

Revised July 8, 2002

FACTORS AFFECTING SEX-SELECTIVE ABORTION IN PUNJAB, INDIA

Robert D. Retherford
East-West Center
Honolulu, Hawaii, USA

T. K. Roy
International Institute for Population Sciences
Mumbai, India

Revised version of a paper presented at the 20th Population Census Conference, Ulaanbaatar, Mongolia, 19–21 June, 2002

This chapter examines the effects of demographic and socioeconomic factors on sex-selective abortion in the state of Punjab. Because reliable statistics on sex-selective abortion do not exist, the sex ratio at birth is used as an indirect indicator of sex-selective abortion.

Abortion has been legal in India since 1972, when the Medical Termination of Pregnancy Act became law. The Act specifies contraceptive failure as one of several legitimate grounds for abortion, in effect legalizing abortion on demand. The Act requires that abortions take place in government-approved facilities, but these are in short supply. Partly because this, the requirement that abortions take place in approved facilities is not enforced. It has been estimated that 5 to 6 million abortions occur annually in India and that roughly 90 percent occur in unapproved facilities (Khan et al. 1996).

Sex-selective abortion occurs in two steps. The first step is assess the sex of the fetus. The second step is to obtain an abortion if the fetus is not of the desired sex. Methods for determining the sex of a fetus became available in the 1970s. Three such methods are commonly used: amniocentesis (normally performed after 15–17 weeks of pregnancy), chorionic villus sampling (more expensive and normally performed around the 10th week of pregnancy), and ultrasound (least expensive and normally performed around the 12th week of pregnancy). In some parts of India, ultrasound is offered illegally by traveling vans. Shortly after the introduction of these tests in India in the 1970s, advertisement and provision of the tests began to spread. Kusum (1993; cited in Patel, n.d.) reports billboards saying “Invest Rs. [rupees] 500 now, save 50,000 later,” designed to encourage prospective parents to abort female fetuses in order to avoid future dowry expenses. Posters in train stations advertised sex-determination tests together with an abortion for as little as Rs. 70.

The Indian government opposes sex-selective abortion, but it took a long time to pass legislation to combat it. In 1994 the Union government enacted the Prenatal Diagnostic Techniques Regulations and Prevention of Misuse Act, which mandates that prenatal diagnostic tests can be conducted only for high-risk pregnant women and only for the purpose of detecting genetic abnormalities in the fetus. The law is easy to circumvent, however, and it is not clear that it has had much effect on the practice of sex-selective abortion.

The prevalence of sex-selective abortion is high in a number of states, especially those characterized by a strong preference for sons; but because of the illegal nature of the procedure, it is not known precisely how high. Sex-selective abortion manifests itself, however, in an altered sex ratio at birth (defined here as the ratio of male births to female births), which in the absence of sex selection is about 1.05. Thus it is possible to study trends and differentials in prevalence of sex-selective abortion indirectly through an analysis of trends and differentials in the sex ratio at birth. This is the approach taken in this chapter.

According to India’s Sample Registration System (SRS), among Indian states during the period 1996–98, the northern states of Punjab and Haryana had the highest sex ratio at birth at 1.23, considerably higher than the all-India average of 1.11 (Premi, in this volume). Our analysis focuses on the state of Punjab. The analysis examines how various demographic and socioeconomic factors affect (or do not affect) the sex ratio at birth in the state, based mainly on data for Punjab from India’s first and second National Family Health Surveys (NFHS-1 and NFHS-2). The analysis is based mainly on data from the two NFHS surveys rather than the SRS partly because the SRS lacks information on many of the variables of interest, and partly because the SRS data are not available in individual-record format and are therefore not amenable to multivariate analysis.

DATA

In Punjab, NFHS-1 was conducted in 1993, and NFHS-2 was conducted in 1998–99. Both surveys were designed along the lines of the Demographic and Health Surveys (DHS) that have been conducted worldwide in many developing countries since the 1980s. Both NFHS-1 and NFHS-2 are nationally representative surveys that include a household sample, covering everyone in the sampled households, and an individual sample, covering all ever-married women age 15–49 (13–49 in the case of NFHS-1) within those households. Corresponding to these two samples are a household questionnaire and an individual questionnaire. In the case of the household questionnaire, the household head or other knowledgeable adult in the household responded for the entire household. In the case of the individual questionnaire, each ever-married woman responded for herself and her children. NFHS-1 and NFHS-2 were designed to provide state-level estimates as well as national-level estimates. In each of the two surveys, the sampling fraction varies from state to state, in order to assure that the sample size in each state is large enough to provide statistically meaningful estimates. Basic survey reports are published for each state as well as for the nation as a whole.

In Punjab, NFHS-1 comprises 3,213 households and 2,995 ever-married women, and NFHS-2 comprises 2,967 households and 2,796 ever-married women. Details of sample design are contained in the basic survey reports for NFHS-1 and NFHS-2 in Punjab (IIPS and PRC Chandigarh 1995; IIPS and ORC Macro 2001).

METHODS

Units of analysis

As mentioned, the sex ratio at birth is used as an indirect indicator of sex-selective abortion. The analysis focuses on births that occurred during the 15-year period before each of the two surveys. A 15-year period is used in order to minimize bias resulting from sex-selective displacement of births to earlier years in the retrospective birth histories. In a previous paper, Retherford and Mishra (2001) have analyzed how displacement biases fertility estimates derived from NFHS-1 and NFHS-2. It turns out that, in Punjab at least, displacement is sex-selective and therefore also biases estimates of the sex ratio at birth for short time intervals.

The distorting effects of displacement on estimates of fertility and the sex ratio at birth for Punjab are apparent in Figures 1 and 2. Figure 1 shows overlapping 15-year trends in the total fertility rate (TFR) estimated from the two surveys, NFHS-1 and NFHS-2. The TFR is defined as the number of births that a woman would bear if, hypothetically, she lived through the reproductive ages 15–49 experiencing the age-specific fertility rates pertaining to a particular time period—5-year time periods as indicated by the plotted points in the two figures. Each trend is piecewise-linear, based on three TFR estimates pertaining to three consecutive 5-year time periods before each survey. The two trends overlap for part of the estimation period. Of particular interest are the second 5-year period before NFHS-2 and the first 5-year period before NFHS-1, which approximately coincide. The TFR estimates for these two periods should agree closely, but they do not. Instead, the TFR estimate from NFHS-2 is considerably higher than the corresponding TFR estimate from NFHS-1. The reason for this discrepancy is that, in the retrospective birth histories, births in the first 5-year period before each survey tend to be displaced into the second 5-year period before each survey, leading to a TFR estimate that is too low for the first 5-year period and too high for the second 5-year period. Such

displacement of births is common in retrospective fertility surveys in many developing countries, particularly in the large countries of South Asia where many respondents have only approximate knowledge of their own birth dates and the birth dates of their children.

- Figures 1 and 2 about here -

Figure 2 shows a similar graph, with the sex ratio at birth (SRB) in place of the TFR. Again, one expects the SRB for the second 5-year period before NFHS-2 to agree closely with the SRB for the first 5-year period before NFHS-1, but again it does not. Instead, the SRB estimate from NFHS-2 substantially exceeds the corresponding SRB estimate from NFHS-1. Evidently male births tend to be displaced backward in time to a greater extent than female births. Why this should be so is not clear, although the strong preference for sons in Punjab undoubtedly has something to do with it.

Bias from displacement is minimized if one aggregates births over a 15-year period before each survey, because then a higher proportion of the displacements that occur are contained within the time period and do not affect appreciably the accuracy of the estimates. For that reason, we aggregate births over 15-year periods rather than 5-year periods in our analysis.

The two piecewise linear trends in Figure 2 also suggest that the rise in the sex ratio at birth over time is slowing and may already have reversed. This is seen by comparing the first (earliest) and third (most recent) data points in each of the two overlapping trends. In the first trend, derived from NFHS-1, the third point is higher than the first point, suggesting an increasing trend. But in the second trend, derived from NFHS-2, the third point is lower than the first point, suggesting a slowing upward trend or perhaps even a decreasing trend, although this is not certain because the third point is downwardly biased, and because the extent of the bias may not be the same in the two surveys. Later we shall examine additional evidence that the rise in the sex ratio at birth is slowing and may have reversed.

Method of multivariate analysis

The underlying response variable in the multivariate analysis of the sex ratio at birth is binary, coded 1 if the birth is male and 0 if it is female. Logistic regression is an appropriate multivariate method when the response variable is binary. It is used here to analyze the effects of selected demographic and socioeconomic factors on the sex ratio at birth. Levels of statistical significance of logistic regression coefficients take into account one level of clustering at the level of the primary sampling unit (rural village or urban block).

Separate logistic regressions are run for NFHS-1 and NFHS-2. In each logistic regression, the units of analysis are births of order 2 or higher that occurred during the 15 years immediately preceding the survey. First births are omitted in the multivariate analysis because the sex composition of earlier children is one of the predictor variables, and there are no earlier children in the case of first births.

The form of the logistic regression is

$$\log [p/(1-p)] = a + b_1 X_1 + b_2 X_2 + \dots + b_k X_k$$

The predicted value of the SRB is specified by the odds of a male birth, $p/(1-p)$, where p denotes the predicted proportion of births that are male. When calculating predicted values of the sex ratio at birth from the logistic regression, the constant term in the fitted regression is first adjusted so that, when mean values are substituted for all the predictor variables, the regression yields a predicted value of the SRB that is identical to the observed value of the SRB. (The coefficients of the predictor variables, representing effects, are not adjusted.) In the present instance, the predictor variables in the logistic regression are all categorical variables specified in terms of dummy variables, as explained in more detail below.

More specifically, predicted values of $\log(\text{SRB})$ —equal to $\log [p/(1-p)]$ —for the various categories of a particular predictor variable are calculated by setting the dummy variable(s) representing that predictor variable to appropriate combinations of ones and zeros in the underlying logistic regression equation while controlling for the other predictors by holding them constant at their mean values in the sample of births on which the regression is based. The predicted value of SRB is then calculated by taking the anti-log of $\log(\text{SRB})$.

FINDINGS

Based on 15 years of births before each survey (including first births), we can calculate a sex ratio at birth from each survey. According to this calculation, the SRB in Punjab increased from 1.14 to 1.20 between NFHS-1 and NFHS-2. These estimates can be compared with estimates from India's Sample Registration System (SRS), as shown in Table 1, but the comparison is not precise because the SRS has not published data that would enable estimation of the SRB for the same 15-year time periods. The SRS has published estimates of the SRB for major states of India for only two time periods: 1981–90 and 1996–98 (Premi, in this volume). In the case of Punjab, these estimates are 1.13 for 1981–90 and 1.23 for 1996–98. The estimates from NFHS-1 and NFHS-2 for the 15-year period before each survey (1.14 and 1.20) are roughly consistent with these estimates from the SRS. This consistency adds to our confidence in the accuracy of the estimates and provides additional justification for basing the analysis in this chapter on data from NFHS-1 and NFHS-2.

- Table 1 about here -

Sex ratio at birth by birth order

If sex-selective abortion is occurring, one would expect it to be much less common for first-order births than for births of higher orders. Table 2 confirms this expectation indirectly by showing that the sex ratio at birth in Punjab (for the 15-year period preceding each survey) increases by birth order in both surveys. In NFHS-1 it increases from 1.09 for first births to 1.22 for births of order four and higher. In NFHS-2 it increases from 1.01 for first births to 1.36 for third births and then falls off slightly to 1.34 for births of order 4 and higher.

- Table 2 about here -

Table 2 also shows that the sex ratio of first births declined from 1.09 to 1.01 between the two overlapping 15-year time periods. The reasons for this decline are not clear, but they almost certainly have nothing to do with sex-selective abortion. Indeed, it seems likely that the implausibly low value of 1.01 in NFHS-2 is due to sampling variability. In the case of higher-order births, the SRB increased

from 1.11 to 1.23 for second births, from 1.17 to 1.36 for third births, and from 1.22 to 1.34 for births of order 4 and higher. Our interpretation of these increases is that they are accounted for by increased use of sex-selective abortion.

Table 2 also shows comparable estimates for 1990 in South Korea, known for having one of the highest levels of son preference in the world. In South Korea, the SRB increased from 1.05 in 1980 to 1.17 in 1990 and then fell to 1.10 in 1998 (Cho and Lee 2000). The peak value of 1.17 in 1990, pertaining to births of all orders, is lower than the corresponding value of 1.20 in Punjab for the period 1985–99, derived from NFHS-2. On the other hand, the SRB for third births is higher in Korea (1.91) than in Punjab (1.36), and the SRB for births of order four and higher is higher in South Korea (2.19) than in Punjab (1.34). Despite the much higher SRBs at birth orders 3 and 4+ in South Korea than in Punjab, the overall SRB is lower in Korea than in Punjab. This apparent inconsistency, which is not real, occurs because fertility is much lower in South Korea than in Punjab, so that births of orders 3 and 4+, for which the SRB is especially high, are a much smaller proportion of total births in South Korea than in Punjab. Overall, the comparison of Punjab and South Korea indicates that the SRB in Punjab is very high by international standards.

Predictor variables

The following predictor variables are included in the logistic regression analysis of factors affecting the sex ratio at birth: birth order \times number of living sons, urban-rural residence, mother's education, religion, caste/tribe, mother's media exposure, household standard of living, mother's age at childbirth, and five-year period in which the birth occurred. Birth order \times number of living sons is defined as a composite, cross-classified variable because birth order and number of living sons interact in their effects on the SRB; i.e., because the effect of number of living sons on the SRB varies by birth order. Each of these predictor variables is defined as a categorical variable, and the categories are shown in Table 3.

- Table 3 about here -

These particular predictor variables are chosen for inclusion in the regression because each of them is likely to have an effect on the sex ratio at birth, and each of them is possibly correlated with birth order \times number of living children, which is the principal variable of interest inasmuch as it has by far the largest effect on the SRB, as will be seen later.

Some of the predictor variables require further explanation. Scheduled castes and scheduled tribes are castes and tribes identified by the Government of India as socially and economically backward and in need of special protection from social injustice and exploitation. Regarding media exposure, NFHS-1 asked questions about exposure to radio (at least once a week), television (at least once a week), and cinema (at least once a month), and NFHS-2 asked not only about these three media but also about reading newspapers and magazines (at least once a week). In our analysis, low media exposure is defined as positive responses on none or one of these four types of exposure, and high media exposure is defined as positive responses on two, three, or four types.

As in the case of media exposure, the standard of living index is defined somewhat differently in NFHS-1 and NFHS-2. In both surveys, standard of living is measured by an index defined in terms of ownership of household goods. In NFHS-1, the index is calculated by adding the following scores:

house type: 4 for pucca (high-quality construction materials throughout), 2 for semi-pucca, 0 for kachha (low-quality construction materials throughout); toilet facility: 4 for own flush toilet, 2 for public or shared flush toilet or own pit toilet, 1 for shared or public pit toilet, 0 for no facility or other facility; source of lighting: 2 for electricity, 1 for kerosene, gas, or oil, 0 for other source of lighting; main fuel for cooking: 2 for electricity, liquid petroleum gas, or biogas, 1 for coal, charcoal, or kerosene, 0 for other fuel; source of drinking water: 2 for pipe, hand pump, or well in residence/yard/plot, 1 for public tap, hand pump, or well, 0 for other water source; separate room for cooking: 1 for yes, 0 for no; ownership of agricultural land: 4 for 5 acres or more, 3 for 2.0-4.9 acres, 2 for less than 2 acres or acreage not known, 0 for no agricultural land; ownership of irrigated land: 2 if household owns at least some irrigated land, 0 if no irrigated land; ownership of livestock: 2 if own livestock, 0 if not own livestock; durable goods ownership: 4 for a car or tractor, 3 each for a scooter/motorcycle or refrigerator, 2.5 for a television, 2 each for a bicycle, electric fan, radio/transistor, sewing machine, water pump, bullock cart, or thresher, 1 for a clock/watch. Index scores range from 0–10 for low SLI to 10.5–20 for medium SLI and 20.5–55 for high SLI.

In NFHS-2, the standard of living index is calculated by adding the following scores: house type: 4 for pucca, 2 for semi-pucca, 0 for kachha; toilet facility: 4 for own flush toilet, 2 for public or shared flush toilet or own pit toilet, 1 for shared or public pit toilet, 0 for no facility; source of lighting: 2 for electricity, 1 for kerosene, gas, or oil, 0 for other source of lighting; main fuel for cooking: 2 for electricity, liquid petroleum gas, or biogas, 1 for coal, charcoal, or kerosene, 0 for other fuel; source of drinking water: 2 for pipe, hand pump, or well in residence/yard/plot, 1 for public tap, hand pump, or well, 0 for other water source; separate room for cooking: 1 for yes, 0 for no; ownership of house: 2 for yes, 0 for no; ownership of agricultural land: 4 for 5 acres or more, 3 for 2.0-4.9 acres, 2 for less than 2 acres or acreage not known, 0 for no agricultural land; ownership of irrigated land: 2 if household owns at least some irrigated land, 0 for no irrigated land; ownership of livestock: 2 if own livestock, 0 if not own livestock; durable goods ownership: 4 for a car or tractor, 3 each for a moped/scooter/motorcycle, telephone, refrigerator, or color television, 2 each for a bicycle, electric fan, radio/transistor, sewing machine, black and white television, water pump, bullock cart, or thresher, 1 each for a mattress, pressure cooker, chair, cot/bed, table, or clock/watch. Index scores range from 0–14 for low SLI to 15–24 for medium SLI to 25–67 for high SLI.

The categories of the other predictor variables shown in Table 3 are self-explanatory.

Table 3 also shows, for each survey, how the sample of births (15 years of births before each survey) is distributed among the categories of each of the predictor variables included in the logistic regressions. Regarding the composite variable, birth order \times number of living sons, the distribution of third births by number of living sons just before the third birth is of particular interest, because the total fertility rate in Punjab is close to three children, which means that many women go on to have a third birth only in order to get a child of the desired gender, which is usually male. In both surveys, the category of third births and no living sons is almost twice as large as the category of third births and two living sons, implying that mothers with no living sons among the first two births are approximately twice as likely to go on to have a third birth as are mothers with two living sons. Later we shall see that results from the logistic regressions confirm this interpretation. A somewhat similar pattern is observed for births of order four and higher.

Regarding the other predictor variables, the table shows that the sample is about three-fourths rural. Three-fifths of the mothers of the births were illiterate in NFHS-1, but only half were illiterate

in NFHS-2, reflecting rapid increases in levels of education in the approximately six years between the surveys. In NFHS-2 about two-fifths of the mothers are Hindu, about half are Sikh, and 6 percent have some other religion. The proportion who are Sikh declined slightly between the two surveys, and the proportion who are Hindu or “other” increased slightly. About one-third of the mothers belong to a scheduled caste or scheduled tribe, and the proportion belonging to these groups increased somewhat between the two surveys. The proportion of mothers with high media exposure increased substantially between the two surveys to slightly more than one-third in NFHS-2. About half of the mothers live in households with a high standard of living as defined here, reflecting the fact that Punjab is one of the more prosperous states of India. Slightly less than half of births occurred to mothers age 15–24, slightly less than half at ages 25–34, and the remaining 3–4 percent at older ages. In each survey, approximately one-third of the births occurred in each of the three 5-year periods before the survey, with a slightly smaller proportion occurring in the first 5-year period before the survey, partly because of displacement of births from the first to the second 5-year period before the survey and partly because of fertility decline.

Multivariate analysis

Table 4 shows estimates of the sex ratio at birth by birth order and number of living sons for births of order 2 and higher occurring in the 15-years before NFHS-1 and NFHS-2 in Punjab. As already mentioned, first births are omitted in the multivariate analysis because the sex composition of earlier children is one of the predictor variables, and there are no earlier children in the case of first births. The overall sex ratio at birth (SRB) for births of order 2 or higher is 1.16 in NFHS-1 and 1.30 in NFHS-2. The first two columns of Table 4 show predicted values of the SRB cross-classified by birth order and number of living sons, calculated from logistic regressions in which the composite variable, birth order \times number of living sons (specified by eight dummy variables representing the nine categories of this variable), is the only predictor variable included in the regression. These predicted values of the SRB are identical (to two decimal places) to the raw values of the SRB (not shown separately) calculated directly for each category of the composite variable. Statistical significance of underlying logistic regression coefficients are indicated in Table 4 for the predicted values. In this context, statistical significance is interpreted to mean that the SRB in a given category differs significantly from the SRB in the reference category at the 5 percent level or better. The last two columns show values of the SRB based on logistic regressions that include, besides birth order \times number of living sons, all the other demographic and socioeconomic predictor variables shown earlier in Table 3.

- Table 4 about here -

None of the other demographic and socioeconomic predictor variables has any statistically significant effects on the sex ratio at birth in either survey. This is true in logistic regressions in which each of these other predictor variables is the sole predictor variable in the regression (these regressions are not shown) as well as in the logistic regressions underlying last two columns of Table 4 in which all the predictor variables are included in the regression. This finding does not necessarily mean that the other predictor variables have no effects, only that the sample size is not large enough for the estimated effects to be statistically significant. Because the other predictor variables have no statistically significant effects on the SRB, predicted values of the SRB for categories of these other variables are not shown in Table 4. Logistic regression coefficients for all the predictor variables in the regressions underlying the last two columns of Table 4 are, however, shown in Appendix Table

1.

The predicted SRBs in the last two columns, which are based on logistic regressions that include all the predictor variables, differ little from the values of the SRB shown in the first two columns, which are the same as the raw values. The close agreement occurs because the dummy variables representing the main predictor variable, birth order \times number of living sons, are mostly uncorrelated with the other predictors, and this is true in both surveys. The highest correlation (absolute value) observed is 0.23, and most of the correlations are well below 0.10. As a general rule, when two predictors are uncorrelated or correlated at a very low level, one predictor can be dropped out of the regression without appreciably affecting the coefficient of the remaining predictor variable. In the present context, this means that predicted values of the SRB derived by logistic regression with the full set of predictors are very close to the raw values of the SRB.

Comparison of the last two columns of Table 4 shows that the SRB tended to increase between the two surveys when the mother did not have any living sons and to decrease when she had all sons. More specifically, in the case of second-order births the SRB increased from 1.13 to 1.40 when the mother had no living sons prior to the birth of the index child and declined from 1.05 to 1.01 when the mother already had a son. Although the latter decline is small, the value of 1.01 in NFHS-2, which is below the expected value of 1.05, suggests that some women who already had a son and who wanted to stop at two children used sex-selective abortion to get a daughter.

The findings for third-order births add to the plausibility of this interpretation. Among third-order births to mothers with no living sons prior to the birth of the third child, the SRB increased from 1.26 to 1.72 between the two surveys, and among those with one living son the SRB increased from 1.18 to 1.38. In contrast, among those who already had two living sons and no daughters, the SRB declined from .99 to .90, indicating a rising level of sex-selective abortion in order to have a daughter before ceasing childbearing.

The pattern for births of order 4 and higher is less clear because these birth orders are grouped. Among births to mothers with no sons, the SRB increased from 1.14 to 1.44 between the two surveys. In the case of one son it declined slightly from 1.54 to 1.44, in the case of two sons it increased from .97 to 1.31, and in the case of 3+ sons it increased from 1.15 to 1.28.

Given that the total fertility rate for Punjab is about three children, the findings in Table 4 suggest that the ideal family preferred by a large proportion of women comprises two sons and one daughter. Actually it is a little less than this on average, because ideal family size tends to be less than actual family size. Answers to survey questions on ideal family size broken down by ideal number of sons and ideal number of daughters indicate that, among ever-married women of reproductive age, the ideal family consists, on average, of 1.5 sons, 0.9 daughters, and 0.2 children of either sex (gender doesn't matter) in NFHS-1, and 1.2 sons, 0.8 daughters, and 0.3 children of either sex in NFHS-2, as shown in Table 5.

- Table 5 about here -

Sex ratio of ideal number of sons to ideal number of daughters

Will the sex ratio at birth increase further in Punjab? One way to look at this is to examine in more

detail the trends in ideal numbers of sons and daughters in Table 5.

Both NFHS-1 and NFHS-2 asked questions on ideal family size. Women with living children were asked, “If you could go back to the time you did not have any children and could choose exactly the number of children to have in your whole life, how many would that be?” Women with no living children were asked, “If you could choose exactly the number of children to have in your whole life, how many would that be?” All the women were then asked, “How many of these children would you like to be boys, how many would you like to be girls, and for how many would the sex not matter?”

As shown in Table 5, the mean ideal number of children fell by 0.3 child between the two surveys, from 2.57 to 2.27 children. The sex ratio of ideal number of boys to ideal number of girls involves assigning values of .512 son and .488 daughter for an ideal child of either sex, consistent with a sex ratio at birth of 1.05. Calculated in this way, the sex ratio of ideal number of boys to ideal number of girls fell from 1.58 to 1.37 between the two surveys. The fall in this sex ratio suggests that son preference is declining. On the other hand, the values of 1.58 and 1.37 are much higher than the actual sex ratio at birth, implying that, if parents are increasingly resorting to sex-selective abortion to realize their preferences, the actual sex ratio at birth may still be increasing at the present time.

Sex ratio of wanted next births

Sex ratios based on ideal numbers of sons and daughter are not the only indicators of gender preference. Questions asked in NFHS-1 and NFHS-2 also allow calculation of the sex ratio of wanted next births among ever-married women age 15–49 who said that they wanted another child. In NFHS-1, 38 percent of ever-married women age 15–49 said that they wanted another child (another 60 percent were sterilized or infecund, and 1 percent were undecided about whether to have another child or said it was up to God), and in NFHS-2, 21 percent said they wanted another child (another 78 percent were sterilized or infecund, and 1 percent were undecided or said it was up to God). The table shows the sex ratio of next births that would occur if, hypothetically, all of the women who want another birth were to go on to have a next birth, and all of them were to use sex-selective abortion to realize the preferred sex of the next birth. Results are based on the survey question on preferred sex of next child, which was asked of women who said that they wanted another child. In cases where the woman said that she was undecided about the sex of the next child or that the sex of the next child was up to God, we assigned values of .512 son and .488 daughter for the next birth, consistent with a sex ratio at birth of 1.05.

Table 6 shows that the overall sex ratio of wanted next births fell from 3.36 in NFHS-1 to 2.45 in NFHS-2. These values are much higher than the sex ratio of ideal number of sons to ideal number of daughters. The likely reason for the discrepancy is that women with either all sons or all daughters are overrepresented among those who want another child, resulting in an upward selection bias. This bias also affects the estimates specified by birth order. The sex ratios specified by both birth order and number of living sons, however, largely control for this bias, although there is still some overrepresentation of women who want large numbers of children.

- Table 6 about here -

Table 6 shows that at birth order 2 among women with one son, a substantial preference for daughters emerged between NFHS-1 and NFHS-2. This finding suggests that women who want to

stop at two children are not only becoming proportionately more numerous but also increasingly want to have one child of each sex. At birth order 3 among women with no son, the sex ratio of wanted next births increased from 48.18 in NFHS-1 to a whopping 105.04 in NFHS-2. In contrast, at birth order 3 among women with one son, the sex ratio declined substantially between the two surveys. At birth order 3 among women with two sons, a very substantial preference for daughters is evident in both surveys, as indicated by a sex ratio of wanted next births of 0.20 in NFHS-1 and 0.27 in NFHS-2, reflecting the modal preference for a completed family size of two sons and one daughter. In most of the categories classified by birth order and number of living sons (and these tend to be the categories with the most births), the trend in the sex ratio of wanted next births is downward.

The finding that the sex ratios of wanted next births specified by birth order and number of living sons are mostly falling also suggests that the actual sex ratio of birth may soon start to fall if it is not falling already. But again, if parents are increasingly resorting to sex-selective abortion to realize their preferences, it is possible that the actual sex ratio at birth is still increasing.

SUMMARY AND CONCLUSION

According to India's Sample Registration system, the sex ratio at birth in Punjab rose from 1.13 to 1.23 male births per female birth between 1981-90 and 1996-98, indicating substantial and increasing use of sex-selective abortion. Compared with other countries, Punjab (although it is a state and not a country) has one of the highest sex ratios at birth in the world.

Multivariate analysis (logistic regression) of data from India's first and second National Family Health Surveys (NFHS-1 and NFHS-2) indicates that socioeconomic factors have no statistically significant effects on the sex ratio at birth in Punjab. Effects appear to be there, but the sample sizes for Punjab in NFHS-1 and NFHS-2 are not large enough to estimate these effects with enough statistical precision to draw firm conclusions.

By far the most important factors affecting the sex ratio at birth are birth order and number of living sons, indicating that gender preferences for sons and daughters are the major factors affecting the prevalence of sex-selective abortion in Punjab. Son preference is much stronger than daughter preference, so that female fetuses are aborted much more frequently than male fetuses. Some mothers who have sons but no daughters, however, selectively abort male fetuses in order to get a daughter. Thus there is some sex-selective abortion aimed at getting daughters. The main predictor variable indicating birth order and number of living sons is mostly uncorrelated with the other predictor variables, with the result that the sex ratios at birth predicted by logistic regression are very close to the raw values of the sex ratio at birth by birth order and number of living sons, calculated directly from the data. The lack of correlation indicates that the effects of the composite variable, birth order \times number of living sons, are mostly independent of the effects of the other predictor variables.

The prospect is that the sex ratio at birth in Punjab will soon start to decline if it is not declining already. On the one hand, son preference is declining, suggesting that the actual sex ratio at birth may also be declining. On the other hand, because the sex ratios at birth implied by stated preferences for sons and daughters are still much higher than the actual sex ratio at birth, and because the likelihood that a woman will resort to sex-selective abortion to realize her gender preferences may still be increasing, it is possible that the actual sex ratio at birth is still increasing. Our analysis of the

actual sex ratio at birth, which uses estimates of the sex ratio at birth for 15-year time periods prior to NFHS-1 and NFHS-2, suggests that the sex ratio at birth may still be increasing, but the 15-year time periods are not sufficiently fine-grained to draw firm conclusions about the trend in the sex ratio at birth over single calendar years close to the survey date of NFHS-2. More definitive conclusions about the current trend could be drawn if the Sample Registration System would publish estimates of the sex ratio at birth on an annual basis.

REFERENCES

Cho, Nam-Hoon, and Lee Sam-Sik. 2000. Low fertility and population policy development in the Republic of Korea. In KIHASA (Korea Institute for Health and Social Affairs) and UNFPA (United Nations Population Fund), eds. *Low Fertility and Policy Responses to Issues of Ageing and Welfare*. Seoul: Korea Institute for Health and Social Affairs.

IIPS (International Institute for Population Sciences) and PRC (Population Research Centre) Chandigarh. 1995. *National Family Health Survey (MCH and Family Planning): Punjab 1993*. Bombay: International Institute for Population Sciences.

IIPS (International Institute for Population Sciences) and ORC Macro. 2001. *National Family Health Survey India: Punjab*. Mumbai: International Institute for Population Sciences.

Khan, M. E., Sandya Barge, and George Philip. 1996. Abortion in India: an overview. *Social Change* 26: 208–225.

Kusum. 1993. “The use of pre-natal diagnostic techniques for sex selection: the Indian scene. *Bioethics* 7: 149–65.

Patel, Rita. n.d. The practice of sex selective abortion in India: may you be the mother of a hundred sons. Manuscript located on the Internet.

Retherford, R. D., and V. K. Mishra. 2001. *An Evaluation of Recent Estimates of Fertility Trends in India*. National Family Health Survey Subject Reports No. 19. Mumbai: International Institute for Population Sciences. Honolulu: East-West Center.

Revised July 11, 2002

Table 1: Trend in the sex ratio at birth (SRB) in Punjab, estimated from India's Sample Registration System (SRS) and first and second National Family Health Surveys (NFHS-1 and NFHS-2)

Date source and time period	Sex ratio at birth
Sample Registration System (SRS)	
1981–90	1.13
1996–98	1.23
National Family Health Surveys	
1979–93 (NFHS-1)	1.14
1984–98 (NFHS-2)	1.20

Note: SRBs estimated from the SRS are based on continuous registration of births. SRBs from the NFHS surveys are based on retrospective birth histories.

Table 2: Estimates of the sex ratio at birth (SRB) for births occurring during the 15 years before the survey, by birth order: Punjab, NFHS-1 and NFHS-2

Birth order	NFHS-1 1979–93		NFHS-2 1984–98		Comparison with South Korea 1990
	SRB	N	SRB	N	SRB
1	1.09	1,667	1.01	1,563	1.09
2	1.11	1,579	1.23	1,437	1.17
3	1.17	1,228	1.36	994	1.91
4+	1.22	1,504	1.34	941	2.19
Total	1.14	5,978	1.20	4,935	1.17

Notes: N denotes the number of births on which the SRB is based. Multiple births and births to women age 13–14 at the time of NFHS-1 are excluded from this table. NFHS-2 did not collect birth histories from women age 13–14 at the time of the survey.

Source: South Korean comparison data are from Cho and Lee (2000).

Table 3: Percent distribution of the samples on the predictor variables included in the logistic regressions, in which the units of analysis are births of order 2 and higher during the 15 years preceding each survey: Punjab, NFHS-1 and NFHS-2

Variable	NFHS-1	NFHS-2
Birth order × No. of living sons^a		
Order 2, 0 sons	19	23
Order 2, 1 son	17	20
Order 3, 0 sons	9	9
Order 3, 1 son	14	15
Order 3, 2 sons	5	5
Order 4+, 0 sons	7	6
Order 4+, 1 son	14	12
Order 4+, 2 sons	9	7
Order 4+, 3+ sons	5	3
Residence		
Urban	26	26
Rural	74	74
Mother's education		
Illiterate	61	50
Less than middle complete ^b	19	23
Middle complete or higher	20	28
Religion		
Hindu	40	42
Sikh	57	52
Other	4	6

Caste/tribe		
Scheduled caste/tribe	30	36
Other	70	64
Mother's media exposure		
Low	72	64
High	28	36
Standard of living		
Low	13	6
Medium	35	46
High	52	48
Mother's age at childbirth^c		
15-24	46	49
25-34	50	48
35-49	4	3
Five-year period in which birth occurred		
0–4 years before survey	30	31
5–9 years before survey	36	35
10–14 years before survey	35	34

Note: Multiple births and births to women who were age 13–14 at the time of NFHS-1 are excluded from this table and from the corresponding logistic regressions. NFHS-2 did not collect birth histories from women age 13–14 at the time of the survey. See text for explanation of how categories of mother's media exposure and household standard of living are defined.

^a Refers to the number of living sons that the mother had just before she delivered the birth of the specified birth order

^b Does not include illiterate

^c "Childbirth" pertains to the births that are the units of analysis in the logistic regressions. The 15–24 age group includes one woman who had her second birth at age 14.

Table 4: Predicted values of the sex ratio at birth by birth order and number of living sons for births of order 2 and higher occurring in the 15 years before each survey, based on logistic regression: Punjab, NFHS-1 and NFHS-2

Birth order Number of living sons ^a	Predicted values from logistic regression without controls ^b		Predicted values from logistic regression with controls ^c	
	NFHS-1 1979–93	NFHS-2 1984–98	NFHS-1 1979–93	NFHS-2 1984–98
Birth order 2				
0 son	1.14	1.44*	1.13	1.40*
1 son [†]	1.07	1.03	1.05	1.01
Birth order 3				
0 son	1.27	1.73*	1.26	1.72*
1 son	1.19	1.36*	1.18	1.38*
2 sons	.99	.88	.99	.90
Birth order 4+				
0 son	1.13	1.39	1.14	1.44*
1 son	1.53*	1.40*	1.54*	1.44*
2 sons	.94	1.25	.97	1.31
3+ sons	1.12	1.27	1.15	1.28

Note: The overall sex ratio at births for births of order 2 or higher is 1.16 in NFHS-1 and 1.30 in NFHS-2. Birth order × number of living sons is a composite variable (with “birth order 2, 1 son” as the reference category) in the underlying logistic regressions. Multiple births and births to women who were age 13–14 at the time of NFHS-1 are excluded from this table. NFHS-2 did not collect birth histories from women age 13–14 at the time of the survey.

[†] Indicates the reference category for the nine-category variable that specifies birth order × number of living sons in the logistic regressions

* Indicates that the coefficient of the underlying dummy variable in the logistic regression is statistically significant at the 5 percent level or higher ($p \# .05$)

^a Refers to the number of living sons just before the mother delivered the birth of the specified birth order

^b Predictor variables are eight dummy variables representing the nine categories of the composite variable indicating birth order cross-classified by number of living sons (as shown in the row labels of the table), with birth order 2 and one living son as the reference category. Predicted values of the SRB in these two columns are the same (to two decimal places) as the raw values of the SRB,

calculated directly for each category of the composite variable.

^c Predictor variables include, in addition to birth order cross-classified by number of living sons, the following control variables: urban/rural residence, mother's education, mother's religion, mother's caste/tribe membership, mother's exposure to mass media, household standard of living, mother's age at the time of the birth, and 5-year period before the survey in which the birth occurred. In the last two columns of the table, these variables are controlled by holding them constant at their mean values (specific to each survey) in the underlying logistic regressions used to calculate predicted values of the sex ratio at birth. See text for how these variables (none of which has a statistically significant effect on the sex ratio at birth in either survey) are specified. The underlying logistic regression coefficients of all the predictor variables in the full model are given in Appendix Table 1.

Table 5: Mean ideal number of children, mean ideal numbers of boys and girls, and sex ratio: Punjab, NFHS-1 and NFHS-2

Indicator	NFHS-1 1993	NFHS-2 1998–99
Mean ideal number of children	2.57	2.27
Mean ideal number of boys	1.48	1.16
Mean ideal number of girls	.91	.82
Mean ideal number of either sex	.17	.29
Sex ratio ^a	1.58	1.37

^a Ideal number of boys divided by ideal number of girls. When calculating the sex ratio, we assigned values of .512 son and .488 daughter for an ideal child of either sex, consistent with a sex ratio at birth of 1.05.

Table 6: Among women who want another child, the hypothetical sex ratio of next births based on the preferred gender of the next birth, by birth order and number of living sons, under the hypothetical assumption that all of these women go on to have a next birth and use sex-selective abortion to realize their gender preference for the next birth: Punjab, NFHS-1 and NFHS-2

Birth order of wanted next birth by number of living sons	NFHS-1	NFHS-2
Birth order 1	2.07	1.59
Birth order 2	3.12	2.15
0 son	13.22	9.88
1 son	1.38	.76
Birth order 3	5.78	7.29
0 son	48.18	105.04
1 son	11.22	3.41
2 sons	.20	.27
Birth order 4+	14.63	14.72
0 son	NC	70.31
1 son	35.89	15.79
2 sons	1.01	3.14
3+ sons	.00	.34
Total	3.36	2.45

NC: Not calculated because there were no preferred female next births in this cell.

Appendix Table 1: Coefficients of predictor variables in the logistic regressions that include all the predictor variables, where the units of analysis are births of order 2 and higher during the 15 years preceding each survey: Punjab, NFHS-1 and NFHS-2

Variable	NFHS-1	NFHS-2
Birth order × No. of living sons^a		
Order 2, 0 sons	.0703	.3271*
Order 2, 1 son [†]	--	--
Order 3, 0 sons	.1765	.5370*
Order 3, 1 son	.1148	.3188*
Order 3, 2 sons	! .0654	! .1139
Order 4+, 0 sons	.0774	.3594*
Order 4+, 1 son	.3785*	.3598*
Order 4+, 2 sons	! .0859	.2627
Order 4+, 3+ sons	.0910	.2374
Residence		
Urban	.0589	! .0063
Rural [†]	--	--
Mother's education		
Illiterate [†]	--	--
Less than middle complete ^b	.0928	.0162
Middle complete or higher	.0148	.1369
Religion		
Hindu [†]	--	--
Sikh	.0894	! .0247
Other	! .2077	.1785

Caste/tribe		
Scheduled caste/tribe	.0633	.0195
Other [†]	--	--
Mother's media exposure		
Low [†]	--	--
High	! .0134	! .1118
Standard of living		
Low [†]	--	--
Medium	.0759	.0300
High	.1481	.1223
Mother's age at childbirth^c		
15-24 [†]	--	--
25-34	.0152	! .0561
35-49	! .0266	.0099
Five-year period in which birth occurred		
0-4 years before survey [†]	--	--
5-9 years before survey	.1178	.0945
10-14 years before survey	! .0838	! .0265

Note: Multiple births and births to women who were age 13-14 at the time of NFHS-1 are excluded from this table. NFHS-2 did not collect birth histories from women age 13-14 at the time of the survey. In the table, the total number of births of order 2 and higher is 4,311 in NFHS-1 and 3,372 in NFHS-2. See text for explanation of how categories of mother's media exposure and household standard of living are defined.

[†] Reference category

-- Reference category, hence no coefficient

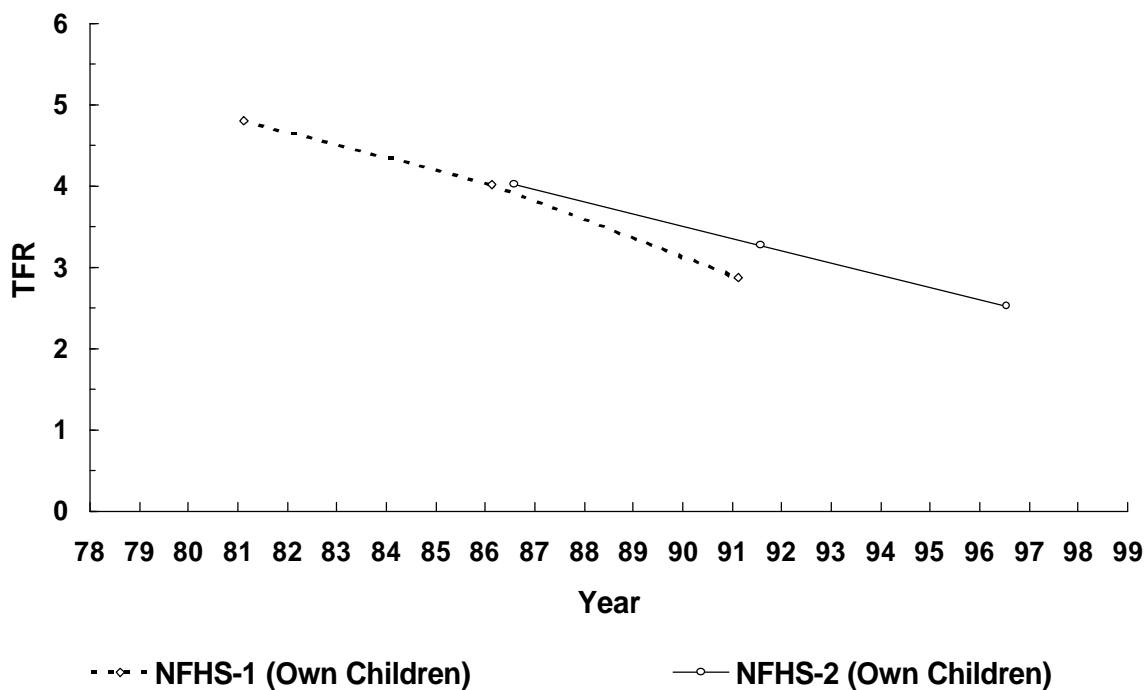
* Denotes statistical significance at the 5 percent level or higher ($p \# .05$)

^a Refers to the number of living sons that the mother had just before she delivered the birth of the specified birth order

^b Does not include illiterate

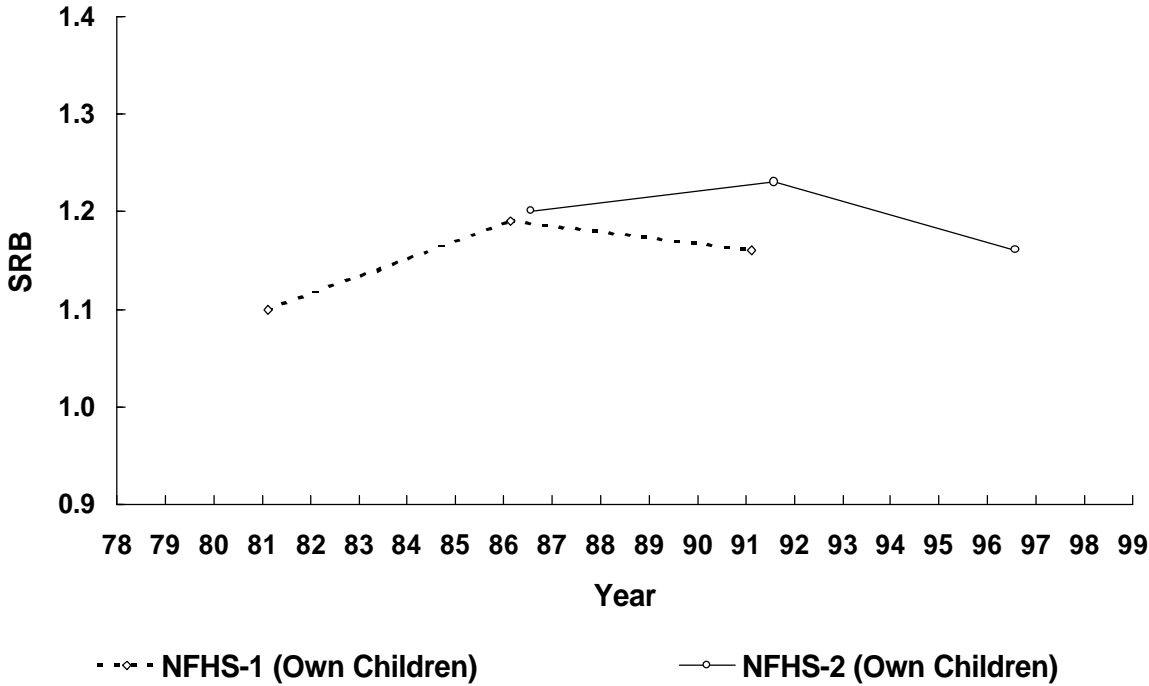
^c "Childbirth" pertains to the births that are the units of analysis in the logistic regressions. The 15-24 age group includes one woman who had her second birth at age 14.

Figure 1: Trends in the total fertility rate in Punjab, estimated from NFHS-1 and NFHS-2



Source: Underlying TFR estimates are from Retherford and Mishra (2001: Table 7).

Figure 2: Trends in the sex ratio at birth (SRB) in Punjab, estimated from NFHS-1 and NFHS-2



Source: Underlying SRB estimates are from Retherford and Mishra (2001: Table 6).