*The China Review*, Vol. 3, No. 2 (Fall 2003), 13–29

# *China's Missing Girls: Numerical Estimates and Effects on Population Growth*[*]

Yong Cai and William Lavely

***Abstract***

The 2000 census of China counted approximately 12.8 million fewer females in the cohorts born between 1980 and 2000 than would be expected if China had experienced normal sex ratios at birth and the gender-neutral mortality rates derived from mainly European-based model life tables. However, this estimate of the "nominally missing" contains a substantial component of females who are alive but hidden in the population. A comparison of cohorts enumerated as small children in the 1990 census with the same cohorts enumerated 10 years later in the 2000 census reveals that fewer than a third of the girls missing in the first enumeration subsequently appear in the second. This comparison

**Yong CAI** is a Ph.D. student in the Department of Sociology at the University of Washington. His research interests include social/historical demography, comparative/historical sociology, and research methodology.

**William LAVELY** is Associate Professor of International Studies and Sociology, and a research affiliate of the Center for Studies in Demography and Ecology, University of Washington.

* Earlier versions of this paper were presented at the 2002 Annual Meeting of the Population Association of America, Atlanta, Georgia, 9–11 May 2002; at the Conference on Chinese Population and Socioeconomic Studies: Utilizing the 2000/2001 Round Census Data, Hong Kong, 19–21 June 2002; and at the Workshop on Chinese Census 2000, held at the University of Washington, 22–23 August 2002. We are grateful for the valuable comments of Li Shuzhuo, Wang Feng, and Zheng Xiaoying.

informs our rough estimate that the number of truly missing girls in the cohorts born 1980 to 2000 is approximately 8.5 million. Although the direct effect of missing girls on population size is small, the long-term influence is considerable because the reproductive potential of the missing girls is also lost. We use cohort component projections to simulate the long-term effects under different assumed scenarios. Girls already missing can be expected to reduce China's future population by 3.2% in 100 years. If missing rates should continue at 2000 levels for a century, population size would be reduced by 13.6%.

The sex ratios of births and children in China, as reported in censuses and surveys, have risen steadily since the late 1970s, a trend that has attracted attention in China and abroad.[1] A preference for sons has long been a hallmark of Chinese culture, and high sex ratios in China are not unprecedented. In various periods and situations, Chinese families of high birth and low, used infanticide to affect the size and gender composition of their families. Some studies suggest that a shortage of females was a chronic feature of Late Imperial China and the Republican era.[2] Sex ratios of cohorts born in the first half of the twentieth century were abnormally high, but fell rapidly among those born in the first decades of the People's Republic. By the mid-1970s cohort sex ratios were normal. The trend reversed in the late 1970s. For cohorts born in the 1980s, sex ratios rose monotonically over the decade, a trend that continued into the 1990s (see Table 1).

In a purely technical sense, the rise must be due to some combination of four proximate causes: (1) excess female mortality in infancy or childhood, presumably due to the discriminatory behaviour of parents; (2) excess female mortality in utero, presumably the result of sex-selective abortion; (3) net out-migration of female children, presumably due to international adoption; and (4) sex-selective undercount of children in censuses and surveys. Although sex-selective abortion, excess female mortality, and sex-selective out-migration have very different social and policy implications, they have an identical demographic effect: girls are truly removed from the population. By contrast, sex-selective undercount has different demographic implications. It implies that the "missing" are present in the population but merely hidden from census officials. In seeking to understand the social or demographic effects of missing girls, for example, for marriage markets or for population growth, it is thus important to distinguish two conceptions of missing-ness, the truly missing

**Table 1: China: Sex Ratios of the Population at Age 0 and Age 0–4, 1953–2000**

| Year of census or survey | Sex ratio of the population age 0 | Sex ratio of the population age 0–4 |
|---|---|---|
| 1953 | 104.9 | 107.3 |
| 1964 | 103.8 | 106.5 |
| 1982 | 107.6 | 107.0 |
| 1990 | 111.8 | 109.8 |
| 1995 | 116.6 | 118.8 |
| 2000 | 117.8 | 120.8 |

Sources: 1953, 1964 and 1982: Guowuyuan renkou pucha bangongshi, Guojia tongjiju renkou tongji si (Population Census Office under the State Council and State Statistical Bureau), *Zhongguo1982 nian renkou pucha ziliao* (Tabulation on the 1982 Population Census of the People's Republic of China) (Beijing: Zhongguo tongji chubanshe, 1985). 1990: Guowuyuan renkou pucha bangongshi, Guojia tongjiju renkou tongji si, *Zhongguo1990 nian renkou pucha ziliao* (Tabulation on the 1990 Population Census of the People's Republic of China) (Beijing: Zhongguo tongji chubanshe, 1993). 1995: Quanguo renkou chouyang diaocha bangongshi (Office of National Population Sampling Survey*), 1995 quanguo1% renkou chouyang diaocha ziliao* (Tabulation on the 1995 One Percent Population Sampling Survey of the People's Republic of China) (Beijing: Zhongguo tongji chubanshe, 1997). 2000: Guowuyuan renkou pucha bangongshi, Guojia tongjiju renkou tongjisi, *Zhongguo 2000 nian renkou pucha ziliao* (Tabulation on the 2000 Population Census of the People's Republic of China) (Beijing: Zhongguo tongji chubanshe, 2002).

and the nominally missing. The truly missing refers to females who have been removed from the population, before or after birth, by some sex-selective mechanism. The nominally missing includes the truly missing as well as girls who are present in the population but who have been omitted from census counts. In this paper, we refer to the truly missing as "missing girls" without specification, and refer to the nominally missing using the full term.

The proportion of the nominally missing that is truly missing has been the subject of debate. In their review of the missing female phenomenon in the five decades prior to the 1990 census, Coale and Banister[3] concluded that nearly all of the missing were truly missing, the result of excess early female mortality or of sex selective abortion. Banister's[4] comparison of 1982 and 1990 census enumerations did not detect sex-selective under-reporting in the 1982 census.[5] Zeng et al.'s[6] comparison of data from the 1987 sample census and the 1988 two per thousand fertility survey with the

1990 census indicated selective underreporting in those surveys. Using data from the 1988 two per thousand fertility survey data, Johansson and Nygren[7] also concluded that sex-selective underreporting arising from informal adoptions contributed to the elevation of reported sex ratios at birth in the early 1980s. An interpretation of the missing girl phenomenon clearly hinges on a better understanding of the proportion of the missing that are alive but hidden in the population.

China's 2000 census provides new information about recent trends in the missing girls phenomenon and an opportunity to assess the proportion of the missing due to sex-selective underreporting. In the following we (1) estimate the nominally missing in the cohorts born between 1980 and 2000; (2) estimate the proportion of the nominally missing that is truly missing from these cohorts; and (3) assess the effect of "missing girls" on population growth by population projections under different scenarios.

## Measurement of the Numbers of Missing Girls

Girls nominally missing from a population may be detected by comparing the age-sex structure of a population to a model population constructed under the assumptions of normal sex ratios at birth and a gender-neutral schedule of mortality. The sex ratio at birth in large populations (which are less subject to chance fluctuations) generally varies between 103 and 107 males per 100 females.[8] Sex ratios at birth of 105.5 or 106 approximate the norm for most populations, and have been used as standards in many studies.[9] We adopt the 106 standard here.

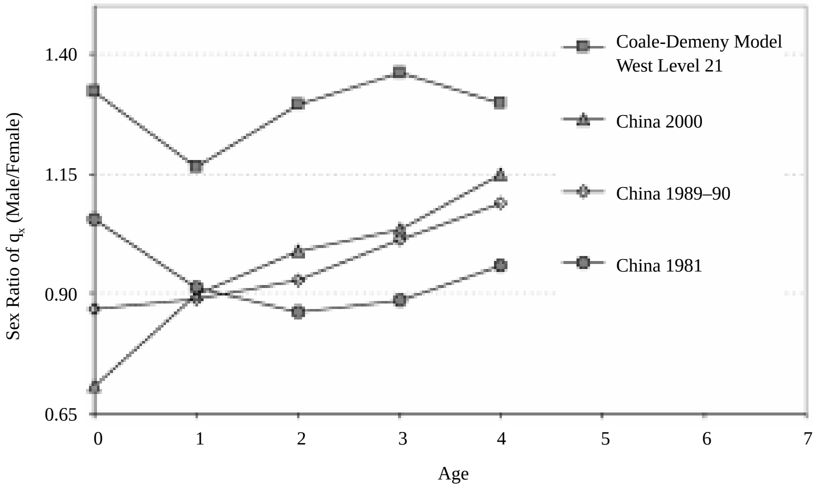
What constitutes a "gender-neutral" mortality schedule is more debatable. Males and females are not biological equals, and not equally subject to the force of mortality. For genetic and biological reasons, females are generally more resistant to disease than males, and this female advantage tends to increase with the life expectancy of a population.[10] A gender-neutral schedule of mortality would reflect biological differences but would exclude environmental influences, such as culturally specific gender bias in parental investment. Because no society has identical sex roles, it is difficult to imagine empirically based sex-specific mortality patterns that would be completely isolated from environmental factors. However, it is possible to find populations in which there is no strong preference for a child of one sex or another, and in which children receive essentially the same treatment regardless of sex.

The Coale-Demeny Model West life tables are a plausible source of gender-neutral mortality schedules.[11] The Model West tables are mainly based on the mortality experience of Western European populations with relatively minor gender bias, but also on the experience of Taiwan and Japan. An equally plausible standard are the gender-neutral mortality schedules proposed by Hill and Upchurch.[12] Hill and Upchurch in essence purge Model West of (presumably gender-biased) non-Western populations and undertake some other technical adjustments. We have adopted the Model West standard because there are solid precedents for its use (e.g., Coale and Banister 1994) and because the results are virtually indistinguishable from those based on the Hill and Upchurch standard.[13]

We considered other life tables as potential standards, including Chinese tables, but discarded them after inspection. The sex ratios of early childhood mortality derived from three recent Chinese censuses are clearly biased against females. Figure 1 portrays the sex ratio of $q_x$ (the life table probability of mortality at age x) for ages 0 through 4 in China in 1981, 1989–90, and 2000, and for Model West Level 21, corresponding to a female average life expectancy of 70.0 years, fairly close to Chinese levels in 1990 and 2000.[14] The Model West standard reflects a significant advantage to females at every age, while in the three Chinese censuses female mortality is much closer to that of male. The 1981 Chinese table is closest in pattern to Model West: males have higher mortality at age 0 but give up some of this advantage at age 1. However, the subsequent Chinese enumerations diverge sharply from this pattern. In 1989–90 infant male mortality is less than 90% that of female, and by 2000, it is less than 70%, indicating a two-decade trend of deteriorating infant mortality of females relative to males. As of the 2000 census, a female infant is almost 50% more likely to die at age 0 than a male infant. Chinese life tables clearly fail as a gender-neutral standard. United Nations model life tables for developing countries were rejected for similar reasons.[15]

Having settled on sex-unbiased standards, we construct a model population against which to compare the actual enumerated population in 2000. We assume a sex ratio at birth of 106, subject the assumed cohorts, separately by sex, to Model West Level 21 mortality,[16] and then derive the sex ratio of the model population at each age. The expected number of females at each age in 2000 can then be derived by dividing the model sex ratios into the number of males actually enumerated in the 2000 census. Assuming that the male population is fully reported, the difference between the reported and expected number of girls is the number nominally missing in each cohort.[17]

**Figure 1: Sex Ratio of $q_x$ in Model West Level 21 and in Three Chinese Census-based Life Tables**



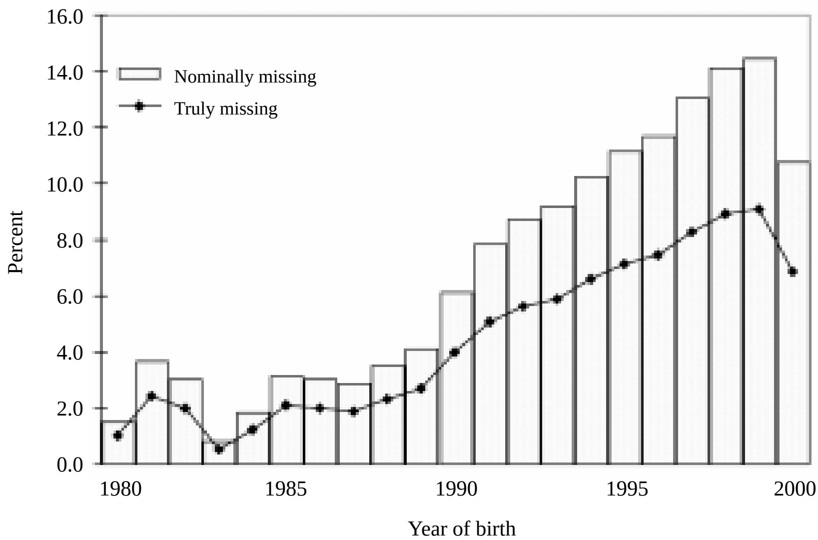

Sources: Coale-Demeny: Ansley J. Coale, Paul Demeny, and Barbara Vaughan, *Regional Model Life Tables and Stable Population* (Princeton: Princeton University Press, 1983). China 1981 and 1989–90 life tables: Huang Rongqing and Liu Yan, *Zhongguo renkou siwang shuju ji* (Mortality Data for China's Population), Zhongguo renkou xinxi yanjiu zhongxin/ lianheguo renkou jijinhui (CPIRC/UNFPA) (Beijing: Zhongguo renkou chubanshe, 1995). 2000 life tables are constructed based on mortality data given by the *Zhongguo 2000 nian renkou pucha ziliao*.

## Estimation

Applying this formulation to the cohorts born between 1980 and 2000 implies that the 2000 census of China enumerated 12.8 million fewer females aged 0 to 20 than would be expected had China had an unbiased sex ratio of births and gender-neutral pattern of mortality in those ages. This represents 6.2% of the expected female population in those cohorts. The estimation is laid out in Table 2. Column 7, containing the nominally missing as a percentage of expected females, shows that the percentage of missing was relatively low for cohorts born before 1990, increased rapidly in the early 1990s, and has remained above 10% since 1994. The pattern of nominally missing by cohort is portrayed by a histogram in Figure 2.

Among the nominally missing girls estimated in Table 2, there could be girls who are alive in the population but merely missed by enumerators.

**Figure 2: Missing Girls as Percentages of Cohorts Born 1980–2000**




An estimate of the number of such hidden girls would permit an estimate of the truly missing. Because children are less likely to remain hidden as they age (a 10-year old child is more conspicuous than an infant), small children who were missed by enumerators are likely to appear in subsequent enumerations. We use the reverse survival method to estimate the number that appears in this way. We first adjust the population enumerated in the 2000 census from the official November reference data back to midyear 2000 so as to match the reference date of the 1990 census.[18] We then project the 2000 midyear population aged 10 to 14 back 10 years to produce an "expected" population age 0 to 4 in 1990, and compare it with the population age 0 to 4 enumerated in the 1990 census.[19]

Back projections from 2000 to 1990 are based on life tables from the 1990 census. We use 1990 mortality schedules because the young cohorts under investigation experienced their highest mortality in the late 1980s and early 1990s; mortality information from the 1990 census is closest to their experience. Survival rates improved in the decade from 1990. The use of the earlier (higher mortality) schedule will produce a smaller expected population and thus increase the rate at which missing girls appear. It thus produces a conservative estimate of the truly missing.

The results of reverse survival reveal that there was significant

**Table 2: Estimated Number of Missing Girls 1980–2000 Based on Reported Sex Ratios in the 2000 Census**

| Birth cohort | Enumerated | | | Expected | | Missing | | |
|---|---|---|---|---|---|---|---|---|
| | Male | Female | Sex ratio | Sex ratio | Female | Nominal | Nominal (%) | Truly (%) |
| | (1) | (2) | (3)=(1)*100/(2) | (4) | (5)=(1)/(4) | (6)=(5)–(2) | (7)=(6)/(5) | (8)† |
| 2000 | 7,460,206 | 6,333,592 | 117.8 | 105.1 | 7,095,113 | 761,521 | 10.7 | 6.9 |
| 1999 | 6,332,424 | 5,162,823 | 122.7 | 105.0 | 6,031,747 | 868,924 | 14.4 | 9.1 |
| 1998 | 7,701,684 | 6,309,027 | 122.1 | 104.9 | 7,345,248 | 1,036,221 | 14.1 | 8.9 |
| 1997 | 7,897,235 | 6,557,101 | 120.4 | 104.8 | 7,538,939 | 981,838 | 13.0 | 8.3 |
| 1996 | 8,257,145 | 6,967,137 | 118.5 | 104.7 | 7,886,691 | 919,554 | 11.7 | 7.4 |
| 1995 | 9,157,597 | 7,775,962 | 117.8 | 104.7 | 8,748,859 | 972,897 | 11.1 | 7.1 |
| 1994 | 8,866,012 | 7,604,128 | 116.6 | 104.7 | 8,471,414 | 867,286 | 10.2 | 6.6 |
| 1993 | 9,590,414 | 8,324,342 | 115.2 | 104.6 | 9,165,440 | 841,098 | 9.2 | 5.9 |
| 1992 | 10,014,222 | 8,737,884 | 114.6 | 104.6 | 9,573,085 | 835,201 | 8.7 | 5.6 |
| 1991 | 10,674,963 | 9,407,063 | 113.5 | 104.6 | 10,207,139 | 800,076 | 7.8 | 5.1 |
| 1990 | 13,811,030 | 12,399,014 | 111.4 | 104.6 | 13,208,763 | 809,749 | 6.1 | 4.0 |
| 1989 | 13,110,848 | 12,026,831 | 109.0 | 104.5 | 12,542,125 | 515,294 | 4.1 | 2.7 |
| 1988 | 12,779,621 | 11,796,570 | 108.3 | 104.5 | 12,228,670 | 432,100 | 3.5 | 2.3 |
| 1987 | 13,619,530 | 12,663,113 | 107.6 | 104.5 | 13,036,374 | 373,261 | 2.9 | 1.9 |
| 1986 | 12,023,710 | 11,166,366 | 107.7 | 104.4 | 11,512,575 | 346,209 | 3.0 | 2.0 |
| 1985 | 10,598,457 | 9,830,866 | 107.8 | 104.4 | 10,151,429 | 320,563 | 3.2 | 2.1 |
| 1984 | 10,468,180 | 9,845,224 | 106.3 | 104.4 | 10,030,528 | 185,304 | 1.8 | 1.2 |
| 1983 | 10,297,987 | 9,789,879 | 105.2 | 104.3 | 9,871,783 | 81,904 | 0.8 | 0.6 |
| 1982 | 12,175,673 | 11,324,942 | 107.5 | 104.3 | 11,677,757 | 352,815 | 3.0 | 2.0 |
| 1981 | 10,146,557 | 9,379,850 | 108.2 | 104.2 | 9,737,252 | 357,402 | 3.7 | 2.4 |
| 1980 | 9,609,380 | 9,085,935 | 105.8 | 104.1 | 9,227,402 | 141,467 | 1.5 | 1.0 |
| Total | 214,592,875 | 192,487,649 | 111.5 | 104.5 | 205,288,334 | 12,800,685 | 6.2 | 4.1 |

Sources: 1990: *Zhongguo 1990 nian renkou pucha ziliao*. 1995: *1995 quanguo 1% renkou chouyang diaocha ziliao*. 2000: *Zhongguo 2000 nian renkou pucha ziliao*.

Note†: (8) = 100*(3*(2)/((5) + 2*(2)) – (2)/(5))

undercount of small children in the 1990 census. It implies that 12% of infants (age 0) were undercounted in the 1990 census (see Table 3). Pertinent to our concerns, girls were somewhat more likely to be undercounted. About 6.2% of boys age 0 to 4 were undercounted in the 1990 census, compared to 7.5% of girls in the same ages. This represents approximately 28.6% of the elevated sex ratio in these ages in the 1990 census.[20] A similar pattern can be observed by comparing the 1995 1% sample census to the 2000 census in the same manner (not shown). The reverse survival analysis reported here uses the same technique applied by Zeng et al. to the 1987–1990 cohorts.[21] Their analysis agrees with our comparison of 1990 with 2000 and of 1995 with 2000: sex selective underreporting is most severe at age 1 and age 2, although they yielded a larger contribution from sex-selective undercount to the elevated sex ratios for age 1 and age 2 in the 1987 sample census.

**Table 3: Under-Enumeration in the 1990 Census and Its Contribution to Elevated Reported Sex Ratio**

| Age in the 1990 census | 1990 census | | Back projected from 2000 census | | Underreporting in the 1990 census (%) | | Contribution to elevated sex ratio |
|---|---|---|---|---|---|---|---|
| | Male | Female | Male | Female | Male | Female | (%) |
| 0 | 12,254,905 | 10,965,946 | 13,906,168 | 12,584,921 | 11.9 | 12.9 | 18.9 |
| 1 | 12,304,824 | 11,027,053 | 13,160,897 | 12,085,311 | 6.5 | 8.8 | 40.8 |
| 2 | 12,672,092 | 11,508,503 | 13,187,806 | 12,185,725 | 3.9 | 5.6 | 36.2 |
| 3 | 12,676,790 | 11,617,575 | 13,197,456 | 12,245,163 | 3.9 | 5.1 | 30.3 |
| 4 | 11,140,519 | 10,270,212 | 11,636,404 | 10,783,212 | 4.3 | 4.8 | 14.9 |
| Total | 61,049,130 | 55,389,289 | 65,088,730 | 59,884,331 | 6.2 | 7.5 | 28.6 |

There are good reasons to expect that the proportion of hidden girls among the nominally missing girls has fallen since the 1980s. First, as observed in Figure 1, female infant survival deteriorated relative to male in the 1990s. Second, techniques for pre-natal sex determination (mainly, ultrasound) undoubtedly continued to spread in the 1990s. Third, international adoptions from China have risen in the 1990s, nearly all of which involve girls. If we assume that a similar relative undercount of females has occurred among infants and children in the 2000 census, it implies that less than a third of the nominally missing girls derived in Table 2 is due to girls hidden in the population.

A potential source of error in our analysis is in the undercounted population in the 2000 census, estimated in a Post Enumeration Survey at 1.81%, or 23 million people. Migrants and out-of-plan births are likely to be over-represented among the under-enumerated. Unfortunately, the age and sex breakdown of the under-enumerated has not been published. The existence of hidden girls among the under-enumerated would be, of course, perfectly compatible with our analysis, as long as the number of the hidden does not exceed our own estimate. However, this is unlikely if the pattern of undercount is similar to that in the 1990 census. As we demonstrated above, undercount in the 1990 census was only mildly skewed towards females.

Based on reverse survival estimates of the 2000 census, and reinforced by knowledge of recent trends in sex-selective mortality, sex-selective abortion and sex-selective out-migration, we conclude that at least two-thirds of the nominally missing are truly missing from China's population.

If we assume that no more than a third of the nominally missing in the

0 to 4 age range is in fact hidden (and certainly less for older ages), we can make a conservative estimate of the truly missing by applying this one-third rule across all age groups. This estimate, shown in column 8 of Table 2 and portrayed as a line superimposed on the histogram in Figure 2, implies that the number of truly missing girls in the 1980–2000 birth cohorts is about 4% or 8.5 million. This estimate is useful because, unlike hidden girls, the truly missing never go to school, never join the labour force, never marry, and never have children. By virtue of their absence they have real demographic effects.

## Effects on Population Growth

The growing dearth of females in the population has numerous demographic, social and policy implications. The potential effect on marriage markets has already attracted considerable attention, as have other potential social costs.[22] In this article we focus on a single consequence that has received little attention, the effect on population growth. Population growth rates are the central focus of China's fertility control policy. The effects of the missing girls phenomenon is thus of interest, even if it is not an intentional adjunct of policy. Any correction of the sex imbalance, to the extent that it affects population growth rates, would also have implications for other aspects of population policy.

The missing girls' influence on population growth is not primarily through the direct removal of girls from the population, but from the removal of their reproductive potential. A missing girl not only does not contribute to the population total, nor will her daughter, nor her daughter's daughter. The reproductive potential of the missing is lost to all future generations.

The direct effects of the truly missing are small relative to China's population. As we have seen, approximately 8.5 million females are missing from the cohorts born from 1980 to 2000, or only about 0.70% of China's total population. The direct influence on population growth is correspondingly small. From 1982 to 2000, China's population increased from 1.00 billion to 1.27 billion, with an annual growth rate of 12.4‰. Adding back 8.5 million missing girls would increase this rate to 12.8‰.

Although the immediate effect is small, the long-term reproductive effects are considerable. Having fewer females of reproductive age dampens reproductivity at any level of children per woman, and even if no more girls go missing, the effects on future growth persist. To assess the magnitude of these effects, we project China's population under four scenarios using the cohort component method. Our projections are used as a simulation device that permits us to compare the outcome of different missing girl scenarios. Because we do not seek to make realistic predictions about China's future population size, our scenarios make simple and mechanical assumptions about the future of China's fertility and mortality. The four scenarios are as follows:

(1) *No missing girls since 1980.* We assume that no girls went missing from 1980 on. We add the estimated missing in the 1980–2000 cohorts to the 2000 census age structure and project forward for 100 years under the assumption of normal sex ratios at birth and a gender-neutral mortality schedule.
(2) *Sudden return to normal sex ratios beginning in 2001*. We assume that the missing girls phenomenon suddenly stops in the year 2000. We take the population as enumerated in the 2000 census and project forward for 100 years assuming normal sex ratios at birth and gender neutral mortality.
(3) *Gradual return to normal by the year 2020*. We assume that sex ratios at birth and female mortality decline linearly from 2000 levels to gender-neutral levels over 20 years. We then project 80 more years from 2020 assuming normal sex ratios and gender-neutral mortality.
(4) *Rate of missing persists at 2000 level*. We assume that sex ratios at birth and excess female mortality remain constant at the levels reported in the 2000 census and project forward for 100 years.

The first scenario provides a hypothetical reference against which the other scenarios can be compared. The second scenario is unrealistic, since high sex ratios could scarcely disappear overnight, but it represents the minimum possible effects of the phenomenon, since it takes account only of missing girls that are already missing from the population and requires no assumption about the future direction of sex ratios at birth and excess female mortality. It will yield an estimate of the long-term effects of the girls missing from the 1980–2000 cohorts. The third scenario, hypothesizing a return to normal over two decades, is plausibly optimistic. It assumes that the phenomenon peaks in 2000 and declines over roughly the same two-decade period it took to rise. The fourth scenario merely assumes that the phenomenon persists at 2000 levels over the next century. A continuation of the status quo does not seem likely, but it is grounded in present reality.

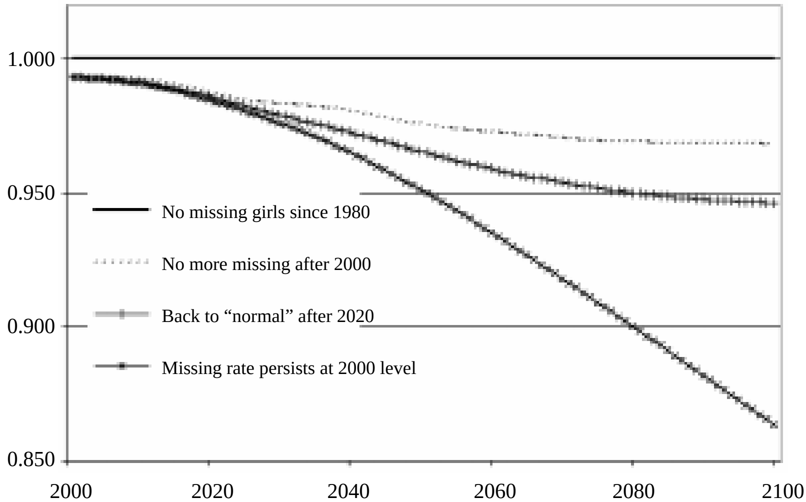
We project population using the 2000 census age-sex structure,[23] the 2000 fertility schedule,[24] and the Coale-Demeny Model West life table Level 21. In applying different assumptions of Total Fertility Rate (TFR), we inflate or deflate the fertility schedule proportionately across ages. To simulate the early excess mortality effect, the sex ratio of $q_0$ to $q_4$ in the 2000 census life table is used to adjust the model mortality schedule. We use the 106 standard for a normal sex ratio at birth. These projections assume a population closed to migration. In Table 4 and Figure 3 we present the results of the projection at TFR = 2.1, indexed to the results of scenario 1. We omit results for other TFR levels because the effects are virtually identical across a plausible range of fertility levels.

The long-term influence of missing girls on growth is considerably larger than the initial effect. Table 4 shows the results, with scenarios 2 through 4 indexed to the first scenario that assumes no missing girls since 1980. The second scenario, which assumes no more missing girls after 2000, implies a reduction of population of approximately 3.2% over 100 years, entirely due to the girls missing between 1980 and 2000. Under the third scenario, that assumes a return to gender-neutral normality by 2020, the reduction would be approximately 5.4%. Should 2000 sex ratios persist, the reduction would be 13.6% in 100 years. The trajectory of growth implied by the three scenarios differs little until 2030, after which

**Table 4: Projected Population by Scenario at TFR = 2.1, Indexed to Population Under Scenario 1 (No Missing Girls since 1980)**

| Year | No missing girls since 1980 | | No more missing after 2000 | Back to “normal” after 2020 | Missing rate persists at 2000 level |
|---|---|---|---|---|---|
| | Number ($10^9$) | Index | | | |
| 2000 | 1.258 | 1.000 | 0.994 | 0.994 | 0.994 |
| 2010 | 1.359 | 1.000 | 0.993 | 0.991 | 0.991 |
| 2020 | 1.431 | 1.000 | 0.988 | 0.986 | 0.985 |
| 2030 | 1.445 | 1.000 | 0.984 | 0.979 | 0.976 |
| 2040 | 1.448 | 1.000 | 0.981 | 0.973 | 0.965 |
| 2050 | 1.412 | 1.000 | 0.976 | 0.966 | 0.951 |
| 2060 | 1.377 | 1.000 | 0.973 | 0.959 | 0.935 |
| 2070 | 1.356 | 1.000 | 0.971 | 0.954 | 0.918 |
| 2080 | 1.333 | 1.000 | 0.969 | 0.950 | 0.900 |
| 2090 | 1.316 | 1.000 | 0.969 | 0.948 | 0.882 |
| 2100 | 1.295 | 1.000 | 0.968 | 0.946 | 0.864 |

**Figure 3: Projected Growth by Scenario Indexed to Scenario 1 (No Missing Girls since 1980), TFR = 2.1**




Note: Projection using cohort component method, based on the China 2000 fertility schedule, and Coale and Demeny Model West Level 21 mortality schedule.

there is rapid divergence. The spread between scenarios 2 and 4 represents the effect of future missing girls on future population size, should nothing change in the next 100 years. After a century this effect amounts to 10%.[25]

Our projections are limited to the national level, but the missing girls phenomenon is not homogeneous in space. The reported national sex ratio at birth in the 2000 census is 116.9. The reported sex ratio at birth of individual provinces ranges as high as 135.6. At this level, the missing girl phenomenon could produce dramatic demographic effects. Of course, provincial populations are not closed to migration. It seems entirely plausible that the local and regional effects of missing girls will occasion population movements and other adjustment responses.

## Summary and Conclusions

The 2000 census enumerated approximately 12.8 million fewer females in the cohorts born 1980–2000 than would be expected if China had experienced normal sex ratios at birth and the gender-neutral mortality rates embodied in largely European-based model life tables. This estimate of the "nominally missing" contains a substantial component of females who are alive but hidden in the population. Comparison of cohorts enumerated as small children in the 1990 census with the same cohorts enumerated 10 years later in the 2000 census reveals that fewer than a third of the missing girls appear in the subsequent enumeration. Under the assumption that girls over age 10 would be difficult to hide from official statistical systems, we assume that one-third of the nominally missing are "hidden" and two-thirds are truly missing from the population.

A conservative estimate of the number of missing girls — the truly missing — for cohorts born between 1980 and 2000 is thus approximately 8.5 million. This amounts to less than 1% of China's total population, but the long-term influence on population size will be greater because the reproductive potential of the missing is also lost. The girls already missing can be expected to decrease China's future population by 3.2% in 100 years. Should rates of missing in 2000 persist for a century, the population would be 13.6% less than it would have been had there been no missing girls. It is possible that China's infant sex ratios are at or near their peak. Beginning in 1995 the reported sex ratio in ages 0–4 was higher than that of age 0, and the reported sex ratio at age 0 has levelled off between 115 and 120, suggesting that the "missing girls" phenomenon has reached a plateau. Current trends in the missing girl phenomenon thus suggest that the effects on population growth are likely to fall between the two extreme scenarios.

Our projections are merely a simulation tool used to estimate the magnitude of population growth effects. They are not intended as forecasts. The projections assume constant fertility and mortality rates and take no account of possible feedback of population growth to fertility rates or to population policy. Higher fertility rates could easily offset the lost reproductivity of the missing girls. Assuming that China's birth planning policy persists into the future, missing girls will contribute substantially to the achievement of population targets or, alternatively, they will permit more relaxed restrictions on the fertility of future generations. However, these effects will entail as yet unmeasured social costs.

## Notes

1. For example, Amartya Sen, "More Than 100 Million Women Are Missing," *New York Review of Books*, 20 December 1990, pp. 61–66. Terence H. Hull,

"Recent Trends in Sex Ratios at Birth in China," *Population and Development Review*, No. 16 (1990), pp. 63–83; Sten Johansson and Ola Nygren, "The Missing Girls of China: A New Demographic Account," *Population and Development Review*, No. 17 (1991), pp. 35–51; Zeng Yi, Tu Ping, Gu Baochang, Xu Yi, Li Bohua, Li Yongping, "Causes and Implications of the Recent Increase in the Reported Sex Ratio at Birth in China," *Population and Development Review*, Vol. 19, No. 2 (June 1993), pp. 283–302; Ansley J. Coale and Judith Banister, "Five Decades of Missing Females in China," *Demography*, Vol. 31, No. 3 (August 1994), pp. 459–79.

2. James Lee, Wang Feng and Cameron Campbell, "Infant and Child Mortality among the Qing Nobility: Implications for Two Types of Positive Check," *Population Studies*, Vol. 48, No. 3 (November 1994), pp. 395–411. James Lee and Cameron Campbell, *Fate and Fortune in Rural China: Social Organization and Population Behavior in Liaoning, 1774–1873* (New York: Cambridge University Press, 1997). William G. Skinner, "Family Systems and Demographic Processes," in *Anthropological Demography: Toward a New Synthesis*, edited by David I. Kertzer and Tom Frick (Chicago: University of Chicago Press, 1998) pp. 53–95.
3. Coale and Banister (Note 1).
4. Judith Banister, "Implications and Quality of China's 1990 Census Data," paper presented at International Seminar on China's 1990 Population Census, Beijing, October 1992.
5. Zeng et al. (Note 1) take issue with Banister's conclusion in footnote 6.
6. Zeng et al. (Note 1), footnote 6.
7. Johansson and Nygren (Note 1).
8. Ingrid Waldron, "Factors Determining the Sex Ratio at Birth," in *Too Young to Die: Genes or Gender?* (New York: United Nations, 1998) pp. 53–63.
9. Johansson and Nygren (Note 1); Zeng et al. (Note 1); Coale and Banister (Note 1).
10. Ingrid Waldron, "Sex Differences in Infant and Early Childhood Mortality: Major Causes of Death and Possible Biological Causes," in *Too Young to Die* (Note 8). Kenneth Hill and Dawn M. Upchurch, "Gender Differences in Child Health: Evidence from the Demographic and Health Surveys," *Population and Development Review*, Vol. 21, No. 1 (March 1995), pp. 127–51.
11. Ansley J. Coale, Paul Demeny, and Barbara Vaughan, *Regional Model Life Tables and Stable Population* (Princeton: Princeton University Press, 1983). Ansley J. Coale, "Excess Female Mortality and the Balance of the Sexes in the Population: An Estimate of the Number of 'Missing Females'," *Population and Development Review*, Vol. 17, No. 3 (September 1991), pp. 517–23.
12. Kenneth Hill and Dawn M. Upchurch, "Gender Differences in Child Health: Evidence from the Demographic and Health Surveys," *Population and Development Review*, Vol. 21, No. 1 (March 1995), pp. 127–51.

13. Hill and Upchurch selected a subset of the empirical life tables underlying the Model West life tables — life tables covering the period 1820 to 1964 for England and Wales, France, the Netherlands, New Zealand, and Sweden — as the basis for examining sex differences in child mortality at different levels of overall mortality. A LOWESS curve was fitted to obtain the expected female to male mortality ratios for ${}_1q_0$, ${}_4q_1$, and ${}_5q_0$. At the same mortality level, Hill-Upchurch and Coale-Demeny ratios are virtually identical.
14. A female life expectancy ($e_0$) of 70.0 is somewhat below the unadjusted mortality levels in China of 71.9 in 1989–90 and 74.3 in 2000, but fairly close after adjusting for mortality undercount. Banister's (Note 4) estimate for 1990 $e_0$ is 66.98. Li Shuzhuo, "80 niandai Zhongguo renkou siwang shuiping he moshi de biandong fenxi" (An Analysis on the Mortality Level and Changing Pattern in 1980s China), *Renkou yanjiu* (Population Research), Vol. 3 (1994), estimates 69.99 for 1990. Li Shuzhuo and Sun Fubin, "Zhongguo dalu 2000 nian renkou pucha siwang shuiping de chubu fenxi" (The Morality Level of Mainland China in the 2000 Census, a Preliminary Analysis), paper presented at the Workshop on Chinese Census 2000, Seattle, 2002, argue that the under-reporting rate of mortality, especially infant mortality in the 2000 census, is at least as high as that in the 1990 census. A higher Model West level would imply an even more demanding sex-neutral mortality standard.
15. United Nations, *Model Life Table for Developing Countries* (New York, 1982), Population Studies, No. 77. The UN model tables are based on empirical data from developing countries, including many populations known to have serious discrimination against females. In fact, the ratio of male to female early childhood mortality in the UN tables is very close to that observed for China in 1982 and portrayed in Figure 1 — clearly not gender-neutral.
16. Although there was a modest increase in life expectancy between 1980 and 2000, we summarize the period with a single Model West level because the crucial issue for our model is the sex ratio of survival at each age. For the range of life expectancies experienced between 1980 and 2000, there are virtually no changes in these ratios in the model tables.
17. The procedure is summarized in the following formula:

$$G_x^{missing} = \frac{P_x^{male}}{SR_x^{model}} - P_x^{female}$$

in which $G_x^{missing}$ is the number of missing girls at age x, $P_x^{male}$ is the enumerated number of males at age x, $SR_x^{model}$ is the sex ratio at age x of the gender-neutral model population, and $P_x^{female}$ is the enumerated number of females at age x.
18. Adjustment was based on mortality data from the 2000 census.
19. We limit the projection to ages 10–14 in 2000 because at age 15 and above we

encounter serious under-enumeration of population due to labour migrations at those ages.

20. Had we used the Model West Level 21 life tables, the overall contribution of sex-selective underreporting to elevated sex ratios reported for age 0–4 in the 1990 would be 23.0%.
21. Zeng et al. (Note 1).
22. Shripad Tuljapurkar, Nan Li, and Marcus W. Feldman, "High Sex Ratios in China's Future," *Science*, New Series, Vol. 267, No. 5199 (10 February 1995), pp. 874–76. Elisabeth Croll, *Endangered Daughters: Discrimination and Development in Asia* (London: Routledge, 2000).
23. The projections do not include the under-enumerated 23 million in the 2000 census, revealed by the post-enumeration survey, because no age-sex breakdown is available. Using military population age-sex structure released with the 100% tabulation, 2.5 million military personnel were added back to the 2000 population. Estimated "Hidden girls" are added to the 2000 census population in all four scenarios, and estimated missing girls are added to the 2000 census population in scenario 1.
24. The fertility schedule derived from the 2000 census, although with a very low total fertility rate (1.22), has an almost identical age-pattern to Chinese fertility reported over the past two decades.
25. An alternative approach to the estimation of growth effects is to use the stable population model. For every combination of fertility and mortality schedules there is a stable equivalent population with a growth rate known as the intrinsic rate of growth. The stable population model has an advantage over the conventional projections because it is independent of the population age structure. The intrinsic rate of growth implied by the fertility and mortality schedules of the 1990 census is 0.3‰. This is the growth rate that would eventually emerge if 1990 rates persisted for several generations. Applying the model to the 2000 census (with a sex ratio at birth of 116.9) implies an intrinsic rate of growth of –24.4‰. If the sex ratio at birth were adjusted to a normal 106, the intrinsic growth rate rises to –22.4‰. This difference seems to be trivial; but after 100 years, the stable population implied by the 2000 census data would be 18% smaller than the corresponding stable population under the assumption of normal sex ratios. The effect would be even larger if sex differential mortality were taken into account. The cohort component model yields a smaller effect because of the effect of China's young age structure. By incorporating the age structure the cohort component model produces a more realistic estimate of the effect.