ELSEVIER

Journal of Econometrics 112 (2003) 225–240

JOURNAL OF Econometrics

www.elsevier.com/locate/econbase

Family planning, gender differences and infant mortality: evidence from Uttar Pradesh, India

Alok Bhargava*

Department of Economics, University of Houston, Houston, TX 77204-5019, USA

Abstract

This paper modeled the proximate determinants of infant survival using the National Family Health Survey data on 11,500 women from the most populous Indian state Uttar Pradesh in the period 1982–1992. A methodological framework was developed for analyzing the inter-relationships between high fertility and infant mortality, gender differences in mortality, and for modeling the effects of health care and family planning variables. Probit models were estimated by maximum likelihood taking into account simultaneity of regressors and unobserved household differences. The proximate determinants of infant survival included maternal education and age at first birth, birth interval, the number of children before family planning was first used, maternal tetanus vaccination, and child's vaccinations. Indicator variables for a boy (girl) born at a birth order higher than the "ideal" number showed that unwanted births exacerbated female mortality.

JEL classification: C5; C33; I12; O12

Keywords: Infant mortality; Gender bias; Family planning; Probit models; Simultaneity

1. Introduction

Children in less developed countries suffer from various forms of under-nutrition; it is estimated that half the children are *stunted* and a smaller proportion suffer from *wasting* (FAO/WHO, 1992). Poor maternal nutritional status, unhygienic home environment, sicknesses during pregnancy, and lack of ante-natal care are likely to adversely affect intra-uterine growth (Kurz et al., 1993; Bhargava, 2000). Thus, for example, babies born in developing countries are shorter in length and weigh less than their

* Tel.: +1-713-743-3837; fax: +1-713-743-3798.
E-mail address: bhargava@uh.edu (A. Bhargava).

PII: S0304-4076(02)00162-8

counterparts in affluent societies (Falkner et al., 1994). While survival chances of retarded newborns are lower, poor access to medical care and immunization programs is likely to exacerbate infant mortality.

The likely causes of child mortality in developing countries have been analyzed in demographic research using data on large numbers of households (e.g. Hobcraft et al., 1983; Cleland and Sathar, 1984; Mohamed et al., 1998). While large surveys cannot observe women's nutritional status during pregnancy, analyses adopting a general framework are useful (Mosley and Chen, 1984). For example, hazard analysis of the data from Sri Lanka (Trussell and Hammerslough, 1983) has influenced subsequent research explaining child mortality by demographic, socioeconomic, and health care variables (e.g. Muhuri and Preston, 1991). The longitudinal study in Bangladesh has led to insights into the causes of excess female mortality (Chen et al., 1981; Koenig et al., 1990).

An important aspect of modeling the proximate determinants of child survival is the use of a framework capable of addressing a wide range of conceptual and methodological issues; conclusions from statistical analyses depend on the postulated model and the estimation method employed. For example, inter-relationships between fertility and child mortality are often discussed in the demographic literature. While some researchers have argued that a high probability of child survival is necessary for couples to use family planning (Taylor et al. 1976), others have suggested that excess or unwanted fertility exacerbates child mortality (Scrimshaw, 1978; Cleland, 1996). From an empirical standpoint, however, it is essential to draw testable implications from such hypotheses and specify statistical models accordingly. The possible direction of causality in the fertility–mortality relationship can be investigated in a framework that tackles methodological problems such as simultaneous determination of explanatory variables, and conceptual issues such as the effects of desired fertility and access to family planning on child survival.

Another important example of the relevance of model specification issues is the work by Muhuri and Preston (1991) and Muhuri (1996) emphasizing the effects of sex composition of children in the household on child mortality in Bangladesh. The model, however, is formulated in a restricted way introducing a limited number of indicator variables to represent the possible sex-composition scenarios. Moreover, the estimation method ignores problems of simultaneity that arise when the number of surviving brothers and sisters are included as explanatory variables; survival chances of older siblings and the index child are determined by similar factors. Consequently, the parameters would be inconsistently estimated.

The structure of this paper is as follows: the data from the National Family Health Survey (IIPS, 1995) are described in Section 2. Uttar Pradesh was selected because in 1991, it had a population of approximately 140 million and the total fertility rate was 5.1. In Section 3, a framework is developed for modeling the effects of demographic, socioeconomic and family planning and health care variables on infant (under-1) survival. First, the inter-relationships between fertility, family planning and child mortality are examined. Second, it is noted that the number of surviving older brothers and sisters should be directly introduced into models explaining infant survival. Third, a random effects estimator is used to tackle endogeneity of regressors and the unobserved between household differences. The empirical results for infant survival in UP are presented in Section 4. The models are separately estimated for the data covering the previous 10 and 5 year periods; the latter contain additional information on ante-natal care and child vaccinations.

2. Data and definitions

The National Family Health Survey (NFHS) was coordinated in 1992–1993 by the International Institute for Population Sciences with support from the U.S. Agency for International Development (IIPS, 1995). The sample was representative of the population and covers approximately 90,000 ever-married women in the age group 13–49 years from 25 Indian states. There are approximately 11,500 women from Uttar Pradesh where roughly 20% of the households reside in urban areas. There was retrospective information on households' background characteristics, land holding, caste, religion, dwelling space, possessions, etc. Information was also gathered on certain sanitation and environmental variables such as the type of toilet used and source for drinking water. While variables such as number of rooms in the house may not correspond to the period during which the index child was born, changes in such variables were likely to be small for households living in poverty.

The data on child mortality were compiled using information on all live births during the previous 15 years. To avoid possible errors in recalling events such as the age of a child at time of death, the present analysis utilized information on births in the period 1982–1992. For children who died, age at the time of death was imputed using alternative methods to minimize measurement errors.

For every woman, information was gathered on education, age at marriage, fertility, and family planning practices. Fertility preferences were investigated by posing hypothetical questions such as "How many of these children [the ideal number] would you like to be boys and how many would you like to be girls?" The answers to such questions were translated into indicator variables. An indicator variable for boys used in the analysis assumed the value one if a boy was born at parity where the preceding number of male births exceeds the ideal number of boys; the corresponding indicator variable for girls was analogously defined. The women were asked "Do you agree or disagree that an Indian family should have no more than two children?" Use of family planning was investigated by recording the point in time when contraceptive was first used.

For the 5-year period (1988–1992), detailed information was available on households' access to immunization programs and medical services. For example, variables such as whether the woman received ante-natal care, was visited by a health care worker during pregnancy, was inoculated against tetanus, etc, were in the data. There was information on whether the child was inoculated against polio, diphtheria, tetanus, etc. Place of delivery and complications during pregnancy were recorded. Two data sets were created for the analysis. In the first case, the models explained probability of infant survival by demographic and socioeconomic variables and by variables reflecting the utilization of family planning services. Secondly, information on health care

Table 1
Child mortality figures for the data from Uttar Pradesh (1982–1992)[a]

	Total	Urban		Rural	
		Boys	Girls	Boys	Girls
Number of children born	21 620	2059	1898	9133	8530
Number of children dead	3165	167	185	1343	1470
	(14.6)	(8.1)	(9.7)	(14.7)	(17.2)
Age at time of death					
Less than 30 days	1578	95	75	736	672
	(7.3)	(4.6)	(4.0)	(8.1)	(7.9)
1–6 months	459	24	31	200	204
	(2.1)	(1.2)	(1.6)	(2.2)	(2.4)
7–12 months	486	17	35	175	259
	(2.2)	(0.8)	(1.8)	(1.9)	(3.0)
13–24 months	379	14	27	137	201
	(1.8)	(0.7)	(1.4)	(1.5)	(2.4)
25–36 months	141	6	14	42	79
	(0.6)	(0.3)	(0.7)	(0.5)	(0.9)
37–48 months	81	7	3	34	37
	(0.4)	(0.3)	(0.1)	(0.4)	(0.4)
49–60 months	41	4	0	19	18
	(0.2)	(0.2)	(0.0)	(0.2)	(0.2)

[a]Figures in parentheses are the percentage of children who died.

and immunization program use in the period 1988–1992 augmented the explanatory variables. Comparison of the results for the two cases can provide insights into the importance of health care programs for infant survival.

Table 1 reports the number of children born in the period 1982–1992 and mortality in the different age groups; figures are presented separately for boys and girls in urban and rural areas. It was evident that approximately 80% of children who died in this period were less than a year old. Also, female mortality in rural areas was higher in age groups 7–12 months and 13–24 months. Further, since information on immunization and health care programs was available during the 1988–1992 period, it seemed fruitful to focus on infant (under-1) mortality. This is because a higher proportion of infant deaths would be included in the infant mortality data than in the situation where, for example, under-2 child mortality figures were used. The data covering the 4-year sub-period (1988–1991) were also analyzed to examine if exclusion of children who may have died in 1992 after the survey affected the empirical results. Excess female mortality in the age group 13–24 months was investigated using under-2 mortality figures (the results for infant and under-2 survival were very similar).

Table 2 reports sample means and standard deviations of selected explanatory variables used in the models for infant survival. Sample means of boys and girls born in this population were very close to those reported for the prospective Matlab study in Bangladesh (Muhuri, 1996).

Table 2
Sample means and standard deviations of selected variables[a]

Variable	Urban	Rural
Mother's education[b]	1.520	0.361
	(1.898)	(0.949)
Total number of rooms	2.518	3.119
	(2.087)	(2.523)
No toilet[c]	0.214	0.929
	(0.410)	(0.256)
Mother's age at first birth[d]	20.076	19.070
	(3.051)	(2.771)
Birth interval[d]	2.265	2.316
	(1.683)	(1.558)
Total number of children	4.259	4.755
	(2.332)	(2.416)
Ideal number of boys	1.796	2.216
	(0.928)	(0.944)
Ideal number of girls	1.176	1.350
	(0.625)	(0.705)
Indian families should have	0.645	0.561
no more than two children[c]	(0.478)	(0.496)
Approve family planning[c]	0.791	0.637
	(0.407)	(0.481)
Use family planning[c]	0.313	0.159
	(0.464)	(0.366)
Number of children before using	3.753	4.413
family planning	(2.406)	(2.439)
Mother vaccinated (tetanus)[c,e]	0.682	0.357
	(0.466)	(0.479)
Child vaccinated[c,e]	0.430	0.311
	(0.500)	(0.463)
Sample size	3957	17 663

[a]Standard deviations are in parentheses.
[b]Categorical variable 0–5.
[c]Indicator variable (Yes = 1, No = 0).
[d]In years.
[e]Based on the period 1988–1992.

3. Modeling the proximate determinants of infant survival

3.1. Family planning and the fertility–mortality inter-relationships

The inter-relationships between fertility and child mortality are dynamic in nature and depend on the economic and social development in the region. Historically, when family planning methods and medical care were unavailable, maternal health and interactions between children's genotype and nutritional factors determined survival. Thus, a large number of children were born and a relatively small proportion survived. The prospects of child survival have improved in recent years even in backward areas such as rural

Uttar Pradesh. This will influence parents' perception of the desired family size. For example, the statistics in Table 2 showed that 65% of urban women and 56% in rural areas agreed that Indian families should not have more than two children. Moreover, 79% of urban women and 63% rural women approved the use of family planning. The figures for actual use were 31% and 16%, respectively, presumably reflecting poor access to such services.

In discussing the effects of child mortality on fertility (or vice versa), it is helpful to indicate the time frame in which implications of the theory are likely to hold. For example, some proponents of the "child survival hypothesis" have suggested that couples are unlikely to adopt family planning unless they are confident that the desired number of children will survive. This can strictly be true for irreversible procedures such as sterilization. More importantly, contraceptive use and its timing, while reflecting parental expectations regarding child survival, affect the survival chances of existing children and future births. For example, parents in affluent societies can be confident of child survival shortly after birth. By contrast, child mortality is high in developing countries and the pattern more complex. The under-5 mortality figures in Table 1 showed that 91% of children in Uttar Pradesh who died were less than 2 years old; it could take 2–3 years for survival uncertainties to be resolved. Thus, irreversible family planning procedures are unlikely to be appealing in developing countries unless couples have surpassed their reproductive goals and access to contraceptives is poor.

Further, if family planning services are unavailable in the period following a birth, then risk of pregnancy would be high, once the post-partum amenorrhea due to breast-feeding is over. The ensuing pregnancy would demand rapid replenishment of vital nutrients such as iron and calcium to support fetal growth (Scrimshaw, 1996). This is difficult in developing countries because bioavailability of such nutrients from staple foods such as cereals is low; nutrient losses due to infections diminish women's capacity to produce healthy infants. Depending on parity, frequent pregnancies reduce the time available for child care and subsistence activities (Bhargava, 1997). Thus, an unmet need for family planning is likely to result in large number of children born in a relatively short interval; short birth intervals are associated with increased risk of mortality (e.g. Hobcraft et al., 1983; Gribble, 1993; Nath et al., 1994). If, however, women use family planning, then births can be spaced. This would be beneficial for health of the surviving children and improve the prospects of carrying the subsequent pregnancy to full-term. Early use of contraceptives is therefore likely to enhance infant survival.

Now, at a given point in time, direction of causality in the fertility–mortality interrelationship has certain asymmetric implications for infant mortality. High mortality at low order births is likely to influence parents' decision to have more than the desired number of children. By contrast, if mortality is the result of a large number of unwanted births, then children born at higher parities would be at greater risk. There may also be gender differences; researchers have reported excess female mortality at higher parities in India (Das Gupta, 1987). However, mortality rates for first-born children are high (Bongaarts, 1987; Trussell, 1988). From a biological viewpoint, there is competition for nutrients between the fetus and the young mother's own requirements for growth (Falkner et al., 1994). Multivariate analyses can partially control for this by including maternal age at first birth as an explanatory variable; some additional methodological issues are discussed in the next two sections.

3.2. Some aspects of model formulation for the infant survival relationship

In the absence of family planning, the total number of live births to a woman will chiefly depend on her fecundity, nutritional status, age at marriage and the breast-feeding patterns. The number of surviving children is affected by demographic, socioeconomic and health care variables. Another important aspect of modeling the proximate determinants of infant survival in south Asian countries is that parents may desire a certain number of sons (Chen et al., 1981; Sen and Sengupta, 1983). This has led Muhuri and Preston (1991) and Muhuri (1996) to postulate that the surviving older brothers and sisters differentially affect survival chances of the index child.

At a conceptual level, differential effects of older boys and girls on index child's survival raise some interesting questions. For, it is common in poor households for young girls to spend a proportion of their time on household tasks. This is not the case for boys especially in rural areas where adult males undertake strenuous work. Further, intra-uterine growth retardation is likely to increase with parity (Al et al., 1997; Bhargava, 2000); mothers also have less time for care of children born at higher parities. Thus older sisters can fill the child care gap. While survival chances of siblings are likely to be positively correlated, care given by girls will strengthen the coefficient of older sisters in the model explaining index child's survival. This phenomenon can be investigated to a limited extent using the data from Uttar Pradesh due to certain methodological issues.

First, Muhuri and Preston (1991) represented households' sex composition by indicator variables for cases such as where 1, 2 or more older brothers or sisters are present. These variables and interactions between them were used as explanatory variables in the model explaining child mortality. The numerous sex compositions in the sample, however, cannot be captured in this way. More importantly, without an appropriate benchmark defining the indicator variables, it is difficult to interpret the coefficients. This problem can be circumvented by including the *actual* number of surviving older brothers and sisters as explanatory variables.

Second, it is difficult to fully capture age differences between older brothers and sisters and the index child since households have different age and sex composition of children. While the age gap reflects capacity of older children to provide care, one cannot easily develop an "index" of child care. The model for infant survival can include birth interval and the number of surviving older brothers and sisters as regressors; unobserved household effects will partially reflect age distribution of children. It would be useful to collect information on time allocation patterns of girls and boys.

Third, survival chances of older brothers and sisters and the index child are determined by a similar set of factors; errors affecting infant survival relationship are likely to be correlated with these explanatory variables. This will lead to inconsistent estimates of the parameters from probit or logit models. Other regressors can also be simultaneously determined. For example, inasmuch as couples have access to family planning, the number of surviving children before contraceptive is used is potentially an endogenous variable. Similarly, while birth interval has been found positively associated with child survival, researchers have argued that short birth intervals are the *result* of mortality of the preceding child. It is important to use suitable econometric methods for estimating the model parameters.

3.3. Sources of endogeneity and the econometric estimation

A major source of simultaneity in the infant survival relationship is that the determinants of survival of older siblings and the index child are similar; the number of surviving children before a family planning method is used is potentially an endogenous variable since it reflects parental preferences about family size. If, however, the estimation method addresses the prime source of endogeneity (i.e. number of surviving children), then it is likely that statistical tests will accept exogeneity hypothesis for the decision to use family planning. Similarly, while birth interval may be affected by death of the preceding child, the problem of reverse causality can be circumvented by controlling for survival status of older siblings.

A probit model explaining the survival to age one year of N live births (y_i) by m endogenous variables (Y_i) and n exogenous variables (X_{1i}) is given by

$$y_i = Y_i'\gamma + X_{1i}'\beta + u_i, \quad (i = 1, 2, \ldots, N), \tag{1}$$

$$Y_i = P'X_i + V_i, \quad (i = 1, 2, \ldots, N). \tag{2}$$

Eq. (1) is the "structural" model for the dichotomous variable y that is one if the child dies before reaching the age of one, and zero otherwise. The coefficients of Y and X_1 are represented by $m \times 1$ and $n \times 1$ vectors γ and β, respectively. Eq. (2) is a "reduced form" for the m endogenous variables in Eq. (1), with a $m \times n$ coefficient matrix P'; V is an $m \times 1$ matrix of reduced form errors. It is assumed that certain exogenous variables are excluded from the structural model for identifying the parameters (e.g. Newey, 1987; Rivers and Vuong, 1988). The errors in Eqs. (1) and (2) are assumed to be jointly normally distributed.

Further, assuming that there are H households in the sample with different number of live births, errors affecting Eqs. (1) and (2) will have an unbalanced structure. Treating the household effects as randomly distributed variables, the errors in Eq. (1) can be decomposed as

$$u_{hj} = \delta_h + w_{hj} \quad (j = 1, \ldots, J_h; h = 1, 2, \ldots, H), \tag{3}$$

where δ's are household specific random effects, w's are randomly distributed variables, and J_h is the number of live births in household h. Similarly, random effects can be introduced in reduced form Eq. (2); the errors on the first endogenous variable in (2) can be written as

$$V_{1hj} = \lambda_{1h} + w_{1hj} \quad (j = 1, \ldots, J_h; h = 1, 2, \ldots, H). \tag{4}$$

Even in the absence of random effects, estimation of the system given by Eqs. (1) and (2) is complicated because the "nuisance" parameters in the reduced form (2) cannot be eliminated by step-wise maximization (Koopmans and Hood, 1953). Thus, researchers have used "conditional" likelihood functions assuming that the errors in Eqs. (1) and (2) are jointly normally distributed (Smith and Blundell, 1986). This entails working with the density of u_i conditional on V_i in (1) and substituting a consistent estimator of P from Eq. (2) in the modified Eq. (1). Model parameters are then estimated by maximum likelihood using a numerical optimization scheme; standard errors are obtained using a Taylor series expansion. The conditional likelihood function with random effects has been developed by Vella and Verbeek (1999); software packages (e.g. LIMDEP, 1995) can be extended to estimate the parameters. Because the errors have a random effects structure, it is necessary to include two covariance terms per endogenous variable in Eq. (1). Exogeneity of a variable is tested by testing the null hypothesis that coefficients of covariance terms computed from the reduced form residuals are zero.

Lastly, in instrumental variables estimation, it is desirable that the instruments are strongly correlated with the endogenous variables (Hotelling, 1936; Sargan, 1958). Because correlations are low in cross-sectional analysis, instrumental variables methods can lead to poor results. This issue was discussed by Bhargava and Sargan (1983) for dynamic models estimated using longitudinal data. Recently, Bound et al. (1995) have underscored its importance; Bollen et al. (1995) discuss certain problems arising in probit models. The approach in this paper reports the results for the cases where endogeneity of regressors is ignored and where it is addressed.

4. The empirical results

Tables 3 and 4 present maximum likelihood estimates of probit models for infant survival for the periods 1982–1992 and 1988–1992, respectively. The tables contains four specifications. In Model 1, the problem of multiple children in households is ignored. Model 2 introduces random effects to capture the unobserved between household differences. The number of surviving older brothers and sisters are treated as endogenous variables in Model 3; the instruments used are the household's possessions and number of older boys and girls born before the index child. Lastly, in last column, the coefficients estimated for Model 3 are converted into the respective marginal effects of the explanatory variables.

4.1. Empirical results for the period 1982–1992

The results for Models 1 and 2 in Table 3 have some interesting features. First, the indicator variable for rural areas was estimated with a negative coefficient indicating lower survival probability in rural areas. The model parameters were broadly similar when the model was separately estimated for urban and rural samples. Second, while the sample statistics in Table 2 showed mortality of girls to be higher in rural areas, coefficient of the indicator variable for girls was not significantly different from zero. Interestingly, the indicator variable that is equal to one if a girl was born at a parity higher than the ideal number was negative and significant in all three specifications.

Table 3
Maximum likelihood estimates of probit models for infant survival in Uttar Pradesh for the period (1982–1992)

Variable	Model 1	Model 2	Model 3	Marginal
Constant	1.126*	1.155*	1.227*	—
	(0.097)	(0.100)	(0.103)	
Indicator for rural areas	−0.146*	−0.139*	−0.137*	−0.020*
	(0.045)	(0.050)	(0.051)	(0.007)
Indicator for girls	0.052	−0.046	0.075	0.011
	(0.038)	(0.040)	(0.045)	(0.006)
Indicator for girls born	−0.228*	−0.219*	−0.251*	−0.036*
after the ideal number	(0.039)	(0.041)	(0.044)	(0.006)
Indicator for boys born	−0.119*	−0.111*	−0.053	−0.008
after the ideal number	(0.035)	(0.037)	(0.040)	(0.006)
Number of older girls	0.134*	0.127*	0.125*	0.018*
	(0.013)	(0.013)	(0.016)	(0.002)
Number of older boys	0.074*	0.066*	0.060*	0.009*
	(0.013)	(0.013)	(0.015)	(0.002)
Mother's education	0.065*	0.068*	0.062*	0.009*
	(0.012)	(0.013)	(0.014)	(0.002)
Mother's age at first birth	0.009*	0.010*	0.006	0.001
	(0.004)	(0.004)	(0.004)	(0.001)
Birth interval	0.170*	0.176*	0.160*	0.023*
	(0.010)	(0.007)	(0.006)	(0.001)
Number of children before	−0.091*	−0.089*	−0.073*	−0.011*
family planning	(0.006)	(0.006)	(0.008)	(0.001)
Indicator for no toilet	−0.131*	−0.143*	−0.147*	−0.021*
	(0.043)	(0.047)	(0.049)	(0.007)
Number of rooms	0.012*	0.014*	0.017*	0.002*
	(0.005)	(0.006)	(0.007)	(0.001)
Covariance term 1 (girls)	—	—	0.078*	0.011*
			(0.003)	(0.001)
Covariance term 2 (girls)	—	—	−0.222*	−0.032*
			(0.009)	(0.001)
Covariance term 1 (boys)	—	—	0.179*	0.026*
			(0.008)	(0.001)
Covariance term 2 (boys)	—	—	−0.271*	−0.039*
			(0.007)	(0.001)
Rho	—	0.267*	0.252*	—
		(0.038)	(0.039)	

There were 21512 live births; Model 3 treats number of older boys and girls as endogenous; the last column reports the "marginal effects" for the parameter estimates in Model 3; asymptotic standard errors are in parentheses; covariances terms are conditional expectations of errors in Eq. (1); Rho is the proportion of variance due to random effects; $^* p < 0.05$.

This suggest that survival chances of girls and boys born at low parities were similar but mortality was greater for girls born at higher parities.

Third, the number of surviving older brothers and sisters were positively associated with survival chances of the index child. Since older sisters are likely to take

Table 4
Maximum likelihood estimates of probit models for infant survival in Uttar Pradesh for the period (1988–1992)

Variable	Model 1	Model 2	Model 3	Marginal
Constant	0.912*	0.912*	1.103*	—
	(0.165)	(0.164)	(0.167)	
Indicator for rural areas	−0.037	−0.037	−0.047	−0.005
	(0.075)	(0.075)	(0.077)	(0.008)
Indicator for girls	0.114**	0.114**	0.170*	0.018*
	(0.063)	(0.063)	(0.068)	(0.007)
Indicator for girls born	−0.208*	−0.208*	−0.261*	−0.028*
after the ideal number	(0.066)	(0.065)	(0.070)	(0.008)
Indicator for boys born	−0.016	−0.016	−0.035	−0.004
after the ideal number	(0.059)	(0.060)	(0.061)	(0.007)
Number of older girls	0.185*	0.184*	0.227*	0.025*
	(0.024)	(0.022)	(0.030)	(0.004)
Number of older boys	0.125*	0.126*	0.162*	0.018*
	(0.024)	(0.023)	(0.032)	(0.004)
Mother's education	0.0001	0.0002	−0.001	−0.0001
	(0.020)	(0.020)	(0.020)	(0.002)
Mother's age at first birth	0.002	0.002	−0.001	−0.001
	(0.007)	(0.007)	(0.007)	(0.001)
Birth interval	0.180*	0.180*	0.186*	0.018*
	(0.016)	(0.014)	(0.014)	(0.002)
Number of children before	−0.124*	−0.124*	−0.138*	−0.015*
family planning	(0.014)	(0.012)	(0.017)	(0.002)
Indicator for tetanus vacc.	0.289*	0.289*	0.282*	0.031*
	(0.047)	(0.047)	(0.049)	(0.005)
Indicator for child vacc.	1.001*	1.001*	0.988*	0.107*
	(0.063)	(0.066)	(0.069)	(0.007)
Indicator for no toilet	−0.117	−0.117	−0.104	−0.011
	(0.072)	(0.070)	(0.072)	(0.008)
Number of rooms	0.007	0.007	0.006	0.001
	(0.009)	(0.009)	(0.009)	(0.001)
Covariance term 1 (girls)	—	—	0.053*	0.006*
			(0.014)	(0.001)
Covariance term 2 (girls)	—	—	−0.280*	−0.030*
			(0.018)	(0.002)
Covariance term 1 (boys)	—	—	0.223*	0.024*
			(0.024)	(0.003)
Covariance term 2 (boys)	—	—	−0.094*	−0.010*
			(0.014)	(0.002)
Rho	—	0.001	0.127	—
		(22.817)	(0.182)	

There were 9674 live births; $^{*}p < 0.05$; $^{**}p < 0.10$; see notes to Table 3.

care of younger siblings, one would expect the coefficient of girls to be higher than the corresponding estimate for boys, which was the case. The null hypothesis that coefficients of boys and girls are the same was rejected by a likelihood ratio test. Maternal education, measured by a categorical variable assuming six different values, was significantly associated with infant survival. There were some non-linearities apparent in the relationships between infant survival and maternal age at first birth and the number of surviving older sisters. For example, the squared term for older sisters was estimated with a negative coefficient that was marginally significant, indicating an adverse effect after couples have had a certain number of girls. While the overall results were similar to those reported in Table 3, the positive coefficient of the number of surviving older sisters should be interpreted in this light.

Fourth, birth interval was significantly positively associated with infant survival. The number of surviving children before a family planning method was used was negatively associated. Note that if the woman did not use family planning, this variable was set equal to the number of surviving children. Because the model accounted for survival status of children born before the index child, early use of family planning was beneficial for infant survival.

Fifth, the indicator variable for not having a toilet was negatively associated with probability of infant survival; having access to piped water was positively associated in certain specifications. The number of rooms was positively associated. Environmental variables are likely to gradually affect health indicators such as anthropometric measurements and morbidity (Bhargava, 1994). Because these indicators are expensive to measure in large surveys, the relationship between environmental variables and infant survival is a useful proxy for the underlying biological relationships.

Sixth, the household specific random effects were statistically significant in Model 2; approximately 25% of the residual variance was due to the unobserved between household differences. However, there were only minor differences in the results for Models 1 and 2. This could be due to the fact that explanatory variables such as number of rooms in the house and maternal education accounted for many of the between household differences. The small unobserved differences suggest that latent factors common to survival chances of siblings may also be small in magnitude.

Seventh, the results for Model 3 that treats the number of surviving older brothers and sisters as endogenous variables, were close to those reported for Models 1 and 2. However, coefficients of the covariance terms were statistically different from zero thereby rejecting the exogeneity null hypotheses for these variables. Coefficients of the number of older brothers and sisters in Model 3 were not noticeably different from the corresponding estimates in Model 2 that ignored endogeneity problems. The remaining estimates were close in all three specifications; the proportion of variance explained by random effects was similar in Models 2 and 3. While minor violations in statistical assumptions are not likely to dramatically alter parameter estimates because of the large sample size, the estimates were dependent on the use of number of preceding male and female births as instruments. These variables were highly correlated with the number of surviving older children and are exogenous in the absence of family planning. Even when family planning is used, it is likely that the unobserved errors affecting infant survival are related to factors such as the maternal nutritional and health status, and environmental factors such as sanitation and hygiene.

Eighth, exogeneity hypothesis for the number of surviving children before a family planning method is used was tested; coefficients of the covariance terms for this variable were not statistically different from zero. This is perhaps not surprising since the prime source of simultaneity was the number of surviving children. Once this is taken into account, statistical procedures were unlikely to find a relationship between family planning decisions and the unobserved factors affecting infant survival. Lastly, because of the presence of random effects and endogenous regressors in Model 3, standard errors of the estimated parameters were computed by a bootstrap procedure; the results were close (Bollen et al., 1995).

4.2. Empirical results for models including immunization variables (1988–1992)

In addition to the variables used in Table 3, the models in Table 4 included indicator variables for whether the mother was vaccinated against tetanus and if the child was vaccinated. The estimated coefficients of these variables were large and statistically significant thereby showing the importance of health care programs for infant survival. The analysis was repeated using data for the period 1988–1991 (excluding the births in 1992) to avoid the problem that some infants may have died shortly after the survey. However, the empirical results for the past 4 and 5 years were very close. Because children are typically vaccinated around the age of 3 months, the indicator for child vaccination was dropped from the models to investigate the robustness of the results. This led to a change in coefficient of the indicator variable for maternal tetanus vaccination though the remaining estimates were similar.

In contrast with the results in Table 3, coefficients of maternal education and age at first birth, indicator variable for not having a toilet and the number of rooms in the house were not statistically significant in Table 4. This could be due to the reduction in the sample size though there was also a decline in the magnitudes of these coefficients. This was not true for variables such as birth interval and number of children before a family planning method was used; the estimated coefficients of family planning related variables were more robust changes in model specification.

The indicator variable for rural areas was insignificant in Table 4. The indicator variable for girls was significant though with a *positive* sign. By contrast, the indicator variable for girls born at a higher parity than the ideal number had a *negative* coefficient. These results suggest that girls born at low parities were not at a disadvantage which could be due to better maternal nutritional status and also because daughters' contribution to housework is viewed favorably. This would not be the case for girls born at high parities who face greater growth retardation and increased competition for resources, including medical care. The indicator variable for boys born after the ideal number was not significant in Table 4, suggesting the possibility of a selective neglect of girls born at high parities.

5. Conclusion

The proximate determinants of infant survival were modeled in this paper using data from Uttar Pradesh, the most populous Indian state. The analysis extended previous demographic research by emphasizing the role of model formulation. Using a suitable parameterization for sex composition of older siblings and addressing issues of simultaneity, the empirical results yielded new insights. The number of older brothers and sisters was found to differentially affect the survival probability of the index child, with older sisters possibly filling in the child care gap. Moreover, indicator variables constructed on the basis of the stated preferences for sons and daughters indicated that unwanted fertility contributes to excess mortality of higher order births, especially of girls. Family planning programs are therefore likely to narrow gender differences in infant mortality. Further, studies explaining variation in *average* fertility rates in developing countries by couples' fertility preferences (Pritchett, 1994) should not be viewed as casting doubt on the efficacy of family planning programs. This is because fertility also depends on socioeconomic variables and access to family planning services; such factors are masked in analyses using cross-country data on average fertility rates.

From the standpoint of the effects of maternal education on infant mortality (Bicego and Boerma, 1993; Murthi et al., 1995), it is likely that educated mothers take better care of themselves and the infant. However, the results in Tables 3 and 4 indicated that family planning and immunization programs had a greater impact on infant survival. Since a majority of women in rural Uttar Pradesh are uneducated, adult education programs would require substantial resources. Investments in family planning and immunization programs are also expensive but are likely to prevent infant mortality in a shorter time frame. Female education should be an important policy goal; parental scores on cognitive tests were predictors of Kenyan children's physical and intellectual development (Bhargava, 1998, 1999). Policies exploiting synergisms between female education and health care and family planning use are likely to be cost-effective.

Finally, public policies are unlikely to have the desired effect on children's lives unless they control unwanted fertility (Cleland, 1996). Even uneducated parents are aware of bleak employment prospects in rural areas; living conditions in urban slums are unlikely to be a motivating factor for large families. While customs can dampen the speed with which family planning programs are adopted, it is essential that such services reach the poor. Infants in developing countries survive if a *minimal* set of maternal health and nutritional requirements are met. Reducing infant mortality is a modest goal. Once this is achieved, policy makers can concentrate on policies that enhance children's physical and intellectual development via nutritional and educational programs (Scrimshaw, 1998).

Acknowledgements

While retaining the responsibility for the views, the author would like to thank Amartya Sen for suggesting this topic and the Mac Arthur Foundation for partial support. Thanks are also due to Erwan Quintin and the two referees for helpful comments.

References

Al, M.D.M., Houwelingen, A.C., Hornstra, G., 1997. Relation between birth order and the maternal and neonatal docohexaenoic acid status. European Journal of Clinical Nutrition 51, 548–553.

Bhargava, A., 1994. Modelling the health of Filipino children. Journal of the Royal Statistical Society A 157, 417–432.
Bhargava, A., 1997. Nutritional status and the allocation of time in Rwandese households. Journal of Econometrics 77, 277–295.
Bhargava, A., 1998. A dynamic model for the cognitive development of Kenyan school children. Journal of Educational Psychology 90, 162–166.
Bhargava, A., 1999. Modelling the effects of nutritional and socioeconomic factors on the growth and morbidity of Kenyan school children. American Journal of Human Biology 11, 317–326.
Bhargava, A., 2000. Modelling the effects of maternal nutritional status and socioeconomic variables on the anthropometric and psychological indicators of Kenyan infants from age 0–6 months. American Journal of Physical Anthropology 110, 89–104.
Bhargava, A., Sargan, J.D., 1983. Estimating dynamic random effects models from panel data covering short time periods. Econometrica 51, 1635–1660.
Bicego, G.T., Boerma, J.T., 1993. Maternal education and child survival: a comparative study of survey data from 17 countries. Social Science and Medicine 36, 1207–1227.
Bollen, K., Guilkey, D.K., Mroz, T.A., 1995. Binary outcomes and endogenous explanatory variables: tests and solutions with an application to the demand for contraceptive use in Tunisia. Demography 32, 111–131.
Bongaarts, J., 1987. Does family planning reduce infant mortality rates? Population and Development Review 13, 323–334.
Bound, J., Jaeger, D., Baker, R., 1995. Problems with instrumental variables estimation when the correlation between instruments and the endogenous explanatory variable is weak. Journal of American Statistical Association 90, 443–450.
Chen, L.C., Huq, E., D'Souza, S., 1981. Sex-bias in the family allocation of food and health in rural Bangladesh. Population and Development Review 7, 50–77.
Cleland, J., 1996. Population growth in the 21st century: cause for crisis or celebration? Tropical Medicine and International Health 1, 15–26.
Cleland, J., Sathar, Z.A., 1984. The effect of birth spacing on childhood mortality in Pakistan. Population Studies 38, 401–418.
Das Gupta, M., 1987. Selective discrimination against female children in rural Punjab. Population and Development Review 13, 77–100.
Falkner, F., Holzgreve, W., Schloo, R.H., 1994. Prenatal influences on postnatal growth: overview and pointers for research. European Journal of Clinical Nutrition 48, S15–S24.
FAO/WHO, 1992. Nutrition and development. A global assessment. Food and Agriculture Organization, Rome.
Gribble, J.N., 1993. Birth intervals, gestational age, and low birth weight: are the relationships confounded? Population Studies 47, 133–146.
Hobcraft, J.N., McDonald, J.W., Rutstein, S., 1983. Child spacing effects of infant and early child mortality. Population Index 49, 585–618.
Hotelling, H., 1936. Relations between two sets of variates. Biometrika 28, 321–377.
IIPS, 1995. National Family Health Survey. International Institute for Population Sciences, Mumbai.
Koenig, M.A., Phillips, J.F., Campbell, O.M., D'Souza, S., 1990. Birth interval and childhood mortality in rural Bangladesh. Demography 27, 251–265.
Koopmans, T.C., Hood, W.C., 1953. Studies in Econometric Methods. Wiley, New York, NY.
Kurz, K.M., Habicht, J.P., Rasmussen, K.M., Schwager, S.J., 1993. Effects of maternal nutritional status and maternal energy supplementation on length of postpartum amenorrhea among Guatemalan women. American Journal of Clinical Nutrition 58, 636–640.
LIMDEP, 1995. Econometric Software Inc., New York.
Mohamed, W.N., Diamond, I., Smith, P.W.F., 1998. The determinants of infant mortality in Malaysia: a graphical chain modelling approach. Journal of the Royal Statistical Society A 161, 349–366.
Mosley, W.H., Chen, L., 1984. An analytical framework for the study of child survival in developing countries. Population and Development Review 10, 25–48.
Muhuri, P., 1996. Estimating seasonality effects on child mortality in Matalab, Bangladesh. Demography 33, 98–110.

Muhuri, P., Preston, S., 1991. Effects of family composition on mortality differentials by sex among children in Matlab, Bangladesh. Population and Development Review 17, 415–434.
Murthi, M., Guio, A.C., Dreze, J., 1995. Mortality, fertility, and gender bias in India: a district level analysis. Population and Development Review 21, 745–782.
Nath, D.C., Land, K.C., Singh, K.K., 1994. Birth spacing, breastfeeding, and early child mortality in a traditional Indian society: a hazards model analysis. Social Biology 41, 168–180.
Newey, W., 1987. Efficient estimation of limited dependent variable models with endogenous explanatory variables. Journal of Econometrics 36, 230–251.
Pritchett, L.H., 1994. Desired fertility and the impact of population policies. Population and Development Review 20, 1–55.
Rivers, D., Vuong, Q., 1988. Limited information estimators and exogeneity tests for simultaneous probit models. Journal of Econometrics 39, 347–366.
Sargan, J.D., 1958. The estimation of economic relationships using instrumental variables. Econometrica 26, 393–415.
Scrimshaw, S., 1978. Infant mortality and behavior in the regulation of family size. Population and Development Review 4, 383–403.
Scrimshaw, N.S., 1996. Nutrition and health from womb to tomb. Nutrition Today 31, 55–67.
Scrimshaw, N.S., 1998. Malnutrition, brain development, learning and behavior. Nutrition Research 18, 351–379.
Sen, A., Sengupta, S., 1983. Malnutrition of rural children and the sex bias. Economic and Political Weekly 18, 855–864.
Smith, R.J., Blundell, R.W., 1986. An exogeneity test for a simultaneous equation tobit model with an application to labor supply. Econometrica 54, 679–685.
Taylor, C.E., Newman, J.S., Kelly, N., 1976. The child survival hypothesis. Population Studies 30, 262–278.
Trussell, J., 1988. Does family planning reduce infant mortality? An exchange. Population and Development Review 14, 171–178.
Trussell, J., Hammerslough, C., 1983. A hazard-model analysis of the covariates of infant and child mortality in Sri Lanka. Demography 20, 1–26.
Vella, F., Verbeek, M., 1999. Two-step estimation of panel data models with censored endogenous variables and selection bias. Journal of Econometrics 90, 239–263.