and of its variance on fertility choices, we have introduced in the equations the mean and the variance of the survival rate at village level. Both variables have negative and significant coefficients. However, as it is well known that experienced child mortality affects the number of births, the number of children who died per woman up to the age of five was also introduced only in the estimates of the second stage. In this way we try to make sure that the effects of expected survival rates are not significant because of their correlation with child survival experienced by each household. Controlling for the actual mortality experienced by the household (that turns out to have a positive and significant coefficient), does not reduce the importance of the expected mean and variance of the survival rate, whose coefficients remain negative and significant.[8]

[8] The presence in the equation of the actual infant mortality experienced by the household might generate

As mentioned above, in order to further test the soundness of these results we have also employed in the estimates the mean and variance of the survival rate computed at the district rather than at the village level. The results (not reported here) show that at district level the negative effect of both mean and variance are reinforced.

Let us now turn to consider the differences between the first and the second stage of the Hurdle estimates. The coefficients of the age of the wife and of the husband are larger in the first stage, indicating that the decision whether to have a child is influenced more strongly by age than the decision about the number of children. Also, the religion of the household's head appear to influence differently the two decisions. Christians are more likely to have children, but also tend to have a small number of children (the dummy for Christians has a positive sign in the first stage and a negative sign in the second stage). The role of education appears less well determined in the first stage. The education of the husband appears to influence positively the decision of having children, while it does not affect the decision relative to the number of children. This could reflect an income effect on decision of having children. Finally, the expected survival rates and its variance are significant with a negative sign also in the first stage. An increase in the variance of children survival, therefore, appears to reduce not only the expected number, but also the probability of having children at all[9].

6. Conclusions

The estimates for India show that the theoretical model proposed appears to find support from the data. If we go back to Table 1 and compare the comparative static effect with the empirical results, we see that the sign we obtain for the relevant variables seems to confirm the validity of the model for the case in which the degree of risk aversion is not "very" high. Our findings have also implication for the debate about the relationship between fertility and mortality. It is often argued, starting from the observation of a positive relationship between fertility and mortality, that an improvement in, for example, general sanitary conditions leads to a reduction of the fertility rate. This conclusion does not necessarily hold when we introduce into the analysis the role of uncertainty about the survival of children, or its empirical counterpart the variance of the survival rate.

---

problems because of the possible endogeneity of this variable. However, the exclusion of the actual mortality does not significantly change the estimates of the other parameters.

One should evaluate the impact of a policy not only looking at its effects on the survival rate, but also on its variance. If an increase in the expected probability of survival is also associated with a reduction of its variance, we have two contrasting effects. On the one hand, the reduction of mortality will lead to a reduction of fertility. On the other hand, the reduction in the risk associated with each child will generate incentives to higher fertility.

Finally, the use of the Hurdle model in the estimates has allowed us to identify a set of variables that affects differently the decision whether to have children and the decision regarding the number of children. This finding highlights the importance of the use of this model in studying processes linked to fertility.

---

[9] - We have also estimated a ZIP model in order to verify the robustness of our results. The second stage estimates do not show any relevant difference with those obtained with the Hurdle model..

References

Chowdury, A.R. (1988) "The Infant Mortality - Fertility Debate: Some International Evidence", *Southern Economic Journal*, 33, p. 666-674

Cigno, A. and Rosati, F.C. (1996) "Jointly Determined Saving and Fertility Behaviour: Theory and Estimates for Germany, UK and USA", *European Economic Review*, no. 40, pp. 1561-1589

-------- and -------- (1997) "Rise and Fall of the Japanese Saving Rate: the Role of Social Security and Intra-Family Transfers", *Japan and The World Economy*, no. 9, pp. 81-92

Gurmu S. and Trivedi P.K., (1996), "Excess Zeros in Count Models for Recreational Trips", *Journal of Business and Economic Statistics*, Vol.14, no. 4, pp.469-77

Lee L.F., (1986), "Specification Tests for Poisson Regression Models", *International Economic Review*, Vol. 27, no. 3 , pp. 689-706

Nguyen-Din, H. (1997) "A Socioeconomic Analysis of the Determinants of Fertility", *Journal of Population Economics*, p. 251-272

Olsen, R.J. (1980) "Estimating the Effect of Child Mortality on the Number of Births", *Demography*, 17, p. 113-121

Rosati, F.C. (1996) "Social Security in a non Altruistic Model with Uncertainty and Endogenous Fertility", *Journal of Public Economics*, no. 60, pp. 283-294

Schultz, T.P. (1993), "Mortality Decline in the Low Income World: Causes and Consequences", *American Economic Review*, Papers and Proceedings, May, pp. 337-342

Shonkwiler, J.S. and Shaw D.W, (1996), "Hurdle Count-Data Models in Recreation Demand Analysis", *Journal of Agricultural and Resource Economics*, Vol. 21, no. 2, pp. 210-19

Winkelmann R. (1995) "Duration Dependence and Dispersion in Count-Data Models", *Journal of Business and Economic Statistics*, Vol. 13, no. 4, pp. 467-741.

Winkelmann R. and Zimmermann K.F.,(1995) "Recent Developments in Count Data Modeling: Theory and Application", *Journal of Economic Surveys*, Vol. 9, no. 1, pp. 1-24

Wooldridge, J.M. (1996) "Quasi-Likelihood Methods for Count Data," forthcoming in the *Handbook of Applied Econometrics*

Table 2 - Summary statistics

Variables	Mean	Variance	Min	Max
Dependent variable				
Number of births	2.912	4.219	0.000	14.000
Age of husband and woman				
Age of wife	30.321	69.061	13.000	49.000
Age of husband	35.309	88.633	14.000	70.000
Age of wife squared / 10000	0.099	0.003	0.017	0.240
Age of husband squared / 10000	0.134	0.005	0.020	0.723
Marriage duration				
Marriage duration	13.492	77.939	0.000	33.000
Mortality and survival variables				
Number of children dead under 5	0.360	0.634	0.000	10.000
Variance of survival at village level	0.177	0.004	0.002	0.330
Variance of survival at district level	0.181	0.002	0.028	0.289
Child survival rate under 5 year at village level	0.901	0.003	0.621	0.995
Child survival rate under 5 year at district level	0.909	0.002	0.757	0.997
Religion and education variables				
Religion 1 - Hindu - Reference group				
Religion 2 - Muslim	0.096	0.087	0.000	1.000
Religion 3 - Christian	0.023	0.022	0.000	1.000
Religion 4 - Other religions	0.038	0.036	0.000	1.000
Education of husband 1 - Up to primary - Ref. group				
Education of husband 2 - Up to middle secondary	0.303	0.211	0.000	1.000
Education of husband 3 - Up to higher secondary	0.161	0.135	0.000	1.000
Education of husband 4 - Graduation and above	0.046	0.044	0.000	1.000
Education of wife 1 - Up to primary - Reference group				
Education of wife 2 - Up to middle secondary	0.184	0.150	0.000	1.000
Education of wife 3 - Up to higher secondary	0.062	0.058	0.000	1.000
Education of wife 4 - Graduation and above	0.008	0.008	0.000	1.000
Income variables				
Income group	1.889	1.304	1.000	5.000
Poverty index	2.676	0.945	1.000	4.000
Household income level / 1000000	0.309	0.121	0.005	3.998
Landsize	3.047	3.436	1.000	6.000

Number of observations after eliminating missing values: 31559

Tab. 3 - Estimates of the number of child births

POISSON MODELS

Variables	**Specification 1**			**Specification 2**		
	Param.	**t-stat (MLE)**	**t-stat (PMLE)**	**Param.**	**t-stat (MLE)**	**t-stat (PMLE)**
Constant	-2.596	-22.040	-24.470	-2.633	-22.270	-24.770
Age wife	0.215	37.460	39.600	0.215	37.470	39.630
Age husband	0.020	4.329	4.985	0.020	4.377	5.046
Age wife squared	-0.003	-37.440	-41.120	-0.003	-37.460	-41.170
Age husband squared	0.000	-4.811	-5.762	0.000	-4.850	-5.815
Religion 2	0.155	14.690	19.420	0.153	14.480	19.150
Religion 3	-0.052	-1.795	-2.026	-0.052	-1.805	-2.038
Religion 4	0.066	3.819	4.744	0.067	3.890	4.835
Education husband 2	0.004	0.478	0.590	0.002	0.266	0.329
Education husband 3	-0.012	-1.072	-1.263	-0.015	-1.279	-1.508
Education husband 4	-0.001	-0.038	-0.043	-0.002	-0.100	-0.113
Education wife 2	-0.052	-4.972	-5.721	-0.053	-5.030	-5.792
Education wife 3	-0.155	-7.837	-8.042	-0.156	-7.870	-8.082
Education wife 4	-0.311	-5.676	-5.248	-0.311	-5.667	-5.243
Household income level	0.006	0.568	0.678			
Y1				0.313	3.490	4.345
Y2				0.149	3.409	4.204
Y3				0.095	3.094	3.764
Y4				0.051	1.905	2.267
Y5				0.021	1.761	2.118
Marriage duration	0.029	24.870	30.810	0.029	24.830	30.760
Children dead under 5	0.190	61.180	104.000	0.191	61.190	104.000
Village variance survival	-0.907	-10.740	-13.140	-0.906	-10.730	-13.130
Village Child surv. rate under 5	-0.783	-8.080	-10.300	-0.787	-8.116	-10.350
District variance survival						
District Child surv. rate under 5						

	Specification 1	Specification 2
Lee (1986) Test 1	-41.7	-41.7
Lee (1986) Test 2	-37.6	-39.7
Wooldridge test (t stat)	-.089 (-50.5)	-.089 (-50.7)
Pseudo R2	0.199	0.200
Log-Likelihood	-53230.2	-53230.9
Degree of freedom	31540	31536

Tab. 4 - Estimates of the number of child births

HURDLE - POISSON MODELS (first stage Logit)

Variables	Specification 1				Specification 2			
	Param. 1° stage	t-stat 1° stage	Param. 2° stage	t-stat 2° stage	Param. 1° stage	t-stat 1° stage	Param. 2° stage	t-stat 2° stage
Constant	-4.996	-17.850	-1.864	-13.790	-5.078	-18.030	-1.895	-13.970
Age wife	0.300	25.330	0.194	28.310	0.300	25.300	0.194	28.310
Age husband	0.053	5.547	0.012	2.222	0.053	5.546	0.012	2.280
Age wife squared	-0.004	-22.670	-0.003	-29.850	-0.004	-22.670	-0.003	-29.860
Age husband squared	0.000	-5.807	-0.000	-2.693	0.000	-5.804	-0.000	-2.743
Religion 2	0.083	2.786	0.165	14.760	0.081	2.706	0.163	14.560
Religion 3	0.165	2.340	-0.090	-2.710	0.167	2.359	-0.091	-2.730
Religion 4	0.125	2.680	0.067	3.525	0.127	2.731	0.068	3.573
Education husband 2	0.023	1.090	0.001	0.075	0.026	1.225	-0.001	-0.163
Education husband 3	-0.023	-0.848	-0.010	-0.782	-0.017	-0.623	-0.013	-1.066
Education husband 4	0.052	1.096	-0.013	-0.560	0.062	1.302	-0.016	-0.683
Education wife 2	0.013	0.518	-0.069	-5.901	0.014	0.566	-0.070	-5.970
Education wife 3	-0.142	-3.565	-0.171	-7.274	-0.142	-3.558	-0.172	-7.313
Education wife 4	-0.322	-3.343	-0.354	-5.066	-0.322	-3.340	-0.354	-5.065
Household income level	-0.024	-0.925	0.008	0.713				
Y1					0.689	2.921	0.255	2.616
Y2					0.127	1.135	0.149	3.142
Y3					0.098	1.263	0.086	2.574
Y4					-0.025	-0.392	0.067	2.051
Y5					0.032	1.042	0.018	1.347
Marriage duration	0.014	6.751	0.037	26.080	0.015	6.892	0.037	26.000
Children dead under 5			0.177	-11.210			0.177	54.780
Village variance survival	-0.172	-0.696	-1.047	-10.000	-0.157	-0.632	-1.052	-10.050
Village Child surv. rate under 5	-0.454	-2.152	-1.035	-11.210	-0.451	-2.135	-1.036	-11.210
Lee (1986) Test 1	-39.1				-39.2			
Lee (1986) Test 2	-37.0				-37.1			
Pseudo R2	0.247				0.247			
Log-Likelihood	-53241.4				-53228.6			
Degree of freedom	31522				31514			

Social Protection Discussion Paper Series

No.	Title
0002	Extending Coverage in Multi-Pillar Pension Systems: Constraints and Hypotheses, Preliminary Evidence and Future Research Agenda
9933	Coordinating Poverty Alleviation Programs with Regional and Local Governments: The Experience of the Chilean Social Fund - FOSIS
9932	Poverty and Disability: A Survey of the Literature
9931	Uncertainty About Children's Survival and Fertility: A Test Using Indian Microdata
9930	Beneficiary Assessment of Social Funds
9929	Improving the Regulation and Supervision of Pension Funds: Are there Lessons from the Banking Sector
9928	Notional Accounts as a Pension Reform Strategy: An Evaluation
9927	Parametric Reforms to Pay-As-You-Go Pension Systems
9926	An Asset-Based Approach to Social Risk Management: A Conceptual Framework
9925	Migration from the Russian North During the Transition Period
9924	Pension Plans and Retirement Incentives
9923	Shaping Pension Reform in Poland: Security Through Diversity
9922	Latvian Pension Reform
9921	OECD Public Pension Programmes in Crisis: An Evaluation of the Reform Options
9920	A Social Protection Strategy for Togo
9919	The Pension System in Singapore
9918	Labor Markets and Poverty in Bulgaria
9917	Taking Stock of Pension Reforms Around the World
9916	Child Labor and Schooling in Africa: A Comparative Study

Social Protection Discussion Paper Series continued

No.	Title
9915	Evaluating the Impact of Active Labor Programs: Results of Cross Country Studies in Europe and Central Asia
9914	Safety Nets in Transition Economies: Toward a Reform Strategy
9913	Public Service Employment: A Review of Programs in Selected OECD Countries and Transition Economies
9912	The Role of NPOs in Policies to Combat Social Exclusion
9911	Unemployment and Unemployment Protection in Three Groups of Countries
9910	The Tax Treatment of Funded Pensions
9909	Russia's Social Protection Malaise: Key Reform Priorities as a Response to the Present Crisis
9908	Causalities Between Social Capital and Social Funds
9907	Collecting and Transferring Pension Contributions
9906	Optimal Unemployment Insurance: A Guide to the Literature
9905	The Effects of Legislative Change on Female Labour Supply: Marriage and Divorce, Child and Spousal Support, Property Division and Pension Splitting
9904	Social Protection as Social Risk Management: Conceptual Underpinnings for the Social Protection Sector Strategy Paper
9903	A Bundle of Joy or an Expensive Luxury: A Comparative Analysis of the Economic Environment for Family Formation in Western Europe
9902	World Bank Lending for Labor Markets: 1991 to 1998
9901	Active Labor Market Programs: A Review of the Evidence from Evaluations
9818	Child Labor and School Enrollment in Thailand in the 1990s
9817	Supervising Mandatory Funded Pension Systems: Issues and Challenges
9816	Getting an Earful: A Review of Beneficiary Assessments of Social Funds
~~9815~~	This paper has been revised, see Discussion Paper Series No. 9923

Social Protection Discussion Paper Series continued

No.	Title
9814	Family Allowances
9813	Unemployment Benefits
9812	The Role of Choice in the Transition to a Funded Pension System
9811	An Alternative Technical Education System: A Case Study of Mexico
9810	Pension Reform in Britain
9809	Financing the Transition to Multipillar
9808	Women and Labor Market Changes in the Global Economy: Growth Helps, Inequalities Hurt and Public Policy Matters
9807	The World Bank Approach to Pension Reform
9806	Government Guarantees on Pension Fund Returns
9805	The Hungarian Pension System in Transition
9804	Risks in Pensions and Annuities: Efficient Designs
9803	Building an Environment for Pension Reform in Developing Countries
9802	Export Processing Zones: A Review in Need of Update
9801	World Bank Lending for Labor Markets: 1991 to 1996

Summary Findings

In this paper we present a non altruistic model of demand for children in the presence of uncertainty about children's survival. Children are seen as assets, as they provide help during old age. If certain conditions are met, both the financial market and the family network are used to transfer resources to old age. Theoretical predictions relative to the change in the mean and variance of the survival rate are derived. The empirical analysis is based on data from the Human Development of India (HDI) survey. Different models for count data variables, such as Poisson, Hurdle and ZI models have been employed in the empirical analysis. The results highlight the importance of the uncertainty about children's survival in determining parental choices, thus showing that realized or expected children's death is not the only dimension that links fertility decision to children's mortality. The policy implications of such findings are briefly discussed.

HUMAN DEVELOPMENT NETWORK