SEX RATIO AT BIRTH AND SON PREFERENCE

Nan Li, Marcus W. Feldman, and
Shripad Tuljapurkar

Paper number 0072

N.L.: Morrison Institute for Population and Resource Studies
Stanford University
Stanford, California 94305

M.W.F.: Department of Biological Sciences
Stanford University

S.T., N.L.: Mountain View Research, Inc.
2251 Grant Road, Suite A
Mountain View, California 94024

1998

In a population with a strong preference for sons, significant fertility decline is often accompanied by a considerable increase in the population sex ratio at birth (SRB, usually recorded as the ratio of the number of male births to the number of female births; Bennet 1983; Zeng et al. 1993; Park and Cho 1995), leading to a projected imbalance in future first-marriage ratios (Tuljapurkar et al. 1995; Park and Cho 1995). The purpose of this paper is to examine the relationship between son preference and SRB. We are particularly concerned about sex selection for sons, by which we mean active intervention (by sex-selective abortion, infanticide, abandonment, or other actions with similar consequences) at a particular birth to ensure that the child is male.

Why does fertility decline cause an increase in SRB? When fertility is high in any population, there are few families with no son. Also, families who do not have a son may be able to adopt one because the average family size is large and sons plentiful. Therefore, even if there is strong son preference, few couples need to employ sex selection in order to ensure a son in the family, and their effect on the SRB is small. When fertility decreases, however, no-son families become more frequent, and it is difficult for them to adopt sons. Thus, there is a large number of families that are likely to undertake sex selection at some birth order(s) to ensure the birth of a son. This behavior may have a substantial effect on the SRB of a whole population.

To clarify the interplay between son preference and fertility decline in producing an increase in SRB, we distinguish two aspects of women's reproductive situation. First, we define conditions under which women may be in a situation (the *sex-selection situation*) in which they would potentially use a sex-selection method for their next birth. We define the *sex-selection pressure* to be the probability that a woman is in this situation. Second, we define the probability that a woman who is in the sex-selection situation will actually use a sex-selection method. We call this conditional probability the *son-preference potency*. In our view, sex-selection pressure is caused by fertility decline and hence is a demographic factor, whereas the son-preference potency is an indicator of the relationship between cultural factors and individual behavior. We decompose the force that increases SRB into two components; the first is the fertility decline, which causes an increase in the sex-selection pressure, and the second is son-preference potency as determined by the cultural background of the population.

This distinction is important because fertility decline does not always cause the SRB to increase. The historical fertility decline in Western countries did not bring about an increased SRB. For example, the SRB of Sweden from 1749 to 1988 remained between 1.05 and 1.06 (Johansson and Nygren 1991). In contrast, recent fertility declines in China, South Korea, and India have led to substantial increases in the SRB. For example the SRB of China in 1989 was 1.14 (Tuljapurkar et al. 1995) and of South Korea in 1990 was 1.17 (Park and Cho 1995). Whenever fertility declines, the sex-selection pressure (as we will show) increases, but the son-preference potency may be low (as in the West) or high (as in China and South Korea) so that the overall effect on the SRB may be small (as in the West) or large (as in China and South Korea). Fertility decline has been imputed to be the basic reason for SRB increase, for example, by Hull (1990). The effect of son preference on SRB has also been pointed out, for example, by Park and Cho (1995). The purposes of this paper are to separate the effects of fertility decline and son preference on SRB and to measure these effects quantitatively.

Fertile women at different ages and with different numbers of daughters may have different chances of giving birth to further children and so face different pressures to carry out sex selection. The main aim of this paper is to quantitatively describe this situation and the effect of these women's reaction to it on the whole population.

In this paper, we provide a method for estimating these components from observed population data, which also produces an estimate of the possible maximum value of SRB. The SRB of China in 1989 was 1.14, much higher than the normal empirical level of about 1.05. Applying our method to data from China in 1989, we estimate that about 58% of women were in the sex-selection situation at some of their fertile ages, 17% of them carried out sex selection, and the possible maximum value of SRB is about 1.77.

A Quantitative Model

We now formalize our speculations in a quantitative model. We aim to describe situations in which fertility has declined, reducing the average maximum number of children that a woman may have. This causes an increase in the probability that a woman has no son in her lifetime. These, together with son preference, the desire to have at least one son, increase the likelihood of sex selection in order to have a sex-selected son.

To capture these qualitative features, we make several assumptions.

ASSUMPTION 1: Sex selection is done only by women who have no previous son and are able (or permitted) to have only one more child.

We refer to a woman under these conditions as being in the sex-selection situation. The probability of a woman's being in this situation at some time of her life will be called the sex-selection pressure. One of the aims of this paper is to provide a method for estimating this pressure from fertility data.

Son preference is manifested as the actions a woman takes to have a sex-selected son when she is in the sex-selection situation. To measure this, we define son-preference potency, the probability that a woman has a sex-selected son, given that she is in the sex-selection situation. This probability is used below to estimate the number of sex-selected sons and hence the SRB.

We base Assumption 1 on several pieces of evidence: (a) SRB changes little with birth order provided there is no sex selection (Chahnazarian 1988); (b) in recent years, a remarkable increase of SRB with birth order was observed in China (Hull 1990; Johansson and Nygren 1991; Coale 1992; Zeng et al. 1993; Park and Cho 1995) and South Korea (Zeng et al. 1993; Park and Cho 1995); (c) in China, women who have no sons are more likely to have additional children (Wang 1995); (d) in South Korea and two Chinese provinces, SRB for women who have not previously had a son is higher than that for other women (Park and Cho 1995); (e) in South Korea and two Chinese provinces, SRB of the last birth (for different mothers the order of this birth may be different) is significantly higher than that of other births (Park and Cho 1995).

These observations suggest that an increase in SRB with birth order is caused by sex selection. Observations (b) and (e) tell us that when the chance of having additional births is decreased (at higher birth orders or before the last birth), women are more likely to have sex-selected sons. Items (b) and (c) show that women who have no son are more likely to have sex-selected sons, because these women contribute a large fraction of higher-order births and the higher SRB appears in these. From (d) and (e) Park and Cho (1995) concluded: "Sex-selective abortion appears especially prevalent among families having only daughters" and "Evidently couples who have girls in the early stage of family building keep bearing children or abort female fetuses until a son arrives and then they stop."

To summarize, women who have no son and are unlikely to have two or more additional children are the most likely to have sex-selected sons. This leads to our Assumption 1.

For any woman, let the random variable C be the total number of children she will have, and let $D(i)$ be the number of daughters among the first i children she has during her life. From Assumption 1, that she is in the sex-selection situation corresponds to a random event Y, defined as

$$Y = [C = 1 \cap D(0) = 0] \cup [C = 2 \cap D(1) = 1] \cup [C = 3 \cap D(2) = 2] \cup \dots \qquad (1)$$

The sex-selection pressure, y, is defined as the probability that a woman is in the sex-selection situation. That is,

$$y = \Pr\{Y\}. \qquad (2)$$

If a woman decides to have no children, she is not in the sex-selection situation, so $y = 0$. If she decides to have only one child, she is in the sex-selection situation before she has this child, so $y = 1$. If she decides to have k children, she will be in the sex-selection situation only if her first $(k - 1)$ children are all daughters, so that y will generally be large if k is smaller, unless k equals zero.

We describe son preference in terms of the likelihood that a woman will employ some method of sex selection. We measure this by the son-preference potency, p, which is the conditional probability that a mother has a sex-selected son, given that she is in the sex-selection situation, Y:

$$p = \Pr\{\text{has a sex-selected son} \mid Y\}. \tag{3}$$

Therefore, the probability that a woman has a sex-selected son during her life is the probability that she is in the sex-selection situation at some point in her life Y and that she employs sex selection,

$$\Pr\{\text{has a sex-selected son} \cap Y\} = \Pr\{Y\} \Pr\{\text{has a sex-selected son} \mid Y\} = py. \tag{4}$$

Let T be the expected number of children a woman has during her life and s_0 be the normal (without sex selection) SRB. The total expected number of children from n women is nT, of which npy are sex-selected sons and the others involve no sex selection, so that the total expected number of sons of the n women is

$$n[py + (T - py)\, s_0 / (1 + s_0)] = n(Ts_0 + py) / (1 + s_0). \tag{5}$$

Similarly, the total expected number of daughters from the n women is $n(T - py) / (1 + s_0)$. Therefore, after sex selection, the expected SRB (i.e., the expected number of sons divided by the expected number of daughters at birth) among all births is

$$s = (Ts_0 + py) / (T - py). \tag{6}$$

As we show below, a fertility decline that corresponds to a decrease in T generally causes an increase in y. Thus, equation (6) suggests that fertility decline yields a higher SRB if p is large. If there is no son-preference potency, $p = 0$, and the SRB is normal. If the son-preference potency reaches its maximum, $p = 1$ (e.g., if techniques for sex selection become freely available and acceptable), then the SRB could be as high as

$$s_{max} = (Ts_0 + y) / (T - y). \tag{7}$$

Note that these results disentangle the effect of y (which is primarily due to the level and timing of fertility) from p (which is primarily driven by son preference expressed in some action to determine the sex of a live birth). Thus, the increase in SRB has been decomposed in terms of our two factors: fertility decline and son-preference potency. We now show how y and p may be estimated from population data.

Population Measures

Denote the birth rates of children of order i to women aged j by $f_j(i)$:

$$f_j(i) = \frac{\text{the number of } i\text{th children to mothers aged } j \text{ in a given year}}{\text{the number of women aged } j \text{ at the middle of the given year}} \,. \qquad (8)$$

These rates can be used to describe a woman's childbearing trajectory if we make Assumptions 2 and 3.

ASSUMPTION 2: The birth rates in equation (8) are constant over a long time.
ASSUMPTION 3: There is no mortality in the reproductive ages $[\alpha,\beta)$.

If Assumption 2 holds, then the birth rates $f_j(i)$ at a given time are also the probabilities of choosing a woman who has her ith child given her age is j, no matter how old she is at the given time and provided she survives from age α to j. Assumption 3 allows us to eliminate the latter proviso and assert that

$$f_j(i) = \Pr\{\text{a woman has her } i\text{th child} \mid \text{her age is } j\}. \qquad (9)$$

This is different from Pr{a woman has her ith child $\cap$ her age is j}, the probability of choosing a woman aged j who has her ith child. Obviously, under our assumptions,

$$\Pr\{\text{a woman has her } i\text{th child} \mid \text{her age is } j\} = (\beta - \alpha) \Pr\{\text{a woman has her } i\text{th child} \cap \text{her age is } j\}. \qquad (10)$$

We may write

$$f_j(i) = \Pr\{\text{has not had } i\text{th child} \mid \text{age}(j-1)\} \frac{\Pr\{[\text{has } i\text{th child} \mid \text{age } j] \cap [\text{has not had } i\text{th child} \mid \text{age } (j-1)]\}}{\Pr\{\text{has not had } i\text{th child} \mid \text{age } (j-1)\}}$$

$$= \Pr\{\text{has not had } i\text{th child} \mid \text{age}(j-1)\} \Pr\{[\text{has } i\text{th child} \mid \text{age } j] \mid [\text{has not had } i\text{th child} \mid \text{age } (j-1)]\}. \qquad (11)$$

Therefore, the ith-order total fertility rate in a given year, T_i, defined as $T_i = \Sigma_{j=\alpha}^{\beta} f_j(i)$, satisfies

$$
\begin{aligned}
T_i &= \Pr\{\text{has } i\text{th child}|\text{age } \alpha\} \\
&+ \Pr\{\text{has not had } i\text{th child}|\text{age } \alpha\} \Pr\{[\text{has } i\text{th child}|\text{age } (\alpha+1)] \mid [\text{has not had } i\text{th child}|\text{age } \alpha]\} \\
&+ \ldots \\
&+ \Pr\{\text{has not had } i\text{th child}|\text{age } (\beta-1)\} \Pr\{[\text{has } i\text{th child}|\text{age } \beta] \mid [\text{has not had } i\text{th child}|\text{age } (\beta-1)]\} \\
&= \Pr\{\text{has } i\text{th child}|\text{age is in the range } \alpha \text{ to } \beta\}. \qquad (12)
\end{aligned}
$$

That is, T_i is also the probability that a woman has an ith child during her life, or the probability that a woman has at least i children during her life.

Let R be the order of a particular child. That is, $\{R = i\}$ is the event that a woman has an ith child during her life. Then,

$$\{R = i\} \supset \{R = (i + 1)\}, \qquad (13)$$

$$T_i = \Pr\{R = i\} = \Pr\{C \geq i\},\ T_0 = 1, \qquad (14)$$

$$
\begin{aligned}
\Pr\{C = i\} &= \Pr\{R \geq i \cap C < (i+1)\} = \Pr\{R \geq i\} \Pr\{C < (i+1) \mid R \geq i\} \\
&= \Pr\{R \geq i\} (1 - \Pr\{R \geq (i+1) \mid R \geq i\}) \\
&= \Pr\{R \geq i\} - \Pr\{R \geq (i+1) \cap R \geq i\} \\
&= \Pr\{R \geq i\} - \Pr\{R \geq (i+1)\} = T_i - T_{i+1}. \qquad (15)
\end{aligned}
$$

Hence, the probability that a woman has i children during her life is $T_i - T_{i+1}$, and

$$1 = T_0 \geq T_i \geq T_{i+1} \geq 0. \qquad (16)$$

Thus, the total fertility rate in a given year, T, satisfies

$$T = \Sigma_{i=1}^{\infty} T_i = \Sigma_{i=1}^{\infty} i(T_i - T_{i+1}) = \Sigma_{i=1}^{\infty} i \Pr\{C = i\}, \qquad (17)$$

which is also the expected number of children a woman has during her life.

Finally, the probability that a woman has an $(i + 1)$th child in her lifetime, given she already has an ith child, is

$$\Pr\{R = (i + 1) \mid R = i\} = \frac{\Pr\{R = (i + 1) \cap R = i\}}{\Pr\{R = i\}} = \frac{\Pr\{R = (i + 1)\}}{\Pr\{R = i\}} = \frac{T_{i+1}}{T_i} = p_i. \qquad (18)$$

The Effect of Son Preference on the Chance of Having the Next Child

The number p_i, defined above, is a parity-progression ratio and is influenced both by constraints of total fertility rate and by the sex composition of previous children. In the conditions we are describing here, fertility constraints combine with strong son preference, so that a woman who has no son is more likely to have a subsequent child than other women (Wang 1995). Let $D(i)$ be the number of daughters among a woman's first i children. Then we expect

$$\Pr\{R = (i + 1) \mid D(i) = i\} = a_i\, p_i\,;\ (1/p_i) \geq a_i \geq 1,\ i \geq 1;\ a_0 = 1, \qquad (19)$$

$$\Pr\{R = (i + 1) \mid D(i) < i\} = b_i\, p_i\,;\ 0 \leq b_i \leq 1,\ i \geq 1;\ b_0 = 1. \qquad (20)$$

Parameters a_i and b_i describe how son preference affects the chance of having the next child; clearly, a larger a_i means stronger son preference, whereas b_i is an associated parameter that expressed in terms of a_i . We may write

$$\begin{aligned} T_{i+1} &= \Pr\{R = i + 1\} \\ &= \Pr\{D(i) = i\}\ \Pr\{R = (i + 1) \mid D(i) = i\} + \Pr\{D(i) < i\}\ \Pr\{R = (i + 1) \mid D(i) < i\} \\ &= a_i\, p_i\, \Pr\{D(i) = i\} + b_i\, p_i\, \Pr\{D(i) < i\}, \qquad (21) \end{aligned}$$

and from Assumption 1, the event $\{D(i) = i\}$ can occur only if there is no sex selection, therefore,

$$\begin{aligned} \Pr\{D(i) = i\} &= \Pr\{D(i-1) = i - 1\}\ \Pr\{i\text{th child is a daughter} \mid D(i-1) = i - 1\} \\ &= \Pr\{D(i-1) = i - 1\}\ \Pr\{[\text{a child is a daughter} \cap R = i] \mid D(i-1) = i - 1\} \\ &= [1 / (1 + s_0)]\ \Pr\{R = i \mid D(i-1) = i - 1\}\ \Pr\{D(i-1) = i - 1\} \\ &= [1 / (1 + s_0)]\ a_{i-1}\, p_{i-1}\ \Pr\{D(i-1) = i - 1\} = \ldots \\ &= T_i [1 / (1 + s_0)]^i\ \Pi_{j=0}^{i-1} a_j\,. \qquad (22) \end{aligned}$$

Further,

$$\Pr\{D(i) < i\} = \Pr\{R = i\} - \Pr\{D(i) = i\} = T_i\{1 - [1 / (1 + s_o)]^i \, \Pi_{j=0}^{i-1} a_j\}. \quad (23)$$

Then, from (21),

$$b_i = \frac{(1 + s_o)^i - \Pi_{j=0}^{i} a_j}{(1 + s_o)^i - \Pi_{j=0}^{i-1} a_j}. \quad (24)$$

No matter how strong the son preference is, it cannot make b_i negative. Therefore, if

$$a_{i0} = \min\{(1 / p_i),\ (1 + s_o)^i / \Pi_{j=0}^{i-1} a_j\}, \quad (25)$$

then

$$a_{i0} \geq a_i \geq 1. \quad (26)$$

Sex-Selection Pressure and Son-Preference Potency

From Appendix A, we have

$$[C = k + 1 \cap D(k) = k] \cap [C = j + 1 \cap D(j) = j] = \Phi,\ \ k \neq j, \quad (27)$$

where Φ is the empty set. That is, a women cannot be in situation $[C = k + 1 \cap D(k) = k]$ and also in situation $[C = j + 1 \cap D(k) = j]$, when k and j are different. It follows that

$$y = \Sigma_{k=0}^{\infty} \Pr\{C = k + 1 \cap D(k) = k\}. \quad (28)$$

Using Appendix D,

$$\Pr\{C = k + 1 \cap D(k) = k\} = (\Pi_{j=0}^{k} a_j)[T_{k+1} - (a_{k+1} + s_o b_{k+1})T_{k+2} / (1 + s_o)] / (1 + s_o)^k, \quad (29)$$

so that

$$y = \Sigma_{k=0}^{\infty} (\Pi_{j=0}^{k} a_j)[T_{k+1} - (a_{k+1} + s_o b_{k+1})T_{k+2} / (1 + s_o)] / (1 + s_o)^k. \quad (30)$$

This allows estimation of y from observed population values of T_i and a_i. Generally, it has been observed that when a population's total fertility rate decreases from a high level, T_{k+1} decreases faster than T_k. Therefore, equation (30) suggests that fertility decline causes y to increase. However, in observed population data, there is usually a value N for which T_N is small, so that it is legitimate to group all higher orders as $T_{(N+1)+}$. Then we may write the low estimate of y as

$$y_{\min} = \Sigma_{k=0}^{N-2} (\Pi_{j=0}^{k} a_j)[T_{k+1} - (a_{k+1} + s_o b_{k+1})T_{k+2} / (1 + s_o)] / (1 + s_o)^k. \quad (31)$$

Now,

$$\begin{aligned} &\Sigma_{k=N-1}^{\infty} (\Pi_{j=0}^{k} a_j)[T_{k+1} - (a_{k+1} + s_o b_{k+1})T_{k+2} / (1 + s_o)] / (1 + s_o)^k \\ &= (\Pi_{j=0}^{N-1} a_j)T_N / (1 + s_o)^{N-1} - \Sigma_{k=N}^{\infty} (\Pi_{j=0}^{k} a_j)s_o b_k T_{k+1} / (1 + s_o)^k \\ &< (\Pi_{j=0}^{N-1} a_j)T_N/(1 + s_o)^{N-1}. \end{aligned} \quad (32)$$

Hence, the high estimate of y is

$$y_{\max} = y_{\min} + (\Pi_{j=0}^{N-1} a_j)T_N/(1 + s_o)^{N-1}. \quad (33)$$

Further, p can then be calculated from equation (6) when the SRB is higher than normal (i.e., when $s > s_o$) as

$$p = \frac{T(s - s_o)}{y(1 + s)}. \quad (34)$$

Otherwise, if $s < s_o$, then p would be negative, meaning that daughters are preferred.

An Application to China

In 1989, China's SRB was 1.14, much higher than the normal expected level of about 1.05 (Johansson and Nygren 1991), while the TFR was 2.25, close to the replacement level. The relevant population data are in Table 1, where the values of T_i are from the Population Census Office under the State Council and Department of Population Statistics, State Statistical Bureau, People's Republic of China, 1993. Assuming that $s_0 = 1.05$, a_{i0} can be computed from equation (25),

TABLE 1. T_i and a_{i0} for China, 1989

i	1	2	3	4	5+
T_i	1	0.73	0.32	0.12	0.08
a_{i0}	1.39	2.25	2.67		

We have no values for a_i in 1989, so we use the average values of 1 and a_{i0} as estimates of a_i. Using these values for a_i, and b_i from equation (24), the conditional probabilities of having the next child are shown in Table 2.

TABLE 2. Conditional Probabilities of the Next Child for China, 1989

i	1	2	3
a_i	1.20	1.63	1.83
b_i	0.81	0.75	0.76
$a_i p_i$	0.87	0.72	0.70
p_i	0.72	0.45	0.38
$b_i p_i$	0.58	0.34	0.29

To evaluate our estimate of a_i in Table 2, the data in Table 3, which are from Wang (1995), can be used to produce reference values for a_i and b_i, which are in Table 4.

TABLE 3. Next-Child Probabilities from a 10% Sample of the 2 per Thousand Fertility Survey of China, 1982

Number of Children	Sex Composition of Children	Number of Mothers	Percent Having Next Child within 6 Years after Birth of Preceding Child
1	f	7205	63.4
	m	7688	55.4
2	f f	1848	61.0
	f m	4190	31.3
	m m	1981	27.6
3	f f f	458	54.3
	f f m	1465	25.6
	f m m	1136	16.8
	m m m	359	24.4

TABLE 4. Conditional Probabilities of the Next Child from the Data in Table 3

i	1	2	3
a_i	1.07	1.64	2.06
b_i	0.93	0.81	0.84
$a_i p_i$	0.63	0.61	0.53
p_i	0.59	0.37	0.26
$b_i p_i$	0.55	0.30	0.21

The constant-fertility assumption does not strictly apply to the data in Table 3. China's fertility decline started in the early 1970's, and fertility was strictly limited from 1980 to 1984. Hence, the p_i (especially p_1) in Table 4 is smaller than in Table 2. Most women whose second child was born within 6 years after the birthday of the first child (in Table 3) would have had their second child before 1980 because it was very difficult to have a second child from 1980 to 1982, and a large fraction of them had had their second child before 1970. Because these women could not anticipate the strong fertility limit that was imposed in 1980, the value of a_1 in Table 4 should be significantly smaller than the value of a_1 in 1989. The fertility decline since 1970 mainly affected the high-order births. Hence, we may infer that women in Table 3 who had had their high-order child experienced strongly limited fertility similar to that in the period after 1980, so that the values of a_2 and a_3 in Table 4 should be close to the values of a_2 and a_3 in 1989. The value of the a_1 in Table 2 is larger than that in Table 4. The values of a_2 and a_3 in Table 2 are close to those in Table 4. Therefore, we may conclude that our rough estimates of a_i in Table 2 are close to the values in 1989 and hence can be used.

From Table 1, we see that $N = 4$ for application of equations (31) and (33). Therefore, we have $0.559 < y < 0.609$, $0.162 < p < 0.177$ and $1.728 < s_{max} < 1.811$. These ranges are small. We use the mean value of each range as our estimate. Thus, a very large fraction of women (about 58%) were in the sex-selection situation at some of their fertile ages, whereas only a small fraction (about 17%) of these had a sex-selected son. This caused the SRB of the whole population to increase from 1.05 to 1.14. If all women in the sex-selection situation were to carry out sex selection at their last birth, the SRB of the whole population would reach 1.77.

China's TFR in the period of 1950–1970 was between 5 and 6, and we take 5.5 as the average value for this period. We do not have the birth-order distributions $(T_i - T_{i+1})$. If we assume that about half the women had 5 children in their life, then $(T_5 - T_6) = 0.5$; and if the other half had 6 children, then $T_6 = 0.5$. This implies that $T_i = 1$ for $i = 1,2,3,4,5$ and $T_6 = 0.5$. Applying our model to this distribution, we have $y = 0.035$. It is clear that the actual distributions may not be of this form, but if numbers of births are approximately symmetric around the mean of 5.5, then 0.035 should still be a reasonable estimate of y because an increase in y due to women having fewer children will be approximately balanced by a decrease in y from women having more children. Under this very weak sex-selection pressure (0.035), son-preference potency has only a very little effect on SRB. For example, we believe that in 1950–1970, the values of p were probably less than 0.169, which is our higher estimate of p in 1989, because it was easier to adopt sons in earlier times; and if we use $p = 0.169$ in 1950–1970, we find that SRB = 1.052, a normal level. Even if $p = 1$ in 1950–1970 (the maximum value), the SRB would equal 1.063, still a normal level.

Based on our estimates for 1989 and 1950–1970, the increase in y is likely to have caused p to become important in its effect on the SRB. We therefore conclude that changes in y are the principal force driving changes in SRB.

We note that our analysis, based on a woman's life-cycle contribution to the population, is strictly valid only if the fertility and son-preference potency remain constant, so that TFR is the total expected number of children a woman has in her life, T. In this case, y and p are the same for women at any age. If fertility and p do not change quickly over time, our model should be a good approximation.

Appendix A

First, for any k,

$(R \neq k) \subset (R \neq k+1)$,

so that for j>k,

$[C=k+1 \cap D(k)=k] \cap [C=j+1 \cap D(j)=j]$

$= [R \neq k+2 \cap R=k+1 \cap D(k)=k] \cap [R \neq j+2 \cap R=j+1 \cap D(j)=j]$

$\leq [R \neq j+1 \cap R=j+1] \cap [R=k+1 \cap D(k)=k \cap R \neq j+2 \cap D(j)=j] = \Phi$.

Appendix B

$\Pr\{C=k+1 \cap D(k)=k\} = \Pr\{R \neq k+2 \cap R=k+1 \cap D(k)=k\}$

$= \Pr\{R=k+1 \cap D(k)=k\} \Pr\{R \neq k+2 \mid (R=k+1 \cap D(k)=k)\}$

$= \Pr\{R=k+1 \cap D(k)=k\} (1 - \Pr\{R = k+2 \mid (R=k+1 \cap D(k)=k)\})$

$= \Pr\{R=k+1 \cap D(k)=k\} - \Pr\{R=k+2 \cap D(k)=k\}$

$= \Pr\{D(k)=k\} (\Pr\{R=k+1 \mid D(k)=k\} - \Pr\{R=k+2 \mid D(k)=k\})$

$= \{T_k[1 / (1+ s_o)]^k \Pi_{j=0}^{k-1} a_j\} (a_k p_k - \Pr\{R=k+2 \mid D(k)=k\})$.

Appendix C

Because

$[D(k)=k] = [R=k+1 \mid D(k)=k] + [R \neq k+1 \mid D(k)=k]$,

therefore

$\Pr\{R=k+2 \mid D(k)=k\} = \Pr\{R=k+1 \mid D(k)=k\} \Pr\{R=k+2 \mid (R=k+1 \mid D(k)=k)\}$

$= a_k p_k \Pr\{R=k+2 \mid (R=k+1 \mid D(k)=k)\}.$

Because

$[R=k+1 \mid D(k)=k] = [D(k+1)=k+1 \mid D(k)=k] + [D(k+1)=k \mid D(k)=k],$

therefore

$$\Pr\{R=k+2 \mid (R=k+1 \mid D(k)=k)\}$$

$$= \Pr\{D(k+1)=k+1 \mid D(k)=k)\} \Pr\{R=k+2 \mid (D(k+1)=k+1 \mid D(k)=k)\}$$

$$+ \Pr\{D(k+1)=k \mid D(k)=k)\} \Pr\{R=k+2 \mid (D(k+1)=k \mid D(k)=k)\}$$

$$= [1/(1+s_o)] \Pr\{R=k+2 \mid (D(k+1)=k+1 \mid D(k)=k)\}$$

$$+ [s_o/(1+s_o)] \Pr\{R=k+2 \mid (D(k+1)=k \mid D(k)=k)\}$$

$$= [1/(1+s_o)] \Pr\{R=k+2 \mid D(k+1)=k+1\}$$

$$+ [s_o/(1+s_o)] \Pr\{R=k+2 \mid D(k+1)<k+1\}$$

$$= [1/(1+s_o)]a_{k+1}p_{k+1} + [s_o/(1+s_o)]b_{k+1}p_{k+1}.$$

Hence,

$$\Pr\{R=k+2 \mid D(k)=k\} = a_k[a_{k+1} + s_o b_{k+1}]p_k p_{k+1}/(1+s_o)$$

Appendix D

From Appendixes B and C,

$$\Pr\{C=k+1 \cap D(k)=k\} = (\Pi_{j=0}^{k} a_j)[T_{k+1}-(a_{k+1}+ s_o b_{k+1}) T_{k+2}/(1+s_o)]/(1+ s_o)^k.$$

References

Bennet, N. G. 1983. *Sex Selection of Children.* Academic Press, New York.

Chahnazarian, A. 1988. " Determinants of the sex ratio at birth: Review of recent literature." *Social Biology* 35: 214–235.

Coale, A. 1992. " Missing females in the population of China." Presented at the East-West Population Institute Seminar, 11 March.

Hull, T. 1990. "Recent trends in sex ratios in China." *Population and Development Review* 16(1): 63–83.

Johansson, S., and O. Nygren. 1991. "The missing girls of China." *Population and Development Review* 17(1): 35–52.

Park, C., and N. H. Cho. 1995. "Consequences of son preference in a low fertility society: Imbalance of the sex ratio at birth in Korea." *Population and Development Review* 21(1): 59–84.

Tuljapurkar, S., N. Li, and M. W. Feldman. 1995. "High sex ratios in China's future." *Science* 267: 874–876.

Wang, Y. 1995. "The effect of son preference on China's fertility." *Population Science of China* 49: 12–15. [In Chinese.]

Zeng, Y., et al. 1993. "Causes and implications of the increase in China's reported sex ratio at birth." *Population and Development Review* 19(2): 283–302.