A Different Perspective on the Imbalance of Reported Sex Ratios at Birth in Rural China

Simona Bignami-Van Assche
University of Pennsylvania

Introduction

The reported sex ratio at birth is the number of male births per 100 female births enumerated in a census, a survey, or civil registration. Theoretically, the reported sex ratio should be equal to the "true" sex ratio (i.e., the number of males per 100 females born in a population), which is biologically stable around the value of 106 in the absence of social and behavioral interference.

Yet, by using survey and census data, many studies have identified discrepancies between true and reported sex ratios at birth in most countries, especially in the developing world. Three main proximate causes of this "imbalance" of reported sex ratios at birth have generally been identified: female infanticide, underreporting of girls, and sex-selective abortion practices.[1] Important distant causes of abnormally high sex ratios at birth are usually acknowledged to be the implementation of family planning policies targeted at fertility reduction and strong preference for sons over daughters. In fact, in countries where son preferences do exist, fertility decline has generally resulted in more skewed sex ratios at birth – a phenomenon that has been observed in Taiwan, Korea,[2] India,[3] Bangladesh, and China.[4]

Almost all studies on the imbalance of reported sex ratios at birth rely on two main implicit assumptions: 1) distant causes provide the context in which the imbalance of reported sex ratios at birth arises as a result of its proximate causes; and 2) the measurement process of the sex ratio at birth – from the actual data collection to the release of the final results – is not systematically affected by the implementation of specific family planning policies and/or by the existence of gender preferences for children. In other words, it is generally assumed that the social, cultural, and political context might create incentives for sex-selective abortion, sex-selective underreporting, and sex-selective infanticide, but would not affect the way through which births are actually measured. The context is

[1] Yet, it is worth maintaining a distinction between sex-selective abortionæwhich can indeed be a possible cause of true variations in the sex ratio at birthæand the other two, which are factors affecting only the reported sex ratio at birth. See Terence H. Hull, "Recent Trends in Sex Ratios at Birth in China," *Population and Development Review*, vol. 16, no. 1 (1990), 63-83; Terence H. Hull and Wen Xingyan, "Rising Sex Ratios at Birth in China: Evidence from the 1990 Population Census," paper presented at the International Seminar on China's 1990 Population Census, October 19-23, 1992, Beijing; and J. Ansley Coale and Judith Banister, "Five Decades of Missing Females in China," *Demography*, vol. 31, no. 3 (1994), 459-479.

[2] Chai Bin Park, "Preference for Sons, Family Size, and Sex Ratio: An Empirical Study in Korea," *Demography*, vol. 20, no. 3 (1983), 333-352.

[3] Monica Das Gupta and P.N. Mari Bhat, "Fertility Decline and Increased Manifestation of Sex Bias in India," *Population Studies*, vol. 51, no. 3 (1997), 307-315.

[4] Chai Bin Park, and Nam-Hoon Cho, "Consequences of Son Preference in a Low-Fertility Society: Imbalance of the Sex Ratio at Birth in Korea," *Population and Development Review*, vol. 21, no. 1 (1995), 59-85.

not supposed to exert any systematic effect on the measurement process of the sex ratio at birth, and its eventual imbalance is supposed to be an "objective" circumstance.

The assumption about the validity of the indicators utilized and the reliability of the methods of data collection – in the analysis of the imbalance of reported sex ratios as well as of most other topics – is indeed a common one in demographic literature. The geographic, political, social, and economic context is generally supposed to affect the quality of demographic data in a random, non-systematic way. Yet, what if a specific context does create incentives, at the individual and administrative level, to purposely misreport and mismeasure the actual number of births, abortions, or infant deaths?

The possibility that the quality of the demographic data used to measure the imbalance of reported sex ratios at birth might be affected by systematic errors caused by contextual influences on the data collection process cannot be excluded *a priori*. Each statistical system develops around complex social relations, involving the roles of respondents, interviewers, administrative personnel in charge of processing the data, and official agencies responsible for disclosing them. As the information travels from one level to another it is susceptible to the effects of the social interaction among individuals with different social interest and constraints. This might be especially true for official statistical systems, where the choice of what to measure, how to measure it, and how to interpret the results is often determined by political judgment.[5]

At least some evidence about the relationship between the imbalance of reported sex ratios at birth and sex selective abortion, sex-selective infanticide, and sex-selective underreporting might therefore be disputable, if there are unmeasured or omitted causes that both affect the outcome and are correlated with its proximate causes.[6]

These considerations are made here because the measurement of the imbalance of reported sex ratios at birth in four counties in rural northern China raises serious questions about the extent to which the social interaction between respondent and interviewer is biased by contextual influences on the data collection process.

> "THE CHINESE FAMILY PLANNING POLICY CONTEXT SEEMS TO EXERT A SIGNIFICANT INFLUENCE ON BOTH INDIVIDUAL AND ADMINISTRATIVE REPORTING. IF THIS IS ACTUALLY THE CASE, THE VALIDITY AND RELIABILITY OF STATISTICAL ESTIMATES BASED ON THESE DATA COULD BE SERIOUSLY COMPROMISED AND CAUTION SHOULD BE ADOPTED IN THEIR INTERPRETATION."

I calculated sex ratios at birth using both linked and unlinked records from retrospective pregnancy histories and vital registration data collected between 1991 and 1996 in four counties in rural northern China. The puzzling result of the analysis is that the same respondents seem to answer the same questions in a systematically different way if they report in a retrospective survey or in a vital registration system. These discrepancies pose a major problem. Is this disagreement related to the specific characteristics of the survey design in the two cases, or to different models of interviewer-respondent interaction

[5] William Alonso and Paul Starr, *The Politics of Numbers* (New York: Russell Sage Foundation, 1983), 34; and Herbert Smith, Tu Ping, Maria Giovanna Merli and Mark Hereward, "Implementation of a Demographic and Contraceptive Surveillance System in Four Counties in North China," *Population Research and Policy Review*, vol. 16, no. 4 (1997), 292-3.

[6] Michael E. Sobel, "Causal Inference in the Social and Behavioral Sciences," in *Handbook of Statistical Modeling for the Social and Behavioral Sciences* (New York: Plenum Press, 1995); and Herbert Smith, "Causation and Demography: Some Comments," paper presented at the Annual Meeting of the Population Association of America, 2001, Washington, DC.

that apply within each context? I suggest that the observed inconsistencies across demographic data collection processes might arise because, in the four counties considered, the Chinese family planning policy context seems to exert a significant influence on both individual and administrative reporting. If this is actually the case, the validity and reliability of statistical estimates based on these data could be seriously compromised and caution should be adopted in their interpretation.

Proximate and Distant Causes of the Imbalance of Reported Sex Ratios at Birth in China

The rising trend in reported sex ratios at birth in China that started during the 1980s (Table 1) has been well-documented by using data from retrospective demographic surveys, such as the 1982 One-per-Thousand Fertility Survey, the 1986 One-Percent Population Survey, and the 1988 Two-per-Thousand Fertility Survey.[7]

Previous research of excessively high sex ratios in China has identified various proximate causes of this trend: 1) excess mortality of girls;[8] 2) underreporting of girls and sex-selective abortion by means of ultrasound and other diagnostic methods;[9] 3) misreporting early deaths of girls as still births;[10] and 4) giving up daughters for adoption.[11]

From a broader perspective, the imbalance of the reported sex ratio at birth is a data collection problem, because girls are being omitted from fertility surveys.[12] This problem might arise because at the local level there are disincentives for reporting pregnancies and birth in the context of the

Table 1. Reported sex ratio at birth in China: 1960-1989

Year	Sex ratio	Year	Sex ratio	Year	Sex ratio
1960	110.3	1970	105.9	1980	107.4
1961	108.8	1971	105.2	1981	107.1
1962	106.6	1972	107.0	1982	107.2
1963	107.1	1973	106.3	1983	107.9
1964	106.6	1974	106.7	1984	108.5
1965	106.2	1975	106.4	1985	111.4
1966	112.2	1976	107.4	1986	112.3
1967	106.6	1977	106.7	1987	111.0
1968	102.5	1978	105.9	1988	n.a.
1969	104.5	1979	105.8	1989	113.8

Source: Sten Johansson and Ola Nygren, "The Missing Girls of China: A New Demographic Account," *Population and Development Review*, vol. 17, no. 1 (1991).

[7] Arthur Haupt, "The Shadow of Female Infanticide," *Intercom*, vol. 11, no. 1 (1983),13-14; Sten Johansson, "A Swedish Perspective on Sex Ratios and Other Intriguing Aspects of China's Demography," in Li Chengrui, ed., *A Census of One Billion People* (Beijing: State Statistical Bureau, 1984); J. Ansley Coale, *Rapid Population Change in China, 1952-1982* (Washington, DC: National Academy Press, 1984); and Judith Banister, *China's Changing Population* (Stanford: Stanford University Press,1987).

[8] Judith Banister, "China: Recent Mortality Levels and Trends," paper presented at the Annual Meeting of the Population Association of America, April 30-May 2, 1992, Denver, Colorado.

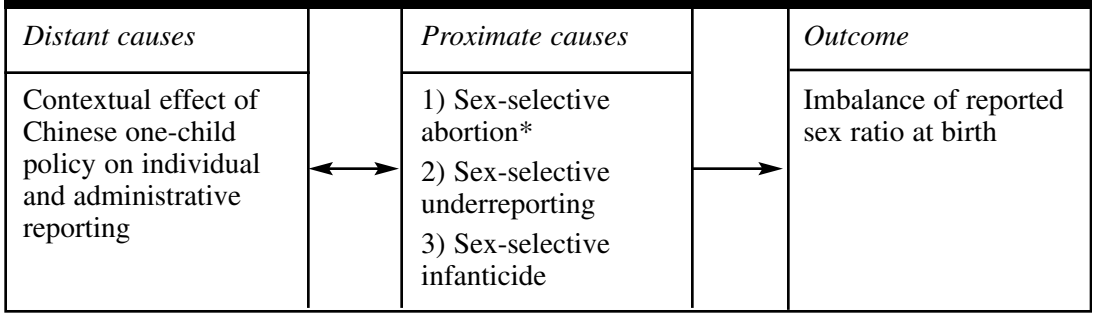
[9] Hull, 63-83; and Zeng Yi, Tu Ping, Gu Baochang, Xu Yi, Li Bohua, and Li Yongping, "Causes and Implications of the Recent Increase in the Reported Sex Ratio at Birth in China," *Population and Development Review*, vol. 19, no. 2 (1993), 283-303.

[10] Coale and Banister.

[11] Sten Johansson and Ola Nygren, "The Missing Girls of China: A new Demographic Account," *Population and Development Review*, vol. 17, no. 1 (1991), 35-51.

[12] Herbert Smith, "Nonreporting of Births or Nonreporting of Pregnancies? Some Evidence from Four Rural Counties in North China," *Demography*, vol. 31, no. 3 (1994), 484.

Figure 1. Suggested framework for the analysis of the imbalance of reported sex ratios at birth in four counties in rural northern China



*See Footnote 1.

Chinese one-child family planning program.[13] In fact, the Chinese statistical system is deeply grounded in the organization of the society over which measurement is desired, and official statistics have traditionally been maintained for administrative and planning purposes (to see that policies are implemented and goals are met. The quality of family planning data in China is therefore likely to have been affected by the establishment of a system of financial awards and penalties to measure how well cadres were able to fulfill birth quotas – the cadre responsibility system – within that policy framework.[14]

For these reasons, it is seldom disputed that the implementation of the one-child policy during the 1980s has been the main distant cause of the contemporary high sex ratio.[15] Yet, this consideration has not been developed much further in the literature. Most demographic research on the imbalance of reported sex ratios at birth in China has focused on its proximate causes (sex-selective underreporting, sex-selective abortion, and sex-selective infanticide) as separate from its main distant cause - i.e., the one-child policy. This article takes a different approach and suggests a framework for the analysis of the imbalance of reported Chinese sex ratios at birth where the interaction between its proximate and distant causes plays a major role (Figure 1). Specifically, I use data from retrospective pregnancy histories and vital registration data collected between 1991 and 1996 in four counties in rural north China to identify the role and the extent of the interaction between proximate and more distant causes of discrepancies between true and reported sex ratios, i.e., the role of the context in which these discrepancies were measured.

Contextual Effects on Individual and Administrative Reporting in Four Counties in Rural Northern China: The Available Evidence

In 1992, the State Family Planning Commission of the People's Republic of China selected four counties in rural northern China to participate in a field study designed for the Introduction of New Contraceptives in Rural China (INCRC) project.

The four counties studied – Huasheng

[13] Karen Hardee-Cleveland and Judith Banister, "Fertility Policy and Implementation in China, 1986-1988," *Population and Development Review*, vol. 14, no. 2 (1988), 245-286.

[14] John Aird, "Recent Demographic Data from China: Problems and Prospects," in *China Under the Four Modernizations*, Part I (Washington, DC: Joint Economic Committee, United States Congress, 1986), 196; and Sian Cui, "Inquiry into the Quality of Population Statistics," *Population and Economics*, vol. 2, no. 2 (1990), 31.

[15] Hull and Wen.

and Pangxie in Hebei Province, and Ciqixian and Shanshui in Shandong Province – were purposely selected to be representative of the kind of model in family planning that the government has been trying to promote.[16] Except for this common feature, the economic and geographical situations of the four counties vary.[17] Huasheng is a semi-mountainous county with strong political connections. Pangxie is a coastal enterprise zone, but the economy still relies mainly on agriculture and has not benefited from the rapid development that has characterized other Chinese counties. Ciqixian is a district in an industrial city. Shanshui is a mountain county designated by the government as underdeveloped. In addition, there is a salient demographic distinction between Huasheng and the other three counties. In fact, Pangxie, Ciqixian, and Shanshui operate under the One-Son-Two-Child policy (which allows couples to have a second child if the first child is a girl), while couples in Huasheng can have two children regardless of the sex of the first child.

The outcome of the INCRC project is composed of a pair of overlapping surveys (the Baseline and the Follow-up Survey, carried out in 1991 and 1994, respectively) and a vital statistics system (the surveillance "Card" system, running over the period 1992-96).

On the basis of these data sources, at least a few studies (described in detail below) found abnormally high sex ratios at first birth and higher parities in the four counties studied. This evidence is in contrast with the situation reported at the national level, in which the imbalance of sex ratios at higher parities is opposed to almost normal levels reported for first births.[18]

When retrospective pregnancy histories (such as those collected by the INCRC project) are used to calculate sex ratios at birth, female infanticide or other forms of differential mortality can generally be excluded as proximate causes.[19] In addition, ultrasound machines (a common device for the detection of the sex of the fetus *in utero*) were not common at the time when the surveys were carried out and the vital registration system was operative,[20] so that sex-selective abortion can also be excluded as a main proximate cause of the imbalance of reported sex ratios at birth in the study counties.[21] The possibility that elevated sex ratios in the INCRC retrospective surveys might be in part the result of misreporting early deaths of girls as stillbirths can also be ruled out.[22] Therefore, it seems that the imbalance of reported sex ratios at birth in the four study areas arise mainly as the result of a substantial tendency to omit female births from retrospective pregnancy histories. Why would this tendency arise?

By using retrospective pregnancy histories from the Baseline Survey, Smith and Tian (1993) found sex ratios at first birth higher than normal at the beginning and the end of the 1980s, but significantly lower than normal from 1987 to 1989, in all counties

[16] Ping Tu, "Contraceptive Use Patterns among Young Rural Women in Four Counties of China," paper presented at the First International Conference on Introducing New Contraceptives in Rural China, June 1993, Honolulu.

[17] Zhengchao Qian, "Progression to Second Birth in China: A Study of Four Rural Counties," *Population Studies*, vol. 51, no. 2 (1997), 221-228.

[18] As reported in Hull and Wen; Johansson and Nygren; and Ping Tu, "The Sex Ratios at Birth in China: Results from the 1990 Census," unpublished manuscript (Beijing: Institute of Population Research, Peking University, 1992).

[19] As they might in census measures of extant sex ratios that are survived back via life table methods, to obtain "implied" sex ratios at births.

[20] Ultrasound machines were only available at county hospitals. Yet, there is a government ban on the use of ultrasound for sex-selection, so that sex-selective abortions are illegal. Couples who wanted a boy could obtain a test only through personal connections, or by bribing medical personnel. See Zeng, et al.

[21] However, a study (described below) suggests that sex-selective abortion might have been more important than usually thought in determining the imbalance of reported sex ratios at birth in the four study areas. See Smith, et al.

[22] Smith, "Nonreporting of Births."

studied.[23] They suggested that the observed pattern was likely to be related to the shifts in local fertility policy during the second half of the 1980s, which allowed couples in Huasheng and Pangxie to have a second birth if the first birth was a daughter. So, they suggested that low first birth sex ratios during the period between 1987 and 1989 were an individual reporting response to policies that anticipated a greater allowance for future fertility among those couples whose first child was a girl.[24]

Analyzing the same data, Qian (1997) focused on the progression to second birth. He found "unnaturally" high sex ratios at second birth arising from the systematic tendency of having a son (rather than a daughter) of those women whose first child was a daughter. He showed that the progression to a boy as second birth is faster for women with senior high school education than for women with only primary school education, and took this evidence as an indirect proof that sex-selective abortion might have played a role greater than usually thought in the four studied counties. Educated women will be, in fact, better informed on ultrasound devices and will have more connections to people in power and easier access to sex-selective abortion. He concluded that the imbalance of reported sex ratios at birth in the four research areas might reflect differences in socioeconomic development, as well as differences in the enforcement of the family planning policy and regulations.

"EDUCATED WOMEN WILL BE, IN FACT, BETTER INFORMED ON ULTRASOUND DEVICES AND WILL HAVE MORE CONNECTIONS TO PEOPLE IN POWER AND EASIER ACCESS TO SEX-SELECTIVE ABORTION."

Using preliminary data for the INCRC Shanshui County Card System, Smith et al. (1997) also found abnormally high sex ratios for higher parities. They stressed that, under the One-Son-Two- Child policy, it would not be difficult to imagine how such a phenomenon might occur in an administrative record-keeping system, since a family whose first child is a boy, but that wants to have a second child, could only do so if that first child were administratively a girl.[25]

In all these studies, the imbalance of the sex ratio at birth in the four research counties seems therefore to emerge mainly as a data quality problem, arising within and because of the Chinese family planning policy context. As Smith et al. note detailing the efforts to implement the INCRC surveillance system:

> Often official statistics are collected and processed to reflect national goals at the possible expense of the accuracy of data . . . [In the INCRC four study counties] the data [statistical workers] collect are often used as the basis for the higher-level administrators to assess their performance, and assign awards or penalties accordingly. Thus "administrative interference" is commonplace and our intended improvement of the statistical system (for the INCRC project) occurs in a context where the data we seek to collect can have real consequences for subjects and administrators alike. [Yet] not all problems with the accuracy of data begin at the administrative end. Individual couples may not report accurately on their

[23] This pattern was not uniform across counties, with Huasheng and Ciqixian showing the greatest variability, while in Pangxie and, to a lesser extent, in Shanshui, sex ratios were overall very close to 1.06, according to Herbert Smith, and Liang Tian, "Variations in Sex Ratios for First Births in Four Counties: Where Does the Variation Begin?" paper presented at the First International Conference on Introducing New Contraceptives in Rural China, 1993, Honolulu.

[24] Smith and Tian, 14.

[25] Smith, et al, 308.

behavior, so as to achieve desired family size or sex compositions that are at variance with the mandates of official policy.

Data and Methods

The INCRC project has three essential components. The first is the Baseline Survey, carried out in the four counties mentioned above in December 1991, before the implementation of the experimental features of the INCRC project. The second is the Follow-up Survey, which was conducted in July 1994, after the experimental changes had been in place for more than two years. The third is a family planning surveillance system ("Card System"), implemented from 1992 to 1996 to provide demographic and contraceptive histories for all women of childbearing age in each township included in the project.

The INCRC Card System is a family planning surveillance system introduced in 1992, which ran throughout 1996 in the four study counties.[26] The INCRC Card System was grafted onto a pre-existing system of administrative records ("Old Card System") and was purposively designed to provide more accurate and detailed information. Like the old surveillance system, the INCRC Card System has five components: family planning surveillance cards and four registration books (marriages, births, deaths, and family planning procedures). A family planning surveillance card is filled out for all women of childbearing age in the township. The information recorded on these cards includes the number of children ever born and children surviving by sex, and complete longitudinal pregnancy and contraceptive histories.[27] Vital events such as marriage, birth, death, abortion, and other family planning operations must be recorded in the registration book within one month from their occurrence and then transferred to the surveillance card at the end of each month. In order to facilitate the input of data being collected by the Card System, the INCRC project provided each township with a microcomputer loaded with a Chinese-character database program specifically developed for this task. The INCRC Card System includes 141,152 records for the period 1992-1996.

Both the Baseline and the Follow-up Survey are complex, multi-stage samples. Yet, the sampling procedure varied somehow across the two surveys.[28] In both surveys, six townships were randomly selected within each county.[29] In the Baseline Survey, eight villages were then randomly selected within each of the 24 study townships, while in the Follow-up Survey five villages were sampled from each township, including three or more villages that had participated in the Baseline Survey. For the Baseline Survey the sampling frame was provided by administrative lists of villages; for the Follow-up Survey it was provided by the surveillance cards of the INCRC Card System. Finally, within each village the Baseline Survey

[26] Smith, et al.

[27] In the Old Card System a surveillance card was filled out only for married women. In addition, information was collected on the most recent pregnancy alone.

[28] Ping Tu, Mark Hereward, Qiu Shuhua and Ni Jiajun, "The Baseline Survey for the Project on Introducing New Contraceptives in Rural China. Part I: Preliminary Findings of the Survey" (Beijing: State Family Planning Commission, 1992); Ping Tu and Mark Hereward, "The Baseline Survey for the Project on Introducing New Contraceptives in Rural China. Part II: Tabulations of the Data" (Beijing: State Family Planning Commission, 1992); and Ping Tu and Mark Hereward, "The Follow-up Survey for the Project on Introducing New Contraceptives with Enhanced Services in Rural China" (Beijing: State Family Planning Commission, 1994).

[29] In both the Baseline and the Follow-up Survey townships were chosen as primary sampling units because for most Chinese women the initial point of contact with family planning service workers is at the township levelæusually at a family planning service station, but potentially at a township hospital (Smith et al., 291). There are between twenty and 35 townships per county, each including about 24 villages and with a total population on the order of 25,000 (implying 4,000-5,000 married women of reproductive age).

randomly selected fifty married women under the age of 35 (all eligible women if there were fifty or fewer available). For the Follow-up Survey, 150 women under the age of 38 were selected (all eligible women if there were less than 150).[30] The Follow-up Survey yielded 11,759 interviews, including 2,676 women who had been interviewed in the Baseline Survey as well.

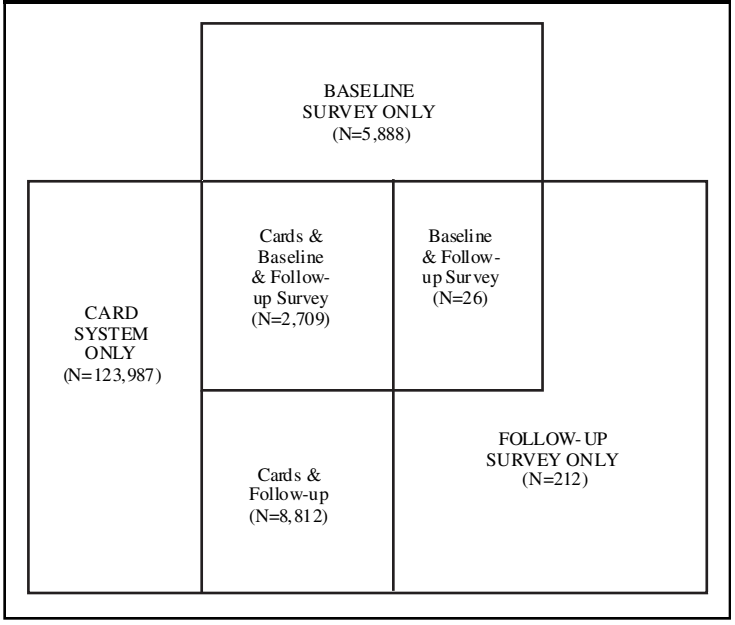
In this article, I use linked records across the Baseline Survey, the Follow-up Survey, and the INCRC surveillance system. The INCRC Card System was indeed designed to link to clinical records and to the project Baseline and Follow-up Surveys.

It has been possible to link 11,521 records (98 percent of the total number of records in the Follow-up Survey) between the Follow-up Survey and the Card System. Out of the total number of linked records between the Card System and the Follow-up Survey, 2,709 observations refer to villages that were also sampled by the Baseline Survey (Figure 2).

The matching of these records is quite good. All counties, townships, villages, and women sampled by the Follow-up Survey are included also in the Card System. Overall, misreporting of a woman's year of birth in the linked records is only ten percent (25 percent for the complete date of birth[31]) and misreporting of a woman's marital status is as low as one percent.[32]

When the matching evaluation is done separately for each county and township (Table 2), it emerges that misreporting of a woman's age is an issue in Shanshui. Despite the analysis below that includes all records linked across the Follow-up Survey and the Card System, caution will therefore be adopted in the interpretation of results for Shanshui.

Figure 2.
Linking records across the INCRC Card System and the Baseline and Follow-up Surveys.



Results and Discussion

Two sets of results are shown below. First, sex ratios at birth by parity, year, and

[30] The extension of the upper age was to ensure that all birth cohorts represented in the Baseline Survey were represented in the Follow-up Survey as well.

[31] In the Card System, the Baseline Survey and the Follow-up Survey age was measured with reference to Western definitions and the Western calendar. Yet, confusion among interviewers and respondents between the Western and Chinese calendars led to inconsistencies in the reporting of age and dates. Particularly, because of the use of the lunar calendar, months were more easily displaced than years.

[32] The Card System recorded vital events regardless of women's marital status. Yet, virtually all women in the Card System are married.

county have been calculated on the basis of all records in the Card System for the period 1992-1996. This first set of results is meant to integrate the work by Smith et al. (1997), which referred to preliminary data for Shanshui only.[33] Results for linked records across the Card System and the Follow-up Survey are then presented.

Results from the INCRC Card System, 1992-1996

A regional pattern of interest seems to emerge in the four study areas over the time period considered (Figure 3 and Table 3). On one side, Huasheng and Pangxie tend to have sex ratios for first births somewhat lower than what it would be expected by chance and

Table 2. Matching evaluation: number of bad matchings as a percentage of the overall number of matched records across the INCRC Card System and Follow-up Survey within each county and township, June 1992 to June 1994.

Huasheng							
	T1	T2	T3	T4	T5	T6	Total
Year of birth	2.83%	4.06%	2.57%	5.04%	1.86%	3.00%	3.32%
Month of birth	2.33%	6.22%	2.34%	4.87%	2.79%	4.38%	3.79%
Complete date of birth	3.99%	8.37%	3.98%	7.64%	4.42%	5.76%	5.71%
N	601	418	427	615	430	434	2925
Pangxie							
	T1	T2	T3	T4	T5	T6	Total
Year of birth	1.66%	0.96%	5.94%	36.06%	8.78%	0.72%	7.53%
Month of birth	0.74%	2.70%	11.17%	60.10%	9.92%	0.00%	11.93%
Complete date of birth	2.22%	3.28%	13.51%	67.31%	11.90%	0.72%	13.85%
N	541	519	385	421	353	277	2496
Ciqixian							
	T1	T2	T3	T4	T5	T6	Total
Year of birth	1.06%	2.14%	3.18%	3.78%	11.62%	14.65%	7.15%
Month of birth	1.85%	3.30%	4.87%	8.11%	40.57%	30.97%	17.03%
Complete date of birth	2.65%	4.85%	6.18%	9.01%	44.30%	36.42%	19.69%
N	378	515	534	555	456	704	3142
Shanshui							
	T1	T2	T3	T4	T5	T6	Total
Year of birth	26.06%	14.47%	12.73%	25.10%	13.59%	34.82%	21.54%
Month of birth	70.64%	68.64%	84.73%	77.35%	59.47%	54.66%	68.38%
Complete date of birth	75.78%	70.08%	84.73%	81.57%	61.30%	63.83%	72.14%
N	545	629	275	522	493	471	2935

Note: For each township and county, percentages are calculated as the number of bad matches over the total number of linked records for that township/county.

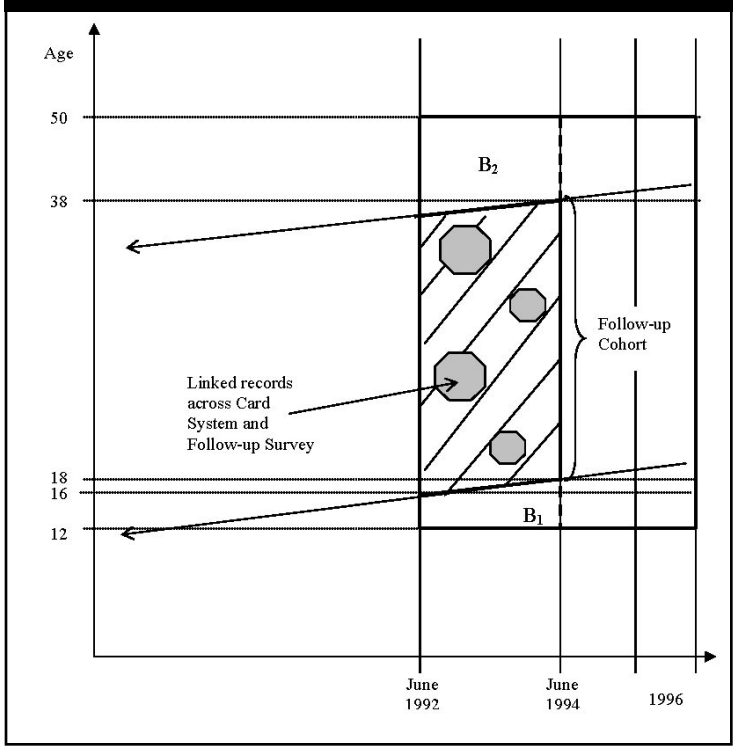
[33] Data collection was still ongoing at the time when the article was written.

implausibly low sex ratios at birth for second births or higher parities. For example, over the period from 1992 to 1996, the sex ratio at birth in Huasheng was five percent lower for parity two than for parity one, with virtually no variability across townships.[34] On the other side, Ciqixian and Shanshui tend to have sex ratios at birth for first births within the "normal" range but impossibly high sex ratios at birth for second births or higher parities. In particular, in Ciqixian the sex ratio increases by five percent at higher parities, while in Shanshui it increases by almost twenty percent. In both cases, there is again virtually no variability across townships.

In Ciqixian and Shanshui (where sex ratios at birth for second and higher order parities are substantially higher than it would be expected by chance) the results seem to be consistent with the One-Son-Two-Child policy, according to which families are allowed to have a second birth if the first birth is female. The ratio of higher births to parity one births (0.4 in Ciqixian and 0.5 in Shanshui) is, in fact, close to the ratio (0.5) implied by this policy. Interestingly enough, the data for Huasheng and Pangxie are also consistent with the One-Son-Two-Child policy (being the ratio of higher births to parity one births equal, respectively, to 0.4 and 0.6), even though the sex ratio at second births and higher parities is, as mentioned before, lower than the sex ratio at first birth. Even more interesting is the fact that Huasheng is actually the only county among those studied where families are allowed to have a second birth regardless the sex of the first birth.

The trend of reported sex ratios at birth by parity shows a second pattern of interest. There is, in fact, no constancy in the increase of the sex ratio with parity. As previously mentioned, this increase can be observed only for Ciqixian and Shanshui, whereas Huasheng and Pangxie show an opposite trend.

Figure 3.
Lexis diagram for Card System and Follow-up Survey: June 1992-1996.



The picture emerging from the analysis

[34] If the year 1992 (when the INCRC Card System was not completely operative yet) is not taken into account, this ratio becomes ten percent lower overall.

Table 3. Results from the INCRC Card System: sex ratios at birth by county, township, and birth order, 1992-1996.

	T1	T2	T3	T4	T5	T6	Total
Huasheng							
All births	1.00	1.05	0.93	0.96	1.09	0.96	1.00
N	982	423	429	1067	941	771	4613
Parity 1	0.95	1.09	0.88	0.98	1.19	1.01	1.01
N	709	368	375	875	599	490	3416
Parity 2+	1.13	0.83	1.35	0.86	0.93	0.89	0.96
N	273	55	54	192	342	281	1197
Pangxie							
All births	0.95	0.80	1.09	1.00	0.90	1.10	0.99
N	772	598	3109	1006	1992	699	8176
Parity 1	1.03	0.84	1.06	1.06	0.89	1.01	0.99
N	507	382	2037	683	1202	436	5247
Parity 2+	0.83	0.74	1.15	0.89	0.91	1.27	0.99
N	265	216	1072	323	790	263	2929
Ciqixian							
All births	1.16	1.06	1.00	1.16	1.10	1.08	1.09
N	450	1690	625	846	1188	673	5470
Parity 1	1.00	1.07	1.03	1.27	1.02	1.03	1.07
N	244	1551	419	490	682	445	3831
Parity 2+	1.40	0.96	0.94	1.02	1.22	1.19	1.13
N	206	139	206	356	506	228	1641
Shanshui							
All births	1.11	0.98	0.96	1.17	1.19	1.21	1.11
N	3112	1607	1269	3201	1912	1129	12230
Parity 1	1.05	0.92	0.89	1.08	1.09	1.18	1.04
N	1993	1139	785	2189	1239	730	8075
Parity 2+	1.23	1.14	1.07	1.40	1.40	1.25	1.27
N	1119	468	484	1012	673	399	4155

of the INCRC Card System's data is therefore puzzling. It is indeed difficult to provide a comprehensive and plausible explanation for the observed trends, which seem to contradict traditional evidence about the imbalance of sex ratios at birth, in China as elsewhere.

To search for such an explanation, reported sex ratios at birth in the four study counties were analysed across linked records between the INCRC Card System and Follow-up Survey.

Results from linked records across the INCRC Card System and Follow-up Survey: June 1992-June 1994

Two interesting results seem to emerge from the analysis of linked records across the INCRC Card System and Follow-up Survey (Table 4).

First, the agreement of linked records (Table 4, columns 3 and 4) on the number of live births, non-live births, and sex ratios at birth is excellent. Not only are the means of live and non-live births virtually identical for those respondents who were interviewed in the Follow-up Survey and recorded on the Card System, but the responses are also in strong accord across the two data collection systems (Table 5).[35] Second, there is a striking disagreement between those in the

Table 4. Means (and standard errors) for live births, sex ratio at birth, and non-live births by sample and county: Card System and Follow-up Survey, June 1992 through June 1994.

	Card System		Follow-up Survey	
	All records	Cards only	Respondents with linked records	
	(1)	(2)	(3)	(4)
Live births				
Huasheng	0.07 (.011)	0.06 (.009)	0.13 (.017)	0.15 (.010)
Pangxie	0.12 (.009)	0.11 (.010)	0.20 (.020)	0.22 (.018)
Ciqixian	0.07 (.009)	0.07 (.009)	0.15 (0.17)	0.17 (.015)
Shanshui	0.12 (.004)	0.11 (.005)	0.20 (.011)	0.24 (.016)
Sex ratio at birth				
Huasheng	1.05 (.037)	1.11 (.052)	0.95 (.093)	1.00 (.082)
Pangxie	1.01 (.054)	1.02 (.056)	0.91 (.071)	0.95 (.070)
Ciqixian	1.10 (.048)	1.15 (.057)	0.95 (.102)	0.98 (.064)
Shanshui	1.11 (.056)	1.09 (.063)	1.22 (.078)	1.20 (.065)
Non-live births				
Huasheng	0.03 (.009)	0.02 (.008)	0.06 (.011)	0.08 (.013)
Pangxie	0.04 (.004)	0.04 (.004)	0.07 (.007)	0.12 (.011)
Ciqixian	0.03 (.006)	0.02 (.006)	0.05 (.008)	0.05 (.007)
Shanshui	0.04 (.005)	0.04 (.004)	0.06 (.012)	0.09 (.008)
Sample size				
Huasheng	24188	21259	2929	
Pangxie	38017	35506	2511	
Ciqixian	26384	23242	3142	
Shanshui	46919	43980	2939	

Note: The mean numbers of live and non-live births and sex ratios at birth have been obtained by using svy commands of the statistical package STATA 6.0. The four study counties were used as four strata and, within each stratum, townships have been considered as primary sampling units (psu). In columns (1)-(3) means and standard errors are also weighted by using the sampling weights of the Card System. Finally, in column (4), means and standard errors are weighted and adjusted for the complex, multi-stage sampling design of the Follow-up Survey. Sample sizes are unweighted.

[35] The only exception seems to be Shanshui County with regard to reporting of non-live births. Yet, as previously mentioned, caution is needed in the interpretation of results for Shanshui because of the poor records matching.

Table 5. Agreement about number of live and non-live births: Card System and Follow-up Survey, June 1992 through June 1994.

	Agreement	Expected agreement	Kappa[36]
Live births			
Huasheng	96.96%	76.21%	0.8723
Pangxie	95.86%	65.87%	0.8787
Ciqixian	96.34%	73.38%	0.8625
Shanshui	94.86%	66.03%	0.8488
Non-live births			
Huasheng	97.03%	88.59%	0.7396
Pangxie	94.62%	84.91%	0.6438
Ciqixian	96.31%	90.75%	0.6011
Shanshui	92.28%	87.56%	0.3790

Card System who were sampled by the Follow-up Survey and those who were not (Table 4, columns 2 and 3). Reporting of live and non-live births doubles from the unmatched surveillance cards and the linked records with the Follow-up Survey. In addition, for all counties except Shanshui, the sex ratios for those in the Card System not interviewed during the Follow-up Survey are higher (and more "normal") than for those included in both the Card System and the Follow-up Survey.

Third, by looking at the first three columns of Table 4, it is interesting to focus on how the overall results for the Card System (column 1) can be broken down between records in the Card System only (column 2) and records linked with the Follow-up Survey (column 3). The "normal" sex ratios at birth for the Card System as a whole result to be primarily a function of the impossibly low sex ratios for respondents sampled by the Follow-up Survey as well, and slightly higher sex ratios for respondents sampled by the Card System only. Absent the linked records, sex ratios at birth for Pangxie and Ciqixian would be definitely too high (column 2), whereas if all records in the Card System are taken into account (column 1) only values for Shanshui are slightly above their "normal" value.

These results are problematic because there was no possibility for those recorded in the Card System to know that the Follow-up Survey would have sampled them as well. So, why does reporting of births in the sub-sample of records drawn from the Card System (i.e., the linked records) differ so much from that of the Card System as a whole but agree so strictly with reporting of births in the Follow-up Survey?

Careful analysis reveals that the disagreement of means of live and non-live births between linked and unlinked records within the Card System (Table 4, columns 2 and 3) is simply due to different sampling designs across the Card System and the Follow-up Survey. This is illustrated in Figure 3. On one hand, the Card system sampled all women aged twelve to fifty in

[36] Kappa is a measure of agreement that is positive when observed agreement exceeds that expected due to chance under the hypothesis of independence.

June 1992, and "followed them up" until 1996. On the other hand, the Follow-up Survey recorded retrospective information for the cohort of women aged 18 to 38 in June 1994. The present analysis has therefore been carried out with reference to the lined parallelogram and the rectangle indicated in Figure 3. The rectangle includes all demographic events recorded by the Card System from June 1992 to 1996, and the shaded parallelogram those recorded by the Follow-up Survey from June 1992 to June 1994. Since the sample for the Follow-up Survey was randomly drawn from the surveillance cards, linked records across the Card System and the Follow-up Survey refer to the dark shaded areas within the lined parallelogram. Demographic events

Table 6. Means (and standard errors) for live births, sex ratio at birth and non-live births by sample and county: linked observations between the Card System and the Follow-up Survey, June 1992 through June 1994.

	Cards	Respondents with linked records in the Cards		
	Only	Total	New Villages	Old Villages
	(1)	(2)	(3)	(4)
Live births				
Huasheng	0.06 (.009)	0.13 (.017)	0.14 (.021)	0.07 (.010)
Pangxie	0.11 (.010)	0.20 (.020)	0.23 (.025)	0.15 (.015)
Ciqixian	0.07 (.009)	0.15 (0.17)	0.16 (.022)	0.10 (.013)
Shanshui	0.11 (.005)	0.20 (.011)	0.23 (.013)	0.10 (.006)
Sex ratio at birth				
Huasheng	1.08 (.037)	0.95 (.093)	0.92 (.118)	1.18 (.158)
Pangxie	1.02 (.056)	0.91 (.071)	0.90 (.068)	0.92 (.325)
Ciqixian	1.15 (.052)	0.95 (.102)	0.95 (.138)	0.95 (.145)
Shanshui	1.10 (.064)	1.22 (.078)	1.18 (.066)	1.58 (.419)
Non-live births				
Huasheng	0.02 (.008)	0.06 (.011)	0.05 (.010)	0.07 (.018)
Pangxie	0.04 (.004)	0.07 (.007)	0.06 (.004)	0.07 (.020)
Ciqixian	0.02 (.006)	0.05 (.008)	0.04 (.010)	0.06 (.007)
Shanshui	0.04 (.004)	0.06 (.012)	0.06 (.013)	0.06 (.010)
Sample size				
Huasheng	22656	2929	2257	672
Pangxie	36907	2511	1852	659
Ciqixian	24667	3142	2366	776
Shanshui	45401	2939	2337	602

Note: The mean numbers of live and non-live births and sex ratios at birth have been obtained by using svy commands of the statistical package STATA 6.0. The four study counties were used as four strata and, within each stratum, townships have been considered as primary sampling units (psu). In columns (1)-(3) means and standard errors are also weighted by using the sampling weights of the Card System. Finally, in column (4), means and standard errors are weighted and adjusted for the complex, multi-stage sampling design of the Follow-up Survey. Sample sizes are unweighted.

recorded only by the Card System refer to the areas B_1 and B_2, and to the points within the lined parallelogram not included in the dark shaded areas.

As it is evident in Figure 3, linked records principally refer to women of reproductive age, who are likely to have the highest fertility in the Card System. On the contrary, records only in the Card System mainly refer to women either at the early beginning (area B_1) or toward the end (area B_2) of their reproductive life span. These youngest and oldest cohorts in the Card System give birth to a smaller number of children, so producing the observed disagreement of means of live and non-live births between linked and unlinked records within the Card System (Table 4, columns 2 and 3, top and bottom panels).

Yet, the different sampling design of the Card System and the Follow-up Survey does not account for differences in sex ratios at birth across the two groups of records. More insight into these discrepancies might be given by recalling that some of the linked records between the Card System and the Follow-up Survey refer to villages sampled by the Baseline Survey as well. In Table 6, these two groups of records are considered separately.[37] "Old Villages" refer to the villages sampled by the Baseline Survey in December 1991, while "New Villages" to those sampled only by the Follow-up Survey in June 1994. It emerges that old villages tend to have higher (and somehow too high) sex ratios at birth than new villages, whereas new villages report more live births than old villages.[38]

The greatest majority of births occurring in new villages (77 percent) are first births while most births occurring in old villages (65 percent) are births of second or higher order parities. The disagreement of sex ratios at birth between linked and unlinked records in the Card System might therefore arise only because of the different parities of the births considered in the two cases. In other terms, differences in reporting between those only in the Card System and those sampled by the Follow-up Survey as well could simply reflect the fact that across the surveillance cards births are more evenly distributed with respect to parity than in the linked records for the Follow-up Survey, which includes mainly first births. The sex ratio at birth for linked records in the Card System would therefore be representative only of sex ratio at first births. Yet, by restricting the analysis only on first births (Table 7), the disagreement between the unlinked and linked records in the Card System emerges as striking as before, suggesting the existence of a real underlying disagreement between the two groups of records.[39]

We record vital events for a group of women. Two years later, we ask a subgroup of those women to report retrospectively the succession of demographic events they experienced during a certain period of time. The sex ratios at birth calculated for the subgroup of women sampled both times strongly agree; yet, they are systematically different from those calculated for the original group of women (which are also more "normal"). Since there was no possibility for the women in the subgroup to

[37] It has to be noted that a "sensitisation effect" induced by the Baseline Survey has to be excluded given the small percentage of people in the linked records who had been previously sampled by the Baseline Survey (2,709 out of 11,521 linked records).

[38] Note that reporting of live births is similar for records only in the Card System (Table 6, column 1) and linked records across the Card System and the two surveys (Table 6, column 4), because both groups of records refer to women at the end of their reproductive life span. Yet, sex ratios at birth are extremely different even in these two cases.

[39] The matching of first births across the Card System and the Follow-up Survey is fairly good (kappa=0.68) despite not as good as for all births considered together (see Table 5).

Table 7. Number of live births and sex ratio at birth by county: First birth order, Card System and Follow-up Survey, June 1992 through June 1994.

	Card System			Follow-up Survey
	All records	Cards only	Respondents with linked records	
	(1)	(2)	(3)	(4)
Live births				
Huasheng	1812	1090	297	323
Pangxie	4710	2478	326	327
Ciqixian	2073	1132	315	346
Shanshui	5592	3327	419	454
Sex ratio at birth				
Huasheng	1.07 (.078)	1.12 (.076)	0.93 (.112)	1.03 (.117)
Pangxie	1.04 (.046)	1.05 (.048)	0.94 (.078)	1.06 (.078)
Ciqixian	1.05 (.051)	1.07 (.062)	0.96 (.157)	1.00 (.059)
Shanshui	1.02 (.061)	1.01 (.069)	1.14 (.055)	1.15 (.068)

Note: Sex ratios at birth have been obtained by using svy commands of the statistical package STATA 6.0. The four study counties were used as four strata and, within each stratum, townships have been considered as primary sampling units (psu). In columns (1)-(3) sex ratios and standard errors are also weighted by using the sampling weights of the Card System. Finally, in column (4), means and standard errors are weighted and adjusted for the complex, multi-stage sampling design of the Follow-up Survey. Numbers of live births are unweighted.

know that they would have been sampled twice, what can be a reasonable explanation for these discrepancies?

The inconsistencies of the results across the Card System and the Follow-up Survey might acquire a different connotation if it is assumed that, within the context of the Chinese family planning policy, demographic data collection during the INCRC project was affected in two main competing ways. First, the political context biased individual reporting towards over-reporting of female births at parity one, because most couples would have not been allowed otherwise to have another child if they already had a son. Second, administrative reporting could have been strongly influenced by the need to conform to the demographic targets set by the Chinese family planning policy, specifically by the need to ensure that the overall ratio of higher births to first births was consistent with that implied by the One-Son-Two-Child policy. These assumptions are consistent with the existing literature on the topic, as it has been illustrated before.

From this point of view, the fact that unlinked records from the Card System give more "normal" results than the Follow-up Survey in terms of the number of live and non-live births and sex ratios at birth (see Table 4, columns 2 and 4) can be interpreted as reflecting the greater influence of administrative over individual reporting. In

order to guarantee that fertility behaviors are consistent with the Chinese family planning policies, administrative reporting would in fact easily be biased toward reporting of more "normal" sex ratios at birth, especially given the increased attention of scholars, policymakers, and the general public to the problem of the increase of reported sex ratios in China.[40] At the same time, lower values of sex ratios at birth in the Follow-up Survey might show evidence of over-reporting of female births for all counties except Shanshui, where the higher value of the sex ratio at birth seem to suggest over-reporting of male births (see Table 3). In Huasheng, Pangxie, and Ciqixian, over-reporting of girls for all parities arise mainly because of over-reporting of first female births; the same is true in Shanshui with reference to over-reporting of male births (see Table 6). In Huasheng, Pangxie, and Ciqixian, this evidence is therefore consistent with the fact that, at the individual level, the most efficient strategy to continue childbearing without violating the Chinese One-Son-Two-Child policy is to report a female birth even when a male birth has already occurred.

Inconsistencies within the Card System (between those who were sampled by the Follow-up Survey and those who were not) might suggest again contextual interferences on the process of data collection. The INCRC Card System indeed provided the sampling frame for the Follow-up Survey, so that the local administrative units were aware of who in the Card System would have been included in the Follow-up Survey. Given the time-lag between recording of demographic events on the cards and transferring this information to a computer, it is not difficult to imagine that the card information for those sampled by the Follow-up Survey were computerized after the end of the survey in order to allow the two types of information to agree. Alternatively, it might not be hard to imagine that, during the actual process of data collection for the Follow-up Survey, the interviewers brought the cards to the interviews and used the information there contained to cross-examine the respondents. This would also explain the extremely high agreement across the linked records, which show virtually no variability across townships.

Conclusion

Statistical systems are supposed to be "neutral" measurement devices that do not interfere with the process of data collection. Yet, statistical systems can rarely be neutral observational devices as they are generally intended to be. This is because such systems develop around a complex social interaction between those who collect the data and those who provide it (such as the interviewer and the respondent in a sample survey), as well as the context in which the demographic data collection process is actually carried out.

The study of the imbalance of reported sex ratios at birth in four counties in rural China seems to show that this complex social interaction can result in a systematic bias in the data and, therefore, in lower quality information. The context in which the demographic data collection process actually takes place seems to exert a crucial influence on individual and administrative reporting in the four study areas. This influence produces systematically different results when the imbalance of sex ratios at birth is measured on the basis of the data obtained from the

[40] For example, see Aird; Banister; Hull; Hull and Wen; and Johansson and Nygren. In June 1986, the State Family Planning Commission organized a seminar on the topic of sex ratios. A summary of the seminar can be found in Xian Zhao, "A Summary Report of the 'Whole China Seminar on Population Sex Ratios'," in *Family Planning Yearbook* (Beijing: People's Hygiene Publishing House, 1987).

same respondents with different methods of data collection.

From these considerations, two major problems seem to emerge. First, which data source should we consider more reliable? There can be no single answer to this question. The observed inconsistencies across different methods of data collection cast doubts upon the validity of the indicators adopted and the reliability of the methods themselves. Second, generalizations about the emerging of the imbalance of reported sex ratios at birth as a result of sex-selective abortion, sex-selective infanticide, and sex-selective underreporting in the study areas might be misleading. In fact, if the measurement process of the sex ratio at birth depends upon the context in which it actually takes place, then the presence of systematic other than random errors in the data seriously limits their interpretation outside that context.