8-2005

Exogamy and Bias Against Daughters in Health-care Provision: A Theory and Evidence from Two Northern States in India

Sajal Lahiri
Southern Illinois University Carbondale

Sharmistha Self
College of Saint Benedict

August 29, 2005

Exogamy and Bias against Daughters in Health-care Provision:
A Theory and Evidence from Two Northern States in India*

By

Sajal Lahiri ‡ and Sharmistha Self §

Abstract

This is a theoretical and empirical paper to analyze possible bias against daughters in the provision of healthcare. Women once married become part of in-laws' families, leading to certain inter-family externalities in household decision making, which in turn result in gender bias in healthcare. We test our theoretical predictions using LSMS household survey data from two Indian states, *viz.* Uttar Pradesh and Bihar. We find strong evidence for the existence of bias against daughters. We also find, consistent with our theory but contrary to conventional wisdom, that the bias is more pronounced among Hindu families (who tend to practice exogamy) than among Muslim families (who very commonly intermarry)

JEL Classification: H52, O10, O16.
Keywords: gender bias, healthcare, marriage, extended family, health-care cost

‡ Department of Economics, Southern Illinois University, Carbondale, IL. 62901-4515, U.S.A. E-mail: lahiri@siu.edu
§ Department of Economics, College of Saint Benedict/ Saint John's University, Saint Joseph, MN 56374, U.S.A.; Email: sself@csbsju.edu

---

* We are grateful to the World Bank for making the *Uttar Pradesh and Bihar Survey of Living Conditions* available to us. The World Bank is not responsible for the estimations reported, and the conclusions drawn, in this paper. We also thank Rick Grabowski for helpful comments on an earlier draft of the paper.

1 Introduction

The primary objective of this paper is to understand a source of gender inequality in South Asia, *viz.*, differential access to healthcare for children. Gender bias in health can be found at multiple levels and in multiple aspects, and it may arise due to reasons not necessarily directly related to health (Okojie, 1994). This bias can have devastating consequences in the developing world where functional and affordable public or private health care system is non-existent.

In the literature there have been numerous attempts to quantify the consequence of gender bias against women. A common approach has been to look at figures for female-to-male ratios in population. In most countries the ratio is close to, or even a little higher than, unity. In most South Asian countries however it is significantly less than unity: 0.934 in China, 0.926 in India and 0.901 in Pakistan, for example. Sen (1992, 2003) concluded that this was due to a higher mortality among women because of discrimination against them in health and nutrition.[1] He coined the phrase "missing women" to explain this phenomenon.[2] He claimed that there were over 100 million missing women in Asia.[3]

There are of course many reasons why there may be discrimination against women in health and nutrition and many alternative explanations can be found in the literature. It ranges from differences in women's childbearing roles, sex-preference of children, lack of autonomy for women, early marriage (Okojie, 1994) to religious preferences, regional factors and civil freedom (Dollar and Gatti, 1999); from denial by governments and communities to excessive female mortality (Croll, 2001) to differences in bargaining power within the household (Quisumbing and Maluccio, 2000; Basu, 2005); from intra-household allocation of

[1] In the popular medium, one also hears of selective abortion and female infanticide as possible reasons.

[2] To be precise, missing women are the extra women population that would be here today but for discrimination.

[3] Klasen (1994) estimated this number to be 89 million; Coale (1991)'s estimate is 60 million. Klasen and Wink (2002) examined whether this number changed since the past decade, and found that the combined estimate of missing women has increased in absolute terms though it has fallen proportionally to total population.

nutrients (Bardhan, 1974; Boserup, 1980; Behrman, 1988) due to son preference or due to parental response to different labor market outcomes (Drèze and Sen, 1991) to differences in investment in boys and girls to dowry related deaths (Johnson, 1996; Prasad, 1996).

In a recent paper, Oster (2005) interestingly finds that many of the missing women are not really missing, but were not born at all. In other words, there are differences in female-to-male ratio at birth. She finds that carriers of hepatitis B virus are more likely to give birth to boys than girls. Remarkably, she concluded that this factor can explain 75% of missing women in China, but a much smaller proportion (less than 20%) in India and Pakistan.[4] Thus, still a large proportion of the missing women in India can potentially be explained by discrimination in health and nutrition.

While many factors, as mentioned above, have been used to explain various aspects of gender bias, most apply to adults. Relatively fewer attempts have been made to explain the gender bias among children. Parents' decisions regarding intra-household allocation of resources may have a direct impact on the health and well-being of children, and the decisions can involve children's education, child labor, or health and treatment of their sick children. Poor health during childhood may have life long consequences.

In this paper our focus is the impact of intra-household decision making by parents regarding children's healthcare. We develop a theoretical model to provide possible explanations of why a parent may choose to discriminate against daughters, and then test the predictions of the model using LSMS micro data from two Indian states: Uttar Pradesh (UP) and Bihar.

Family decisions on healthcare may be influenced by certain socio-cultural practices prevalent in India. One such practice is that of the extended family system. According to this system, once daughters are married, they leave the parents' house and live with their in-laws. Any income of the daughter after marriage becomes a property of the in-law's family.

[4] Her estimates bring down the total number of missing women in Asia to 32 miilion.

Sons continue to live with parents after getting married and the income of a son and his wife becomes part of the income of the parents' family. Moreover, among Hindu families in the northern states of India exogamy is widely practiced. In such situations, parents of brides and grooms were more often than not unknown to each other when those brides and grooms were younger. These parents therefore are unlikely to co-ordinate among themselves and take the future incomes of daughters-in-law into consideration when making decisions on intra-family distribution of resources. Using this failure of parents to internalize the inter-household externality of their potential daughters-in-law's education and income potential, we show that there would be gender bias in terms of how many children receive medical care within a family. Additionally we show that as the cost of medical care rises, the bias against girls gets worse and the bias works insofar as it interacts with health care costs.

The layout of the paper is as follows. The following section develops the theoretical framework and a number of testable hypotheses. In section 3 we test the hypotheses using LSMS data from two northern provinces in India. We use two econometric techniques: Ordinary Least Squares and bivariate Logit regressions. These two are discussed respectively in subsections 3.2 and 3.3. In subsection 3.4 we extend our analysis to examine whether there is any religion-specific difference in family behavior with respect to gender-specific health-care provision. Finally, some concluding remarks are made in section 4.

2 The formal framework

We consider a society in which there are a number of household each with N_m number of boys and N_f numbers of girls to start with. There are two time periods. In period 1 all the children go to school and enjoy leisure. However, during this period a proportion of these children become ill. The proportion of boys and girls getting ill are denoted by i_m and i_f respectively. These proportions are known to the families. However, some of the illnesses are severe and some are not, and the families face some uncertainty about which of the sick children are severely ill and which are not. We assume that the total number of severely sick boys and girls — denoted by X_m for boys and X_f for girls — follow a binomial distribution with the probability that a child is severely sick denoted by p, and that X_m and X_f are independently distributed. We also assume that a family that cannot identify a severely sick child, takes t_m proportion of sick male children and t_f proportion of sick female children, chosen at random, to healthcare specialists and this costs the family the amount c per child. These two variables are chosen optimally by the family, and this optimality problem will be considered later on. Thus, $t_m X_m$ of the severely sick boys and $t_f X_f$ severely sick girls receive treatment. A proportion δ, representing the quality of healthcare, of these children do not survive. All the severely sick children who do not receive treatment — $(1-t_m)X_m$ boys and $(1-t_f)X_f$ girls — also die. Total number of boys and girls that die — denoted by D_m and D_f respectively — are therefore given by

$$D_m = \delta t_m X_m + (1-t_m)X_m, \tag{1}$$

$$D_f = \delta t_f X_f + (1-t_f)X_f. \tag{2}$$

The children who do not die work in the second period and earn an income — w_m for each boy and w_f for each girl.

Marriage plays an important role in our analysis. We assume that the surviving young men and women get married at the beginning of period 2 to people from outside the family. The daughters leave home to live with their in-laws and the daughters-in-law move in with the husbands' families. However, as you will see later on, the number of surviving boys will outnumber the surviving girls and therefore some of the boys will remain unmarried in the second period. Since for the daughters-in-law, investment in healthcare is made by other families, a family has no control over how many daughters-in-law there are. For expositional simplicity, we shall assume that there are only two families in the society, and mark the variables for the other family with an asterisk as a superscript. Since there are only $N_f - D_f^*$

number of surviving girls in the other family, $(N - D_m) - (N_f - D_f^*)$ number of boys in the family we focus on will not get married.

Total family income net of healthcare costs in terms of period 2 prices, y, is given by

$$y = w_m[N_m - D_m] + w_f[N_f - D_f^*] - (1+r)c[t_m i_m N_m + t_f i_f N_f]. \tag{3}$$

The first term on the right hand side of (3) is the income of surviving sons, the second term is by daughters-in-law, and the third term is healthcare costs. Since healthcare costs are incurred in period 1, they are multiplied by the interest factor $1 + r$.

Apart from utility from income, the families also suffer disutility from the death of children. In order to keep the analysis tractable, we assume a very simple rule for converting disutility from bereavement into monetary values, and write the net utility of the family, u, as

$$u = U(y - \phi(D_m + D_f)), \tag{4}$$

where $U' > 0,\ U'' < 0$, and ϕ is a constant parameter representing marginal disutility from bereavement.

We turn now to the treatment of uncertainty. We utilize the concept of certainty equivalence using the Markowitz's model of mean-variance analysis of portfolio selection. The certainty equivalence of the family's utility u^c is written as[5]

$$u^c = E(y - \phi(D_m + D_f)) - \gamma Var(y - \phi(D_m + D_f)), \tag{5}$$

where γ is the measure of relative risk preference. We assume that the economic agents are risk averse, so that $\gamma > 0$.[6]

[5] See, for example, Newbery and Stiglitz (1981, ch.6) for a discussion on this concept.

[6] As has been shown in Newbery and Stiglitz (1981), the formulation given in (5) does not need any approximation if the utility function is of a particular type and the random variable follows a Normal distribution. However, they have also shown that this formulation provides a good approximation with any utility function or distribution provided the variance of the random variable is small. We have chosen a Binomial distribution as it seems natural for the problem at hand. However, as it is know from the Central Limit Theorem that a binomial distribution is asymptotically Normal.

Since the number of severely sick boys and girls — X_m and X_f — are assumed to follow a binomial distribution, we have

$$\begin{aligned} E(X_m) &= N_m i_m p, && (6)\\ E(X_f) &= N_f i_f p, && (7)\\ Var(X_m) &= N_m i_m p(1-p), && (8)\\ Var(X_f) &= N_f i_f p(1-p). && (9) \end{aligned}$$

Using (1), (2) and (3), we write

$$\begin{aligned} y - \phi(D_m + D_f) &= (w_m N_m + w_f N_f) - (w_m + \phi)(\delta t_m + 1 - t_m) X_m \\ &\quad - [w_f(\delta t_f^* + 1 - t_f^*) + \phi(\delta t_f + 1 - t_f)] X_f \\ &\quad - (1+r)c[t_m i_m N_m + t_f i_f N_f], \qquad (10) \end{aligned}$$

and therefore

$$\begin{aligned} E(y - \phi(D_m + D_f)) &= (w_m N_m + w_f N_f) - (w_m + \phi)(\delta t_m + 1 - t_m) N_m i_m p \\ &\quad - [w_f(\delta t_f^* + 1 - t_f^*) + \phi(\delta t_f + 1 - t_f)] N_f i_f p \qquad (11) \\ &\quad - (1+r)c[t_m i_m N_m + t_f i_f N_f], \\ Var(y - \phi(D_m + D_f)) &= (w_m + \phi)^2 (\delta t_m + 1 - t_m)^2 N_m i_m p(1-p) \\ &\quad + [w_f(\delta t_f^* + 1 - t_f^*) + \phi(\delta t_f + 1 - t_f)]^2 N_f i_f p(1-p). \qquad (12) \end{aligned}$$

Substituting (11) and (12) into (5) and then taking partial derivatives of the resulting equation we obtain the first order conditions for t_m and t_f as

$$\begin{aligned} \frac{\partial u^c}{\partial t_m} &= (w_m + \phi) N_m i_m p(1-\delta) - (1+r)c N_m i_m \\ &\quad + 2(1-\delta)\gamma N_m i_m p(1-p)(w_m + \phi)^2 (\delta t_m + 1 - t_m) = 0, \\ \frac{\partial u^c}{\partial t_f} &= \phi N_f i_f p(1-\delta) - (1+r)c N_f i_f \\ &\quad + 2(1-\delta)\gamma N_f i_f p(1-p)\phi(w_f + \phi)(\delta t_f + 1 - t_f) = 0, \end{aligned}$$

which can be simplified as

$$p[(w_m + \phi) + 2\gamma(1-p)(w_m + \phi)^2(\delta t_m + 1 - t_m)] = \frac{(1+r)c}{1-\delta}, \tag{13}$$

$$p[\phi + 2\gamma(1-p)\phi(w_f + \phi)(\delta t_f + 1 - t_f)] = \frac{(1+r)c}{1-\delta}. \tag{14}$$

The right hand side of the above two equations are the marginal costs of child getting medical care (corrected for the quality of medical care). These costs are the same for boys and girls. The left hand side of the two equations are the marginal benefits. The second terms in the two equations arise via changes in the variance of income. The first terms differ between the two equations in a very substantive way. In (13), the first term represents two benefits. If a son does not die, then the family receives an income (w_m) and does not suffer disutility of bereavement (ϕ). However, in (14), which is the first order condition for daughters, the benefit from wage income is absent as the daughters become part of the in laws' families in period 2 and the income of daughters-in-law are taken as given in the optimization problem of the families.

From (13) and (14) the closed-form solutions the optimum levels of t_m and t_f are found as

$$t_m^o = \frac{p(w_m + \phi) + 2\gamma p(1-p)(w_m + \phi)^2 - \beta}{2\gamma p(1-p)(w_m + \phi)^2(1-\delta)}, \tag{15}$$

$$t_f^o = \frac{p\phi + 2\gamma p(1-p)\phi(w_f + \phi) - \beta}{2\gamma p(1-p)\phi(w_f + \phi)(1-\delta)}, \tag{16}$$

$$\frac{t_m^o}{t_f^o} = \frac{\{p(w_m + \phi) + 2\gamma p(1-p)(w_m + \phi)^2 - \beta\}(w_f + \phi)}{\{p\phi + 2\gamma p(1-p)\phi(w_f + \phi) - \beta\}(w_m + \phi)^2}, \tag{17}$$

where

$$\beta = \frac{(1+r)c}{1-\delta}.$$

From (15) and (16) the following two propositions follow

Proposition 1 *Ceteris paribus, a larger proportion of sick boys than girls receive medical treatment.*

Intuition behind proposition 1 follows from the discussions after (14). Since parents do not internalize the externality of daughters-in-law's income potential, changes in family income coming from the daughters-in-law do not appear in the first order condition associated with daughters. In other words, the marginal benefit of providing medical care is higher for sons than for daughters. The marginal costs, as mentioned before, are the same. Thus a larger proportion of sons get medical care than daughters. This bias would not have occurred if all the families coordinated their actions and took more of their sick daughters to hospitals/doctors with the understanding the parents of the future daughters-in law would be doing the same.

We now examine how changes in some of the parameters affect bias against girls in the provision of health care. In particular, we shall examine the effects of a change in either the cost of healthcare (c), the discount rate (r), or the quality of healthcare (δ) on the relative attention the sons and daughters receive, given by the ratio of dt_m^o and dt_f^o. From (17), it is clear that an increase in either of these three parameters can be represented by an increase in the parameter β.

Differentiating (15) and (16), and since $1-\delta$ in the denominators of (15) and (16) disappear when take the ratio of t_m^o and t_f^o, treating this term as constant, we get

$$2\gamma p(1-p)(1-\delta)\cdot\frac{dt_m^o}{d\beta} = -\frac{1}{(w_m+\phi)^2} < 0, \tag{18}$$

$$2\gamma p(1-p)(1-\delta)\cdot\frac{dt_f^o}{d\beta} = -\frac{1}{\phi(w_f+\phi)} < 0. \tag{19}$$

From (15), (16), (18) and (19) we obtain

$$\left.\frac{d(t_m^o/t_f^o)}{d\beta}\right|_{w_m=w_f} = \frac{t_m^o}{t_f^o}\left[\frac{dt_m^o}{d\beta}\cdot\frac{1}{t_m^o} - \frac{dt_f^o}{d\beta}\cdot\frac{1}{t_f^o}\right]$$
$$= \frac{w_m}{p\phi(w_m+\phi)[1+2\gamma(1-p)(w_m+\phi)]} > 0. \quad (20)$$

Formally,

Proposition 2 *Ceteris paribus, an increase in the unit cost of medical care (c) or the interest rate (r) increases the bias against girls in the provision of healthcare.*

The above result can be explained with the help of a diagram. In figure 1, the line $\bar{c}\bar{c}$ is the marginal cost of healthcare which same for both sons and daughters (the right hand sides of (13) and (14)). AA and BA are respectively the marginal benefit curve for sons and daughters. Note that AA is steeper than BA and they intersect the horizontal axis at the same point (A). The initial equilibrium values of the treatment rates for sons and daughters are given by t_m^o and t_f^o respectively, and it is clear that $t_m^o > t_f^o$. An increase in either c or r shifts the marginal cost line upwards to $\bar{c}^*\bar{c}^*$, and the resulting new equilibrium for the two variables are t_m^* and t_f^*. It should be clear from the diagram that the bias is more in the new equilibrium than in the initial one in the sense that the difference between the two variables is higher in the new equilibrium than in the old one, i.e., $t_m^* - t_f^* > t_m^o - t_f^o$.[7]

Figure 1 here

From (15), (16) and the above discussion, it should be clear that there would be no discrimination if $w_m = w_f$ and $c = \beta = 0$, i.e., health-care cost is zero and there is no discrimination in the labor market. The reason for this is that parents do not have to incur any cost for taking their daughters to health-care facilities and can potentially suffer

[7]Since $t_m^o > t_m^*$, it is also true that $t_f^o/t_m^o > t_f^*/t_m^*$.

disutility if the daughter dies because of lack of health care. An implication of this result is that discrimination will only occur insofar as it interacts with health-care cost. In the next section we shall test this hypothesis and the ones in proposition 2 using micro data from India.

3 Empirical Estimation

In this section, we empirically test the theoretical predictions of the model presented above. For this purpose we utilize a dataset from the World Bank's household and community surveys modeled after the Living Standards Measurement Study (LSMS) surveys and test if there is a bias against daughters in health-care provision and how this bias interacts with health-care costs.[8] The econometric methodologies we use are Ordinary Least Squares (OLS) and bivariate Logit analysis. This section is divided in three subsections. In the first subsection the dataset is discussed. Then subsections 3.2 and 3.3 discuss results from OLS and Logit regressions. Finally, in section 3.4 we examine if discrimination against daughters is more among Muslim families than among Hindu families.

3.1 The Data

The data was collected by the World Bank from a two-part study of rural poverty carried out in 1997-98 in south and eastern Uttar Pradesh and north and central Bihar. The study utilized both qualitative methods such as rapid rural appraisal (RRA), participatory rural appraisal (PRA) methodologies, and semi-structured interviews as well as quantitative methods drawing upon the data collected. The available data are from the quantitative component of the study. The data was collected through household and village level questionnaires in 120 villages from a sample of 25 districts in Uttar Pradesh and Bihar. A total of 2,250 households were interviewed covering over 13,000 individual interviews. For the

[8]The source of the dataset is the *Uttar Pradesh and Bihar Survey of Living Conditions*. This dataset can be downloaded from the World Bank website: http://www.worldbank.org/lsms/guide/select.html.

purpose of this paper we utilize some information on the socio-economic characteristics of households, some information at the village level, and, most importantly, health information on individual children within each household. The health information utilized here include illnesses of individual children within a family such as the quality of care that the sick children received, the cost of treating each sick child, and the type of illness being treated. Illness data is quite detailed and comes in eleven different categories - from relatively less severe illnesses such as injury, fever,[9] diarrhea and cataract or problems with eye sight to mental illness, respiratory problems, tuberculosis, blood pressure, heart problems and permanent disability. The quality of care or the type of treatment also has several categories - from faith healer and quack to village nurse, government doctor and private doctor. For simplicity, all the non-traditional forms of medical care have been given a score of 0 and the traditional ones have been scored in ascending order of importance.[10] Data on health expenditure or the cost of healthcare is the amount spent by a family for a particular child over the time period of a year.

The final sample consists only of unmarried children and grandchildren who have suffered some kind of illness in the year prior to the survey. All adults and servants have been eliminated. The most common illness reported was fever followed by diarrhea. About 42% of the children reported fever and about 16% reported diarrhea. In terms of the types of treatment or quality of care received by the sick children, about 40% received traditional treatment while 60% were taken for some method of non-traditional treatment. That is, 60% of the sick children were treated by a quack, an indigenous practitioner, a village chemist, or a faith healer. Among all different types of treatment quacks were the most sought after. About 48% of sick children were taken to quacks to receive treatment. In terms of the incidence and distribution of illness, out of 1,993 families with sick children, on average each family had about 3 sick girls and about 4 sick boys, but only 0.6 girls and 0.9 boys received

[9]Fever is considered severe if it has afflicted the patient for over a month.

[10]All government doctors have been classified under one category though the original data had three different categories of government doctors based on where the doctor was practicing medicine.

traditional treatment. The average amount spent on traditional healthcare per person was around 128 Rupees with total average household spending on medical care being about 860 Rupees.

In terms of the variables used for household characteristics, we have used household head or father's age, mother's age, father's and mother's education, caste, the type of home structure, and the number of people that live in the household. Additionally, we have also tried to isolate whether gender bias was more prevalent in the poorer state of Bihar compared to Uttar Pradesh. On average, the age of a household head or father was 47 years, the average mother's age was 42 and the average child's age (includes children and/or grandchildren) was 8.8 years. For the sample as a whole, about 50% of all household heads were illiterate, about 80% of mothers were illiterate and about 43% of the children were illiterate. The average education level of the household head was between being literate with no formal schooling to less than primary education. Mothers' average education was somewhere between being totally illiterate to being literate with no formal schooling. For the children, the average education was the same as the household head except it was skewed toward being literate with no formal schooling rather than toward less than primary education.

Tables 1 and 2 here

Caste is broken down by religion in this sample. The two religions are Hindus and Muslims with Muslims making up 11% of the population. The Muslim's have two castes,[11] an upper and one backward caste while the Hindus have 5 castes with the lowest being the scheduled castes and scheduled tribes who were traditionally known to be the untouchables. The largest concentration of people was in the scheduled castes and scheduled tribes accounting for about 26% of the total population. The average household size in this sample was a little over 8 people in each household. In the final sample there are close to 2000 households and a little over 13,000 sick children and grandchildren. Only 1200 children, including

[11] Although Islam officially has no caste, but *de facto* the caste system is practiced by all groups of Indians.

grandchildren, were not ill in the year preceding the survey. In the absence of any wealth data and poorly reported income data, the type of home structure was used as a proxy for family wealth. These also show a lot of variation across the sample. The home structures are classified under 5 categories ranging from a thatched and completely temporary structure to a permanent and stable structure. The average structure/dwelling for this sample ranged between a semi-permanent and semi-temporary structure. The largest concentration of people is found in the group that lives in katcha/tile housing which falls short of being semi-temporary and just above the poorest group with completely temporary structures. Thus, while the wealth varied across the sample, on average, the people participating in the survey were quite poor.

We also use a variable that represents village-level characteristics. The survey asked how far one would have to travel to receive treatment for five different types of treatment: (i) complicated surgery, (ii) injections, (iii) minor surgery (iv) treatment of broken bones, and (v) treatment of TB. We created a variable called 'distance' which is the average distance over the five types. The mean value of this variable is 14.2 km for Bihar and 20.6 km for UP (see Table 1 and 2 above).

3.2 Ordinary Least Squares

We now turn to econometric analysis. As we mentioned before, we shall both OLS and Logit regressions. In this section we consider OLS; Logit will be taken up in section 3.3.

In this subsection the dependent variable is the quality of care or treatment type (treattype).[12] As explained earlier and in the Appendix, treattype is determined using a scoring system. The scoring system indicates that the higher the score the better the quality of care. The dependent variable is not a gender specific variable, and we account for the gender bias through some of the explanatory variables described below.

[12]All the variables used in our regression analysis are described in the Appendix.

In our estimation we have used a number of control variables and those will be discussed a little later. But, the two most central variables are a male dummy — a dummy variable that takes the value 1 if the child is male, and 0 otherwise — and a child-specific health-care cost variable (healthcost). These two variables help in tracking gender bias, if any, and in testing the hypotheses that we developed earlier. As we noted in our theoretical section, in the absence of any health-care cost, there will be no bias against girls. Our theory also predicts that bias, if any, would work only via an interaction with health-care cost. Thus, to test the hypothesis that there is no bias we introduce an interaction term, mhcost, which is a product of the male dummy and healthcost.

We have run a number of OLS regressions based on variants of the following equation:

$$\begin{aligned}\text{treattype} &= \beta_0 + \beta_1 \text{ healthcost} + \beta_2 \text{ mhcost} + \beta_3 \text{ mhcostbh} + \beta_4 \text{ male}\\ &\quad \beta_5 \text{ illtype} + \beta_6 \text{ hhsize} + \beta_7 \text{ headedu} + \beta_8 \text{ medu} + \beta_9 \text{ headage}\\ &\quad + \beta_{10} \text{ mage} + \beta_{11} \text{ hometype} + \beta_{12} \text{ caste} + \beta_{13} \text{ Hindu} + \beta_{14} \text{ distance},\end{aligned}$$

where some of the variables have already been discussed. Additionally, hhsize refers to the number of people in each family, headedu is the education level of the household head or father, medu is the education level of the mother, headage is the age of the household head, illtype is a score variable representing the severity of illness of a child, mage is the age of the mother,[13] hometype refers to the kind of structure that the family lives in, caste refers to which caste the family belongs to, Hindu is dummy variable which takes the value 1 if the child belongs to a Hindu family,[14] distance (which is a village-level characteristics) is the average distance of the village from a medical facility, and mhcostbh is a product of male, healthcost and Bihar, the latter being a dummy variable representing households that reside in the state of Bihar.

[13]there are 2166 male headed households and 87 female headed households. However when we created the headedu and headage variables we considered the education and age of the head and where we say medu and mage we considered the education and age of the spouse of the head. Thus, in a very few cases, the medu and mage variables actually related to the spouse of the female head.

[14]In our sample, 93% of the families are Hindus in UP, 86% in Bihar. The overall percentage of Hindu families is 90.

The results are presented in Table 3. Inter-firm heteroskedasticity is taken into account in the estimations.

Table 3 here

The OLS results give strong support to the theory presented in this paper. We see that the coefficient of male dummy is consistently positive, but insignificant. In these equations the coefficient of the healthcost is consistently negative and significant implying that an increase in the cost of care reduces the quality of care a child – boy or girl – receives. More importantly, we also find that the coefficient of the interaction variable mhcost is positive and highly significant in all the regressions. This result is found to be robust to changes in model specification. This has two implications. First, there is a bias against girls in health-care provision as long as health-care cost is positive. Second, the degree of bias increases as health-care cost increases. This supports proposition 2. Finally, on the interaction terms, the coefficient for mhcostbh is significant and positive (though only at 90% confidence level) implying that, as long as health-care cost is positive, the extent of bias is higher in Bihar than in UP.

All the control variables also have the expected signs. Illtype is positive and significant implying that the more serious the illness is better the is the type of treatment the child receives.[15] Hometype, which controls for the level of wealth/income for a family, has a positive and often significant impact, as one would expect.[16] Father's and mother's level of education have significant positive effects on the quality of healthcare a child receives. Father's age has a significant positive impact while household size and mother's age have negative impact on the quality of health care. In other words, younger mothers seem to be wiser than the older ones. The caste variable is also positive and significant implying that

[15] We have tried alternative scoring system for illtype, giving the category "other" (see Table I in the Appendix), a score 5. Since none of the results changed qualitatively, they are not reported here.

[16] As an alternative, we have also tried landholding as a control for wealth/income. Since the results are very similar qualitatively, they are not reported here.

the higher the caste of the sick child the more likely that the child receives a better quality healthcare. Finally, the Hindu dummy has a negative and significant effect implying that a Hindu family, on an average, goes to a less qualified medical professional than a Muslim family after controlling for a whole gamut of economic and socio-economic factors. We did not have any prior expectation on this effect.

3.3 Bivariate Logit

Although the OLS regression results strongly support our theoretical predictions, the econometric methodology, for the problem at hand, can be subject to criticisms. One reason for such a criticism is that decisions facing the families are discrete ones, selecting from a finite number of discrete alternatives. Another potential criticism is that our dependable variable treattype is a score variable (see Table II in the Appendix), and there is some degree of arbitrariness in the choice of the equi-distant scores. In this section, we shall therefore carry out our analysis using a binomial logit estimation process. In particular, we assume that a family chooses one of the following two options: (a) to take a sick child to a non-traditional person such as a village quack, and (b) to take a sick child to a qualified medical professional for treatment. The choice (a) which is given the number 0 represents the first four type of treatment in Table II, and choice (b) is given the number 1 and represents the last four categories.

We estimate the Binomial Logit model specified as follows.

$$P(y = j/x) = \frac{\exp{(x'\alpha_j)}}{1 + \sum_{k=0}^{1} \exp^{(x'\alpha_k)}}, \qquad (j = 0, 1).$$

The choice of the firm is denoted by y. As for the regressor, vector x explaining a firm's choice y, we use the same explanatory variables used in the OLS analysis.

The estimation results from the binomial response model are reported in Table 4. Like in OLS regressions, inter-firm heteroskedasticity is taken into account in the estimations. As the response probability of the choice j relative to the base category (choice (a)) is given by $p_j/p_0 = \exp(x'\alpha_j)$, the significantly positive estimate mhcost indicates that parents are more likely to take their sons for better types of treatment than their daughters. In general all the coefficients are qualitatively the same – both in terms of their signs and significance levels – as in the OLS regressions.[17] Therefore, the robustness of our results carries through to the present case of binomial Logit analysis.

Table 4 here

3.4 Do Muslims discriminate more than Hindus?

Since our sample represents two distinct religious groups, *viz.*, Hindus and Muslims, with the latter forming a significant minority: 11% of the observations, one may be tempted to ask if there is any difference in the behavior of the two groups with regard to their daughters *vis-à-vis* their sons. It is possibly not wrong to say that the conventional wisdom — at least outside the Islamic World — is that bias against women is more pronounced among Muslims than among other religious groups. In this subsection, we look into this question.

Before turning to the empirical questions, we should note that our theoretical model would predict a *lower* bias among Muslim families than among the Hindu ones. This is because, in our theoretical model, the presence of inter-family externality in family decision-making process leads to bias against daughters insofar as health-care provision is concerned. This externality would be internalized if different families coordinate their actions. Since whereas Hindus in Northern India tend to practice exogamy in marriages, intermarrying is very common among Muslim, one would expect more coordination among Muslim families than among Hindu ones. This should lead to less bias among Muslims and Hindus.

Returning to the empirical issues, an obvious approach for testing the existence of possible differential bias among the two religious groups would have been to include another

[17]The only exception is the coefficient for the Bihar dummy (mhcostbh) which is insignificant here.

variable which is the product of three variables healthcost, male dummy and Hindu dummy, and look at the sign of its coefficient. A significant negative value of this coefficient would have confirmed the conventional wisdom. Unfortunately, however, we were unable to do so as this variable and mhcost (which is the product of only the first two of the three variables) are highly correlated with the correlation coefficient being 0.94.

As a first step in our analysis in this regard, given that we cannot use the dummy-variable technique, we computed female-to-male ratios for the two groups separately and find, surprisingly, that the ratio is 0.88 for Hindus and 0.93 for Muslims. We therefore find *prima facie* evidence that the conventional wisdom may not be true after all.

The next step in our analysis is to divide the total sample in to two subsamples, one consisting observations on Hindus only and the other Muslims, and then run the regressions separately for the two groups. We have run both OLS and Logit regressions as in the preivious subsections, and all the specifications therein. However, since all the results are very similar, for the sake of brevity in Table 5 we only present a selection of estimations: Logits regressions corresponding to columns (5)-(8) in Table 4.

As can be seen from Table 5, the results are very similar to those in Tables 3 and 4, with one important difference.[18] The coefficient for mhcost is statistically significant (at 99% confidence level) for Hindus, but insignificant (even with 75% confidence level) for Muslims. This implies that the evidence is is in sharp contrast to conventional wisdom: there is bias against daughters among Hindu families, but there is not enough evidence to suggest that Muslim families discriminate against their daughters insofar as health-care provision is concerned. More interestingly, these findings are consistent with the predictions of our theoretical analysis.

Table 5 here

[18]It is also interesting to note that the coefficients of household size (hhsize) are insignificant for Hindus, but are significantly negative for Muslims. In contrast, the coefficient for healthcost is significant for Hindus and significantly negative for Muslims.

4 Conclusion

Much has been written on possible bias against women and daughters and their consequences, in Asia. There are many mechanisms via which a bias can manifest itself. One such form could be in the provision of healthcare for sick children. In this paper, we first of all provided a theoretical explanation for why a family may treat their sons and daughters differentially in health-care provision, and tested our theoretical predictions with micro data from two provinces of India.

There are two important elements to our theoretical model. The first is specific to the nature of the problem at hand, *viz.*, healthcare. In particular, the fact that often parents cannot be sure about the extent of severity of an illness of a child, is important for our analysis. The second element has to do with a particular social institution that is prevalent in South Asia, *viz.*, a woman, once married, becomes a part of the in-law's family. We show that these two aspects lead to the presence of inter-family externality in family decision-making process and this in turn leads to bias against daughters insofar as health-care provision is concerned. It is to be noted that the externality would be internalized if different families coordinate their actions. However, coordination may be difficult for a society in which exogamy is practiced in marriages. We also find that the bias exists only in the presence of positive health-care costs, and the extent of the bias increases as costs increase.

The data for our empirical analysis come from the World Bank's household LSMS surveys carried out in 1997 and 1998 in south and eastern Uttar Pradesh and north and central Bihar. Both these states are toward the north of the country where the social institutions fit very well with our theoretical construct. Econometric analysis using both an Ordinary Least Squares method and a binomial Logit analysis give strong backing to our theoretical predictions. In particular, we find strong evidence of a bias against daughters in health-care provision. Additionally, the empirical results show that gender bias gets worse as the cost of care increases. Finally, we find that whereas the bias is significant among Hindu families who tend to practice exogamy, there is not enough evidence to suggest that bias exists among Muslim families who often intermarry.

One of the policy implications of our analysis is that it is not only important that the quality of healthcare improves, but it is also very important that while such improvements take place governments should make sure that costs of quality health-care provision to poorer sections of the population is kept as low as possible.

Figure 1: The Discrimination Equilibrium

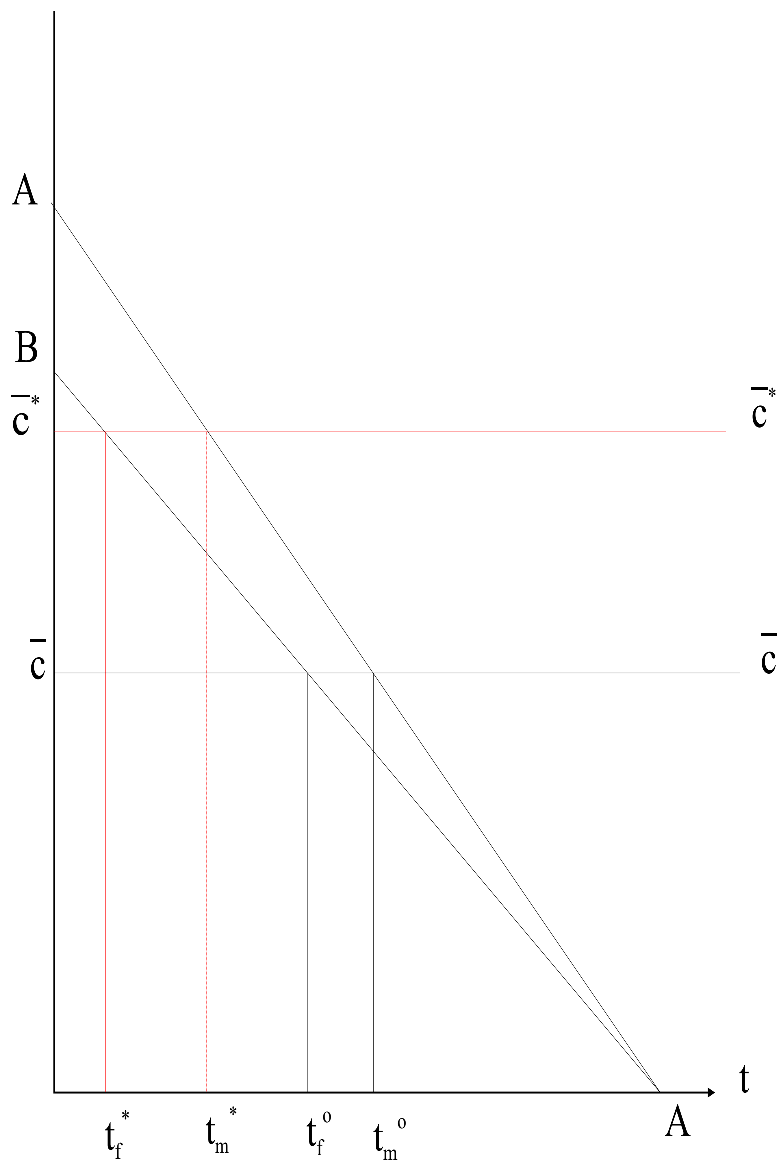

Table 1: Summary statistics for UP

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
illtype	4662	3.512227	2.20553	1	11
hometype	7151	2.852608	1.379705	1	5
landholding	6464	3.898864	6.774413	0	93
healthcost	7162	198.0595	1361.904	0	35000
headage	7146	49.5883	13.99864	7	95
mage	6300	43.72571	12.81158	7	90
headedu	7146	2.845788	2.343075	1	11
medu	6300	1.47746	1.431164	1	11
distance	7152	20.60733	11.35347	5.2	72.8
hhsize	7227	8.453438	4.331367	1	29
hhkids	7162	4.248255	2.362741	1	17

Table 2: Summary statistics for Bihar

Variable	Obs.	Mean	Std. Dev.	Min.	Max.
illtype	3089	4.295565	2.533912	1	11
hometype	6154	2.155671	1.302142	1	5
landholding	4271	2.319766	3.145107	0	20
healthcost	6161	209.2282	1231.977	0	30000
headage	6161	47.11816	12.92628	3	87
mage	5694	41.85739	12.35055	0	80
headedu	6161	3.057621	2.265729	1	11
medu	5 694	1.574113	1.32098	1	8
distance	6121	14.24454	8.687809	4.4	47.7
hhsize	6202	7.57288	3.273768	1	25
hhkids	6161	3.860737	1.969677	1	15

Table 3: OLS Regressions

Dependent Variable: Treatment Type (treattype)

	1	2	3	4	5	6	7	8
Const.	1.214*	1.176*	1.022*	0.979*	1.071*	0.997*	1.322*	1.284*
	(17.7)	(15.8)	(13.1)	(10.0)	(11.0)	(10.1)	(11.2)	(11.0)
Male	0.058	0.060	0.055	0.064	0.059	0.059	0.301	0.038
	(1.4)	(1.5)	(1.3)	(1.5)	(1.3)	(1.3)	(0.7)	(0.9)
Illtype	0.139*	0.138*	0.133*	0.139*	0.139*	0.139*	0.136*	0.136*
	(16.0)	(16.0)	(14.5)	(14.8)	(14.9)	(14.8)	(14.5)	(14.5)
healthcost	-0.156*	-0.149*	-0.164*	-0.150*	-0.165*	-0.159*	-0.161*	-0.150*
	(3.2)	(3.0)	(2.9)	((2.9)	(3.1)	(3.0)	(3.1)	(2.9)
hometype	0.063*	0.058*		0.055*		0.058*	0.033***	0.028***
	(4.1)	(3.7)		(3.2)		(3.4)	(1.9)	(1.7)
mhcost	0.0001*	0.0001*	0.0001*	0.0001*	0.0001*	0.0001*	0.0001*	0.0001*
	(5.7)	(5.6)	(4.2)	(4.8)	(4.8)	(4.8)	(4.6)	(2.5)
distance	-0.020*	-0.020*	-0.018*	-0.020*	-0.019*	-0.019*	-0.017*	-0.018*
	(11.2)	(11.1)	(9.7)	(10.5)	(9.8)	(10.2)	(9.2)	(9.3)
hhsize		0.006	-0.001		-0.004	-0.007	-0.012**	
		(1.2)	(0.1)		(0.8)	(1.4)	(2.1)	
headedu			0.104*					
			(9.9)					
medu			0.049*					
			(2.9)					
headage				0.020*	0.019*	0.020*	0.013**	0.014**
				(3.4)	(3.3)	(3.4)	(2.4)	(2.4)
mage				-0.016*	-0.013**	-0.015**	-0.009	-0.011***
				(2.5)	(2.2)	(2.4)	(1.5)	(1.8)
Caste							0.529*	0.523*
							(9.5)	(9.3)
Hindu							-0.318*	-0.309*
							(4.6)	(4.5)
mhcostbh								0.0001***
								(1.6)
$\bar{R}^2$	0.06	0.06	0.08	0.06	0.06	0.06	0.08	0.08
F-Stat.	79.9	68.3	76.1	55.2	53.2	49.4	53.4	52.3
$\sqrt{\text{MSE}}$	1.75	1.75	1.75	1.76	1.76	1.76	1.75	1.75
Obs. #	7623	7620	6787	6782	6787	6779	6779	6779

* Significant at 99% level of confidence
** Significant at 95% level of confidence
*** Significant at 90% level of confidence

Table 4: LOGIT Regressions

	1	2	3	4	5	6	7	8
Const.	-0.679* (8.3)	-0.735* (8.2)	-0.897* (9.4)	-0.887* (7.6)	-0.767* (6.7)	-0.870* (7.4)	-0.552* (4.0)	-0.593* (4.3)
Male	0.049 (1.0)	0.052 (1.1)	0.046 (0.9)	0.061 (1.2)	0.055 (1.1)	0.056 (1.1)	0.028 (0.5)	0.031 (0.6)
Illtype	0.147* (14.1)	0.147* (14.1)	0.139* (12.6)	0.144* (12.9)	0.143* (13.0)	0.144* (12.9)	0.143* (12.7)	0.143* (12.7)
healthcost	-0.213* (3.5)	-0.202* (3.3)	-0.208* (3.2)	-0.200* (3.1)	-0.216* (3.3)	-0.208* (3.2)	-0.215* (3.3)	-0.212* (3.2)
hometype	0.087* (4.7)	0.080* (4.3)		0.077* (3.9)		0.081* (4.0)	0.055* (2.7)	0.051* (2.5)
mhcost	0.0001* (3.9)	0.0001* (3.8)	0.0001* (3.3)	0.0001* (3.5)	0.0001* (3.5)	0.0001* (3.5)	0.0001* (3.5)	0.0001* (2.5)
distance	-0.022* (9.1)	-0.022* (9.1)	-0.019* (7.8)	-0.021* (8.4)	-0.019* (7.8)	-0.020* (8.2)	-0.019* (7.4)	-0.019* (7.4)
hhsize		0.009 (1.5)	-0.001 (0.2)		-0.003 (0.4)	-0.007 (1.1)	-0.012*** (1.7)	
headedu			0.129* (10.3)					
medu			0.056* (2.7)					
headage				0.024* (3.2)	0.023* (3.2)	0.024* (3.3)	0.017* (2.6)	0.017* (2.6)
mage				-0.021* (2.6)	-0.018** (2.3)	-0.020* (2.5)	-0.014** (1.9)	-0.015** (2.1)
Caste							0.546* (8.5)	0.539* (8.4)
Hindu							-0.313* (4.0)	-0.302* (3.8)
mhcostbh								0.0001 (1.5)
Pseudo R^2	0.04	0.04	0.05	0.04	0.04	0.04	0.05	0.05
Wald χ^2	324.0	324.9	413.0	288.6	274.9	288.8	365.2	358.3
Obs. #	7623	7620	6787	6782	6787	6779	6779	6779

* Significant at 99% level of confidence

** Significant at 95% level of confidence

*** Significant at 90% level of confidence

Table 5: LOGIT Regressions for Hindus and Muslims

	1 Hindu	1a Muslim	2 Hindu	2a Muslim	3 Hindu	3a Muslim	4 Hindu	4a Muslim
Const.	-0.716* (5.9)	-0.895* (2.5)	-0.825* (6.5)	-1.128* (3.1)	-0.780* (6.1)	-1.045* (2.8)	-0.788* (6.3)	-1.230* (3.4)
Male	0.040 (0.7)	0.133 (0.8)	0.039 (0.7)	0.129 (0.8)	0.019 (0.4)	0.089 (0.5)	0.015 (0.3)	0.172 (1.0)
Illtype	0.131* (11.3)	0.240* (4.6)	0.131* (11.2)	0.235* (4.5)	0.132* (11.3)	0.245* (4.7)	0.131* (11.2)	0.238* (4.9)
healthcost	-0.240* (3.5)	-0.016 (0.1)	-0.229* (3.3)	-0.041 (0.2)	-0.242* (3.5)	-0.038 (0.2)	-0.254* (3.6)	-0.099 (0.5)
hometype			0.080* (3.8)	0.305* (4.4)	0.043** (2.0)	0.283 (3.9)	0.045** (2.1)	0.163 (2.7)
mhcost	0.00012* (3.5)	0.00005 (0.9)	0.00012* (3.4)	0.00006 (1.0)	0.00012* (3.4)	0.00008 (1.2)	0.00008* (2.5)	0.00002 (0.01)
distance	-0.016* (6.3)	-0.044* (5.1)	-0.017* (6.8)	-0.050* (5.7)	-0.015* (5.7)	-0.056* (6.3)	-0.014* (5.6)	-0.059* (6.7)
hhsize	0.008 (1.1)	-0.074* (3.8)	0.004 (0.51)	-0.098* (5.0)	-0.001 (0.13)	-0.088* (4.4)		
headage	0.022* (2.8)	0.059* (3.3)	0.023* (2.8)	0.069* (3.5)	0.018** (2.3)	0.069* (3.6)	0.018** (2.4)	0.057 (3.3)
mage	-0.020** (2.4)	-0.033*** (1.7)	-0.022* (2.6)	-0.047** (2.2)	-0.019** (2.2)	-0.052* (2.5)	-0.019** (2.3)	-0.047* (2.5)
Caste					0.548* (7.8)	0.682* (3.9)	0.543* (7.8)	0.794* (4.6)
mhcostbh							0.0001 (1.6)	0.0001 (0.3)
Pseudo R^2	0.03	0.13	0.03	0.14	0.04		0.04	0.17
WALD χ^2	201.8	90.1	212.6	97.4	262.0		257.5	116.1
Obs. #	6035	752	6027	752	6027		6027	752

* Significant at 99% level of confidence
** Significant at 95% level of confidence
*** Significant at 90% level of confidence

APPENDIX

Definition of Variables

Variable Name	Definition
pdtreatment	takes the value 1 if the child is taken to a traditional doctor for treatment and 0 otherwise.
male	takes the value 1 if the child is male and 0 otherwise
healthcost	amount spent by a family for a particular child over the time period of a year
illtype	Score variable on type of illness (see Table I below)
treattype	Score variable on the type of treatment (see Table II below)
mhcost	Interaction between 'male' and 'healthcost'
distance	Average distance from the village to five types of medical facilities
hometype	Score variable on the quality of the house the child lives in (see Table IV below)
hhsize	total number of people in the family
headage	age of father
mage	age of mother
headedu	Score variable on father's education level (see Table III below)
medu	Score variable on mother's education level (see Table III below)
caste	takes the value 1 if the child in high or middle-high cast, and 0 otherwise
Bihar	takes the value 1 if the child in from Bihar, and 0 otherwise
religion	takes the value 1 if the child in Hindu and 0 otherwise
mcostbh	Interaction between 'male', 'healthcost' and 'Bihar'

Scoring System

Table I

Illness Type	Score
Injury	1
Fever	2
diarrhea	3
cataract	4
mental	5
respiratory	6
Tuberculosis	7
blood pressure	8
heart problem	9
perm disability	10
other	11

Table II

Treatment type	**Score**
non-traditional:	
indigenous	0
faith healer	0
quack	0
chemist	0
traditional:	
charitable doctor	1
village nurse	2
government doctor	3
private doctor	4

Table III

Education	**Score**
illiterate	1
literate but no formal schooling	2
less than primary	3
primary	4
middle	5
matriculate	6
intermediate	7
bachelor's degree	8
masters	9
professional degree	10
diploma	11

Table IV

Home type	**Score**
"katcha/thatch" – temp structure	1
"katcha/tile" – temp but more stable	2
"semi pucca" – relatively stable	3
"pucca w/weaker sector" – relatively stable with weak areas	4
"pucca" – stable	5

References

[1] Bardhan, P, 1974, 'On Life and Death Questions', *Economic and Political Weekly* 9, 1293-1305.

[2] Basu, K., 2005, Gender and Say: A Model of Household Behavior with Endogenously-Determined Balance of Power,' forthcoming in *Economic Journal*.

[3] Behrman, J., 1988, 'Intrahousehold Allocation of Nutrients in Rural India: Are Boys Favored? Do Parents Exhibit Inequality Aversion?', *Oxford Economic Papers*, New Series, 40(1), 32-54.

[4] Boserup, E., 1980, *Women's Role in Economic Development*, New York: St. Martin's Press Ltd.

[5] Coale, A., 1991, 'Excess female mortality and the balance of the sexes: an estimate of the number of 'missing females',' *Population and Development Review*, 17, 517-523.

[6] Croll, E., 2001, 'Amartya Sen's 100 Million Missing Women', *Oxford Development Studies*, 29(3), 225-244.

[7] Dollar, D., and R. Gatti, 1999, 'Gender Inequality, Income, and Growth: Are Good Times Good for Women?', Policy Research report on Gender and Development, Working Paper Series No. 1, World Bank Development Research Group, World Bank.

[8] Drèze, J. and A. Sen, 1991, *Hunger and Public Action*, Oxford: Oxford University Press, ch. 4.

[9] India Registrar General, 1995, 'SRS Based Abridged Life Tables 1988-92', Occasional Paper No. 4, New Delhi, India.

[10] Johnson, C., 1996, ‘Violence Against Women in India: The Law as Protector?’, paper presented at the Association of Women in Development International Conference, Washington, D. C.

[11] Klasen, S., 1994, ‘Missing women reconsidered,’ *World Development*, 22, 1061-1071.

[12] Klasen, S., and C. Wink, 2002, ‘A Turning Point in Gender Bias in Mortality? An Update on the Number of Missing Women, *Population and Development Review*, 28(2), 285-312.

[13] Newbery, David M.G. and Joseph E. Stiglitz, 1981, *The Theory of Commodity Price Stabilization: A study in the Economics of Risk*. The Clarendon Press, Oxford.

[14] Okojie, C., 1994, ‘Gender Inequalities of Health in the Third World’, *Social Science Medicine*, 39(9), 1237-1247.

[15] Oster, E., 2005, ‘Hepatitis B and the case of the missing women,’ mimeo., Harvard University.

[16] Prasad, S., 1996, ‘Instituting Measures to Address Violence Against Women in India’, paper presented at the Association of Women in Development International Conference, Washington, D. C.

[17] Quisumbing, A. and J. Maluccio, 2000, ‘Intrahousehold Allocation and Gender Relations: New Empirical Evidence From Four Developing Countries’, FCND Discussion Paper No. 84, International Food Policy Research Institute, Washington, D.C.

[18] Sen, A. 2003, ‘Missing women - revisited’, *British Medical Journal*, 327, 1297-1298.

[19] Sen, A, 1992, ‘Missing Women’, *British Medical Journal*, 304, 586-587.