Sampling Bias in the China Population Censuses since 1982

Christina Jenq *

September 23, 2015

PRELIMINARY AND INCOMPLETE. PLEASE DO NOTE CITE OR DISTRIBUTE WITHOUT AUTHOR'S PERMISSION.

Abstract

There is a growing recognition that micro-level data collected and published by China's various statistical agencies are of unclear reliability. This paper is a preliminary exploration of the consistency of samples of the China Population Census since 1982. By comparing cohort measurements of sex ratios and educational attainment across Census waves, I document the growing extent of sampling bias since the 1982 Census. Specifically, I find increasing under-sampling of males and the less-educated since 1982, with the exception of 2010. After considering various explanations for the cohort discrepancies, the evidence seems to support the explanation of increased under-sampling due to increasing migration interacting with institutional incentives to evade tracking by statistical authorities. I then calculate cohort sex ratios and cohort educational attainment that are adjusted for tis sampling bias, and provide a few estimates of the extent to which different Censuses under-sampled males and the less-educated. Lastly, I discuss a few welfare implications for this systematic under-tracking, and suggest directions for future research.

Key words: China Population Census, Sampling Bias
JEL classifications:

*Institute for Emerging Market Studies (IEMS) and Institute for Advanced Study (IAS), Hong Kong University of Science and Technology, email: cjenq@ust.hk

1 Introduction

There is a growing recognition that both aggregate and micro-level data collected and published by the various statistical agencies of China, the world's largest economy, are of unclear reliability. While recent academic work has analyzed the representativeness and reliability of Chinese statistics related to GDP and industry, little in-depth work has been done to evaluate the representativeness of samples of micro-level household surveys.

This article analyzes the consistency of samples of the China Population Census, a set of data on which not only social science researchers rely to assess social and demographic indicators of interest, but also on which mainland Chinese government agencies rely to recalibrate various population and demographic estimates.

I document and quantify the growing extent of sampling bias in the various waves of the China Population Census since 1982[1]. By conducting a simple counting exercise of the relative numbers of males and females and the high- and less-educated for particular cohorts across Census waves since 1982, I find evidence of increasing under-sampling of males and the less-educated after 1982. The exception is the 2010 Census, which, while exhibiting some evidence of under-sampling males, does not seem to under-sample the less-educated.

While it is difficult to precisely determine why the sampling deteriorated after 1982, I consider various explanations for cohort discrepancies across samples and show that the most plausible explanation is increased migration interacting with institutional incentives to evade tracking by authorities. I also hypothesize that the under-sampling of male migrants relative to female migrants is related to the increased difficulty of tracking individuals in industries that male migrants typically work in.

I then develop a simple method for correcting the under-sampling of males by exploiting the observation that cohorts were more consistently tracked when they were less mobile and most likely to be tracked by authorities, approximately age 5-15. I also develop a simple method for correcting the under-sampling of the lower-educated.

Lastly, I generate estimates for the relative amount of males and less-educated that were under-sampled in each Census wave. I estimate that for the 1990 Census, males were undersampled by 0.2 to 0.5%, while in the 2000 Census, males were under-sampled between 1.3 to 3%, and between 8 to 11.6% for the 2005 Census. I also estimate that the lower educated were undersampled by 11-19% if it is assumed that the under-sampled had 3 years of education.

The larger goal of this article is to not only to equip social science researchers with tools

[1]Analysis on the 1953 and 1964 China Population Census is excluded due to the lack of availability of sufficiently detailed tabulations of this data

to more properly analyze and interpret Census data, but also to spur the re-examination of results using Census data.

Section 2 reviews related literature and Section 3 describes the China Population Census in more detail. Section 4 describes in more detail the empirical strategy and and presents results of the analysis. This section also includes corrected measures of sex ratios and share of highly-educated. Section 5 discusses the implications of these results for a few selected papers and published official statistics. Section 6 concludes.

2 Literature

Several articles and books have recently been published that have called attention to the reliability of China's micro-level and macro-level data.

Rawski (2001), Rawski (2002), Adams and Chen (1996), and Xiaolu and Lian (2001) question the reliability of GDP statistics, with Rawski focusing especially on the years since 1998. In a series of articles, Holz (2014), Holz (2013), Holz (2008), Holz (2004), Holz (2003), and Holz and Lin (2001) analyze published industrial, output, and GDP statistics.

Scharping (2001) explains the meaning and limitations of Chinese official population statistics.

More recently, Crabbe (2014) is a recently published book that gives a solid introduction to understanding Chinese official published statistics as products of political, social and economic realities. Koch-Weser (2013) evaluates the reliability of China's GDP statistics. It discusses how economic measures often remain unrepresentative and their methodology and definition opaque and notes that many data series are subject to revision without adjustment to historical data.

3 Background

The following section provides some background on the China Population Census.

The statistical agency of the People's Republic of China, currently known as the National Bureau of Statistics (NBS), conducted Census data collection in 1953, 1964, 1982, 1990, 2000, and 2010, with a "mini-Census" collected in 1995 and 2005. The 1953, 1964, 1982, and 1990 Census set July 1 as a reference date, but the 2000 Census set the reference date to November 1 to improve the probability of enumerating migrants at their actual rather than legal residence.

Banister (1984), in analyzing the 1982 China Population Census, notes that with the exception of military age males, the age and sex structure of the population is remarkable

consistent across the 1953, 1964, and 1982 waves, and that the 1982 Census is notable for its enumaration accuracy. Indeed, as noted by Lavely (2001) , China's social and political conditions before the post-1978 economic reforms once favored high quality enumeration. The population was relatively immobile and under the surveillance of a powerful and pervasive bureaucracy via the household registration (*hukou*) system that was started in 1958. Perhaps more importantly, there were few disincentives to full and accurate reporting for citizens or for officials. Under these circumstances China was able to produce census data with remarkable accuracy.

But then, as commentators like Lavely (2001) and Wu and He (2015) have noted, several reforms occurred that made population tracking increasingly difficult. Economic reforms in the countryside jump-started agricultural productivity, freeing formerly agricultural workers to pursue non-agricultural activities. Migration slowly increased through 1990, and then skyrocketed afterwards. Figure 3b shows the relative size of migration flows across provinces and between regions of China between 1982-1990, and then 1990-2000.

Officials had disincentives to track these rural-to-urban migrants. Lavely (2001) gives two examples of these disincentives. First, local officials were being evaluated by GDP per capita measures, which were larger the smaller the local population count. Secondly, the introduction of the One-Child fertility policies in the late 1970's gave both families and local officials substantial disincentives to report out-of-plan births.

As for the 2010 Population Census, Wu and He (2015) notes that an unprecedented amount of administrative effort was put into accurate enumeration, and that postenumeration checks suggest a low undercount rate.

Appendix Section C.2 provides further details on the China Population Census.

4 Empirical Strategy and Results

The data sources used in this analysis are a 1% sample of the 1982, 1990, and 2000 Population Census, a 1-in-500 sample of the 2005 Mini-Census, and individual data from the 2010 China Family Panel Survey (2010 CFPS). The survey weights included in the 2010 CFPS are calibrated to the results from the 2010 Census. More details about the China Population Census are given in Appendix Section C.2

The empirical strategy of this paper is to compare the measurements of two largely static demographic characteristics, gender and educational attainment, for birth cohorts tracked at different points in time in successive Census waves.

We first examine the measured cohort sex ratios[2] across Census samples, shown in Figure

[2]Cohorts are smoothed by including birth years 3 years above and below the reference birth year.

1. Cohorts aged between 5 and 40, inclusive, during each year the Census was taken are included for comparison. We examine cohorts age 5 and above to abstract from issues of sharp shifts in cohort sex ratios between ages 0-5 due to either systematic under-reporting of young female children or excess female mortality in the early years. Excess female mortality in the early years is documented by Coale and Banister (1994). Under-reporting could be due to parents who, facing strict fertility limitations, hide these births from authorities in order to avoid birth fines from trying to have a male. Indeed, Lavely (2001) notes that parents conceal excess births to avoid fines, while officials keep them off the books because cadre job evaluations are based primarily on birth planning performance. Presumably, these females would still be alive and eventually be tracked by institutions once they reach school age.

We limit the cohorts to a maximum of age 40, as life tables from Banister and Hill (2004) show that excess male mortality rates sharply increase after this age.

Figure 1 shows large discrepancies in measured cohort sex ratios across the Census samples [3]. Specifically, each cohort in the Census samples since 1982, with the exception of the CFPS 2010 sample, seems to be increasingly female as cohort sex ratios dip. We note that biologically natural sex ratios occurring in the absence of sex-selective abortion, ranges between 105-107 males per 100 females. In fact, it is extremely strange to see sex ratios dipping below this range, as male preferences have generated cohort sex ratios higher than this range.

We note in particular dips in cohort sex ratios for those cohorts in their early 20's. In 1990, there is a dip in cohort sex ratios around the 1970 birth cohort (age 20), around 1976 for 2000 (age 24), and around 1982 for 2005 (age 23), and around 1986 (age 24) for 2010.

One explanation is that males in the military around their early 20's were excluded from these samples. While this is likely true for the 1982 Census, given thatBanister (1984) note that military members were excluded and find the 1982 Census to be be of high quality, this reason cannot completely explain the discrepancies of the other Census samples, as the implied magnitudes of the size of the military are implausible. Assuming all the "missing males" went to the military implies, back-of-the-envelope estimates show[4] this implies an increase in the size of the military of about 35 million males, whereas the entire size of the military is only about 3 million and the only increase in the Chinese military manpower, as documented by the World Bank[5], is about 45,000.

Another explanation is gender differences in migration out of China. A growth of the

[3]The CFPS 2010 sample results are much less smoother due to a much smaller sample size

[4]See Appendix C.1 for more details.

[5]http://data.worldbank.org/indicator/MS.MIL.TOTL.P1

outmigrant population of about 35 million males would be necessary to explain these results, but, as with the military explanation, this is also implausible as the total number of overseas Chinese was only about 35 million as of 2007[6].

Another explanation is that this may be due to increased relative male mortality between the Census waves. But, according to mortality rates calculated by Banister and Hill (2004)[7], shifts in cohort sex ratios due to gender differences in mortality should be no more than about 1 less male per 100 females between the ages of 5 to 40[8]. The discrepancies for the 1980 birth cohort, for example, are on the order of 5 to 20 less males per 100 females relative to the 1982 Census sex ratios. Furthermore, this explanation would suggest greater discrepancies for older age cohorts rather than a greater discrepancy around the age 20 cohort.

Another hypothesis is that officials may over-count females to lower measured fertility rates, which is something local officials may be implicitly or explicitly evaluated on. But this is not a likely explanation since it does not explain why females around age 20 would be over-counted relative to females in their 30's.

Yet another hypothesis is that the statistical bureaus possess accurate data but only selectively release data of men who are low-income and low-educated due to political concerns about damning statistics on inequality. But this is unlikely since the official published tabulations also show a similar qualitative trend in cohort sex ratio discrepancies (see Appendix Figure C.1). This figure shows that cohort sex ratios since 1982 are becoming more female, with large discrepancy for the 2005 published Census tabulations.

The last, and more plausible, explanation is that males were progressively more under-sampled in Censuses. In fact, we hypothesize this under-sampling is related to increasing migration flows, as the dips in cohort sex ratios in the 1990, 2000, and 2005 samples all correspond to about age 20, an age at which people are most likely to migrate. Figure 3a lends some support for this hypothesis. It shows the change in sex ratios for each cohort for each region[9], given in log units. Regions above 0 represent a male-biased shift in cohort sex ratios, and regions below 0 represent a female-biased shift in cohort sex ratios between 1982-2000. The figure shows that for those born 1950 and afterwards, the West, North/Northeast and South-Central regions experienced a male-biased shift in cohort sex ratios, while the East region experienced a sharp female shift in cohort sex ratios. Because the provinces in the Eastern region are more populous than those in the other regions, the overall result is

[6]http://www.china-embassy.org/eng/qwgz/t297510.htm

[7]Table A3, which provides mortality rates adjusted for incomplete death reporting

[8]For example, Table 2 from Coale and Banister (1994) shows that for those born in the 1947-1951 birth cohort, the sex ratios drop by about 1 male per 100 females due to male-biased mortality rates by the time this cohort reaches about age 40.

[9]For this and subsequent analysis, each individual is assigned the province of birth to abstract from migration issues

a downward shift in cohort sex ratios. These patterns suggest that females were migrating from the other regions to the Eastern region. However, we note that even if there were no under-sampling and migrants were simply mis-identified in terms of birth province, a discrepancy in sex ratios would not be observed at the aggregate level, as shown in Figure 1.

Meanwhile, Figure 3b compares interprovincial migration flows in 1985-1990 with those in 1995-2000. It shows that interprovincial migration, particularly to the Eastern provinces and within the Eastern region provinces, significantly increased in the latter time period. We find that the magnitude of these discrepancies correspond to roughly the magnitude of rural-to-urban migration observed since 1982. Between 1982 and 1990, there was little rural-to-urban migration and, notwithstanding the 1982 cohorts of military age, sex ratio discrepancies were relatively small. But these discrepancies become much larger when comparing the 2000 and 2005 samples to the 1982 sample, while migration flows were much larger during these periods.

Commentary from Lavely (2001) provides some potential reasons for the under-sampling. The article notes that the floating population is not easy to count in the best of circumstances, analogous in difficulty to counting undocumented aliens or the homeless in the U.S. Lavely (2001) also notes that while census workers in the 1990 and 2000 waves went to unusual efforts to enumerate urban migrants, including an extensive pre-enumeration of sojourners sleeping outdoors and in public officials, local officials, for various reasons, tried to minimize the population of their administrative domain. Low rates of population growth reflect success in population control, and small population totals boost per capita measures of income and productivity. Thus even as census workers valiantly sought out the floating population by scouring construction sites and searching under bridges. local officials in some places subjected migrants to "census fees", "security fees", and "temporary residence fees" in an attempt to discourage them from being counted.

Indeed, Lavely (2001) notes that the 2000 Census was dogged by rumors of vast undercounts. Some provinces, such as Henan, Hunan, and Shaanxi, were said in newspaper accounts to have counted millions fewer people than expected.

4.1 Cohort Educational Attainment

Figure 2 presents further evidence that the un-tracked individuals in the Census samples consisted of a disproportionate share of migrants. It gives cohort mean years of education for Census samples from 1982 to 2010. 3-year cohorts are constructed to smooth out roughness in the raw data. Only cohorts aged 19-50 are included. 19 is chosen as the beginning age year as most have completed high school education by this time.

We see that for the birth years between approximately 1940-1962, 1990, 2000, and 2005 measurements of cohort educational attainment are increasing relative to the 1982 measurements. Yet, for some reason the 2010 measurement gives educational attainment below the 1990-2005 measurements and, in fact, seems to overlap with the 1982 measurement.

Differences in mean years of education across Census samples could be due to the same reasons discussed for sex ratios, i.e. educational differences in out-migration, educational differences in mortality rates, and under-sampling of the disproportionately lower or higher educated. Another mechanism, additional education attained between successive Census waves, would result in higher measured levels of education by cohort.

The explanation of cohorts receiving additional education between the Census waves is implausible for several reasons. If we assume that there was no biased sampling in 1990 and 2000 and that the differences are merely due to educational attainment shifts, then the discrepancies imply that those born between roughly 1940-1960 attained significantly more education. Take for example, the birth cohort of 1950. In 1982, this cohort would have been 32 years old. In 1990 this cohort would have been between 40 years old, and 50 in 2000. The discrepancy implies that this cohort raised their mean years of education by one year between the ages of 32 and 40. This implication seems implausible since there is little evidence, anecdotal or otherwise, that people in their 30's went back into the formal educational system to attain more formal schooling.

Furthermore, the 1960-1980 cohort measurements show that it seems as if they became more educated between 1990 and 2005, but then suddenly became less educated by 2010. One explanation is that the 1990,2000, and 2005 samples are correct and these cohorts did become more educated in that period, and that the 2010 sample under-sampled the educated. Another explanation is that a significant portion of educated people died between 2005-2010. Another explanation is that the 1990-2005 samples increasingly under-sampled the less-educated and the 2010 sample is more correct.

The first two explanations are implausible. The 2010 Census under-sampling the educated is unlikely since the highly educated are usually tracked by more formal institutions. By definition, the highly educated would at least have been tracked by more educational institutions. Likewise, the possibility that many educated people suddenly died between 2005-2010 is implausible, since mortality rates among the less-educated are always higher than the more-educated.

Given the previous evidence about sex ratio discrepancies and migration activity, the second explanation is much more plausible. In addition, several sources have noted the relative enumeration accuracy of the 1982 and 2010 China Population Censuses relative to the other Census samples in between (Banister, 1984; Wu and He, 2015). While some of

the discrepancy between the 1990, 2000, and 2005 cohort years of education could be due to additional completed schooling between the time of the 1990 Census and the 2000 and 2005 Censuses, the existence of a measurement from 2010 of a substantially lower cohort years of education, lower even than the 1990 measurement, provides strong evidence for under-sampling of the less-educated. Lastly, the 1982 calculated years of education and the 2010 calculated years of education seem to overlap each other quite well.

It is difficult to conclusively determine the reasons behind why the lower-educated and male were disproportionately under-sampled, given the opaqueness of the sampling procedures of China's National Bureau of Statistics. The evidence presented suggests that many of the under-sampled were probably migrants. Yet, if this is true, the question of why male migrants were more likely to be under-sampled relative to female migrants remains.

We hypothesize that male migrants tended to work in industries that made them less likely to be tracked by Census enumerator. Perhaps the fact that the nature of work was structured so that the workers are more mobile and less likely to be tracked. Appendix Table B.1 shows the breakdown of industry for those who have moved location since 1995 in the 2000 Census, and the sex ratio of employment. It shows that among migrants, males make up the preponderance of construction workers, a type of job that is more mobile and renders them less likely to be tracked by the statistical bureaucracy. Migrants in other male-heavy industries, such as certain manufacturing and mining sectors, may also be difficult to track by enumerators.

4.2 Calculating Adjusted Cohort Sex Ratios

The hypothesis that under-sampling of males and the lower-educated was related to differing migration propensities over time and space motivates an algorithm for constructing cohort sex ratio measurements that are adjusted for these sampling biases.

If migration is the main issue behind biased sampling, then the most accurate measurements of cohort sex ratios are taken when individuals are too young for migration. Because of other studies showing that age of first migration begins to happen around 15 years old, we choose the age range of 5-15 years as a period of relatively more accurate measurement. The lower range of age 5 accounts for the possibility of excess female mortality between birth and 5 years, as documented in Banister (1984), or for the deliberate hiding of female births.

Hence, we calculate adjusted cohort sex ratios using the following rules: 1) 1982 measurements are assumed accurate. 2)For the 1990, 2000, and 2005, cohort sex ratio measurements taken during ages 5-15 are considered more accurate than other measurements. 3) In cases where there are multiple sex ratio measurements for a given cohort that fall within the age

5-15 range, the maximum cohort sex ratio is taken.

The algorithm is as follows. Given the enumeration accuracy of the 1982 Census, cohorts from birth years 1940-1976 are used as accurate measures of cohort sex ratios as of 1982. Then for the birth years of 1977 to 1985, the maximum of the sex ratios measured in the 1982 and 1990 Census waves is taken as the accurate measures of the cohort sex ratios. In practice, the cohort measurements are taken from the 1990 Census as the sex ratios from the 1982 Census are lower, presumably due to the exclusion of military males. For birth years 1985 to about 1995 measurements from the 2000 Census are used. A bit of interpolation and extrapolation is included in the calculations.

Figures 4a and 4b show the results of this algorithm for calculating cohort sex ratios adjusted for biased sampling. Figure 4a shows how the adjusted cohort sex ratios are the maximum of the measured cohort sex ratios for the various Census samples between 1982-2005. Figure 4b also gives the adjusted cohort sex ratios by region. It shows cohorts born in the South-Central and Eastern regions had the highest increase in sex ratios.

Appendix Figures C.2 gives these adjusted sex ratio measures by source province. Provincial level municipalities are not included due to the changing city limits during the time period.

We note that these adjusted sex ratios are cohort sex ratios at a particular point in time, not sex ratios at birth. One would be able to construct sex ratios at birth by assuming a particular gender difference in mortality rates between the time of the Census measurement and birth.

4.3 Validating with Sex Ratios at Birth

We now validate these adjusted sex ratios with other independent measurements of sex ratios at birth (SRB) that are widely regarded as relatively accurate. Figure 5 compares the calculated cohort sex ratios adjusted for sampling bias with the recorded sex ratio at birth from two separate fertility surveys by five-year birth cohorts, excerpted from Table 3 of Coale and Banister (1994). If sampling were unbiased in all data sources, the only discrepancy between the reported Sex Ratios at Birth from the fertility surveys and the calculated "adjusted" sex ratio would be due to gender differences in mortality rates [10].

[10]One was a 1-in-1000 fertility survey from 1982, and the other was a 2-per-1000 survey from 1988. Both surveys asked a large sample of married women (311,000 age 15-67 in 1982, and 459,000 age 15-57 in 1988) for a lifetime history including each respondent's date of birth, her date of marriage, and the birth date of each child that she had borne. Coale and Banister (1994) and Banister (1984) report that that the 1982 fertility survey has proven to be consistent with the 1982 Census in terms of matching age-specific fertility rates with the number of children ever born. Births through 1982 are from the 1982 1-per-1000 survey, for 1982-1987 from the two-per-thousand survey, and for 1989 from the sample tabulation of the 1990 census, with the 1988 value being interpolated. Coale and Banister (1994) also notes that while in actuality the sex

We see that the two series are relatively close to each other, within 1-2% of each other. We note that the "choppiness" of the SRB series is due to the significantly smaller sample size of the fertility surveys.

4.4 Calculating Adjusted Cohort Years of Education

Cohort years of education corrected for sampling bias are also calculated. As previously discussed, the 1982 and 2010 measurements seem to be consistent with each other, and other researchers have noted the relative enumeration accuracy of the 1982 and 2010 Census. Hence, we rely upon the 1982 and 2010 measurements of cohort years of education to generate cohort educational attainment adjusted for sampling bias. Specifically, for the birth years of 1930-1962, the 1982 measurements are used, and for the 1965 through 1991 birth cohorts, the minimum of the 1990 and 2010 measurements are used, as the 1990 and 2010 measurements are relatively close to each other for certain birth cohorts. Cohort measurements are smoothed across series by interpolation.

Figure 4c shows the results of this calculation, by region. Interestingly, we see that educational attainment of the provinces in the Eastern region for those born before 1980 lagged behind the provinces in the North/Northeast region and the South-Central region, but then for those cohorts born after 1980, the educational attainment of the Eastern region outstripped the other provinces.

4.5 Quantifying Sampling Bias

We now quantify the amount of sampling bias in the Census samples since 1982. Table 1 gives various estimates of the relative amount of under-sampling of males and the less-educated given a few assumptions.

Columns 1 and 2 estimate the amount of males under-sampled relative to females in percentage points, while Columns 3 and 4 estimate under-sampling of the lower-educated populations. Columns 1 and 2 estimate the average discrepancy between the given Census year measurements and the calculated adjusted cohort sex ratios for a given range of birth years[11]. Visually, it would be the sum of the "area" between the Census birth cohort sex ratio series and the adjusted cohort sex ratio series for a particular range of birth years. Column 1 calculates the average discrepancies for the cohort sex ratios calculated from the

ratio at birth was close to the abortion-free ratio of 106 until about the early to mid 1980's, after which sex ratios at birth rose significantly due to sex-selective abortion, the high correlation between 1990 Census sex ratios and the sex ratios at birth reported in the fertility surveys indicate that female children who died early were neither reported in the Census nor recalled by their mothers in the fertility surveys.

[11]The birth years are chosen to correspond to age 5-40.

1990, 2000, and 2005 Census micro-data samples, while Column 2 calculates the average discrepancies for the sex ratios calculated from the published tabulations of the 1990, 2000, and 2005 Census[12]. Columns 1 and 2 show that while the published tabulations had less male under-sampling relative to the Census samples, each successive Census wave measured under-sampled males successively. While the 1990 Census under-sampled males between 0.2 to 0.5 percentage points, the 2000 Census under-sampled males by 1-3 percentage points, and the 2005 1% Mini-Census under-sampled males by 8-12 percentage points.

Column 3 shows the average discrepancy between the adjusted cohort years of education and the measured years of education for the 1990, 2000, and 2005 samples. This is simply the weighted average of the difference between the measured years of education and the adjusted years of education for a particular range of birth years, with the weights being the relative size of each birth cohort.

Column 4 calculates the implied percentage of the population that was under-counted if the un-tracked had 3 years of education. These estimates can be thought of as a lower bound for the percentage of the population that is untracked, since assuming a low level of education for the untracked decreases the relative amount of people needed to lower the average measured educational attainment. This percentage is calculated by solving for the amount of individuals possessing 3 years of education that would need to be added to the population to lower the measured educational attainment to the adjusted years of education. The birth year ranges are chosen to correspond to ages 19-60.

These calculations show that if the untracked all had about 3 years of education, then about 11-19% of the population was untracked.

5 Discussion

There are several implications of this research. Any study results using Census data from the 1990, 2000, or 2005 wave, whether in tabulation or micro-sample form, should be interpreted with the knowledge that these data samples increasingly under-tracked males and the less-educated. In particular, studies calculating cohort sex ratios from these data sources need to be revisited, as this article finds that cohort sex ratios are likely too low given the under-sampling of males.

In addition, because historical time series of population statistics are adjusted every time a Census is taken, it is quite likely that population estimates until the 2010 Census are also skewed by the under-sampling of males and the less-educated.

[12]Published tabulations are downloaded from China Data Online

Since the National Bureau of Statistics runs a variety of other household surveys, it is also quite likely that other household surveys run by them also under-sample males and the less-educated.

There are also implications for measures of poverty and inequality. Poverty and inequality estimates typically use household surveys with sampling weights calibrated to a comprehensive measurement of the population, such as the Census. To the extent that poverty and inequality estimates use sampling weights based on biased Census samples, they may understate the true extent of poverty and inequality.

There are also some interesting welfare implications for the un-tracked. Banister (1984) notes that enumeration is closely tied to hukou registrations, and that hukou registrations are closely tied to the rationing of grains, cotton, cloth, and other scarce items. While there have been many administrative reforms since the 1980's, many benefits are still tied to having official hukou registration in an area. The growing number of untracked in the Census since 1982 suggests that since 1982 there has been a relatively growing population of people who are not able to access the benefits tied to hukou registration.

Lastly, we note that the 2010 China Population Census seems to have somewhat improved upon its enumeration accuracy relative to the 1990, 2000, and 2005 Census. While the sex ratio measurements do seem to be female-biased, especially for the early 20's age cohort, the 2010 measurements of cohort educational attainment are consistent with the accurate 1982 measurements. Further, Wu and He (2015), in their article about the evolution of the China Population Censuses, notes that the 2010 China Population Census has significantly improved relative to previous Census waves.

6 Conclusion

This article has attempted to evaluate the quality and representativeness of the China Population Census since 1982. It has documented systematic and increasing under-sampling of males and the less educated since 1982. We present some preliminary evidence showing that increasing migration may have something to do with the increased sampling. Then we exploit the fact that migration propensities are lowest below age 15 to calculate cohort sex ratios adjusted for sampling bias. We also calculate cohort years of education adjusted for sampling bias by using the fact that the 1982 and 2010 measurements seem to be the most accurate.

This article suggests various future avenues for research. Because these findings call into question the representativeness of not only the 1990-2005 Census, but also that of other household surveys conducted by the NBS, we deem it a useful exercise to generate more

correct sampling weights for household survey data. Then these survey weights could be used to re-estimate measures of poverty and inequality, such as Gini coefficients for income and wealth.

References

Adams, F. G. and Y. Chen (1996). Skepticism about chinese gdp growth–the chinese gdp elasticity of energy consumption. *Journal of Economic and Social Measurement 22*(4), 231–240.

Banister, J. (1984). An analysis of recent data on the population of china. *Population and Development Review*, 241–271.

Banister, J. and K. Hill (2004). Mortality in china 1964–2000. *Population studies 58*(1), 55–75.

Coale, A. J. and J. Banister (1994). Five decades of missing females in china. *Demography 31*(3), 459–479.

Crabbe, M. (2014). *Myth-Busting China's Numbers: Understanding and Using China's Statistics*. Palgrave Macmillan.

Fan, C. C. (2005). Interprovincial migration, population redistribution, and regional development in china: 1990 and 2000 census comparisons. *The Professional Geographer 57*(2), 295–311.

Holz, C. A. (2003). "fast, clear and accurate": How reliable are chinese output and economic growth statistics? *The China Quarterly 173*, 122–163.

Holz, C. A. (2004). China's statistical system in transition: Challenges, data problems, and institutional innovations. *Review of Income and Wealth 50*(3), 381–409.

Holz, C. A. (2008). China's 2004 economic census and 2006 benchmark revision of gdp statistics: more questions than answers? *The China Quarterly 193*, 150–163.

Holz, C. A. (2013). Chinese statistics: classification systems and data sources. *Eurasian Geography and Economics 54*(5-6), 532–571.

Holz, C. A. (2014). The quality of china's gdp statistics. *China Economic Review 30*, 309–338.

Holz, C. A. and Y.-m. Lin (2001). Pitfalls of china's industrial statistics: inconsistencies and specification problems. *China Review*, 29–71.

Koch-Weser, I. N. (2013). The reliability of china's economic data: An analysis of national output. *US-China Economic and Security Review Commission Staff Research Project*.

Lavely, W. (2001). First impressions from the 2000 census of china. *Population and Development Review 27*(4), 755–769.

Rawski, T. (2002). Measuring china's recent gdp growth: where do we stand. *China Economic Quarterly 2*(1), 2002.

Rawski, T. G. (2001). What is happening to china's gdp statistics? *China Economic Review 12*(4), 347–354.

Scharping, T. (2001). Hide-and-seek: China's elusive population data. *China Economic Review 12*(4), 323–332.

Wu, X. and G. He (2015). The evolution of population census undertakings in china, 1953–2010. *China Review 15*(1), 171–206.

Xiaolu, W. and M. Lian (2001). A reevaluation of china's economic growth. *China Economic Review 12*(4), 338–346.

8 Figures

Figure 1: Cohort Sex Ratios, by Census Year

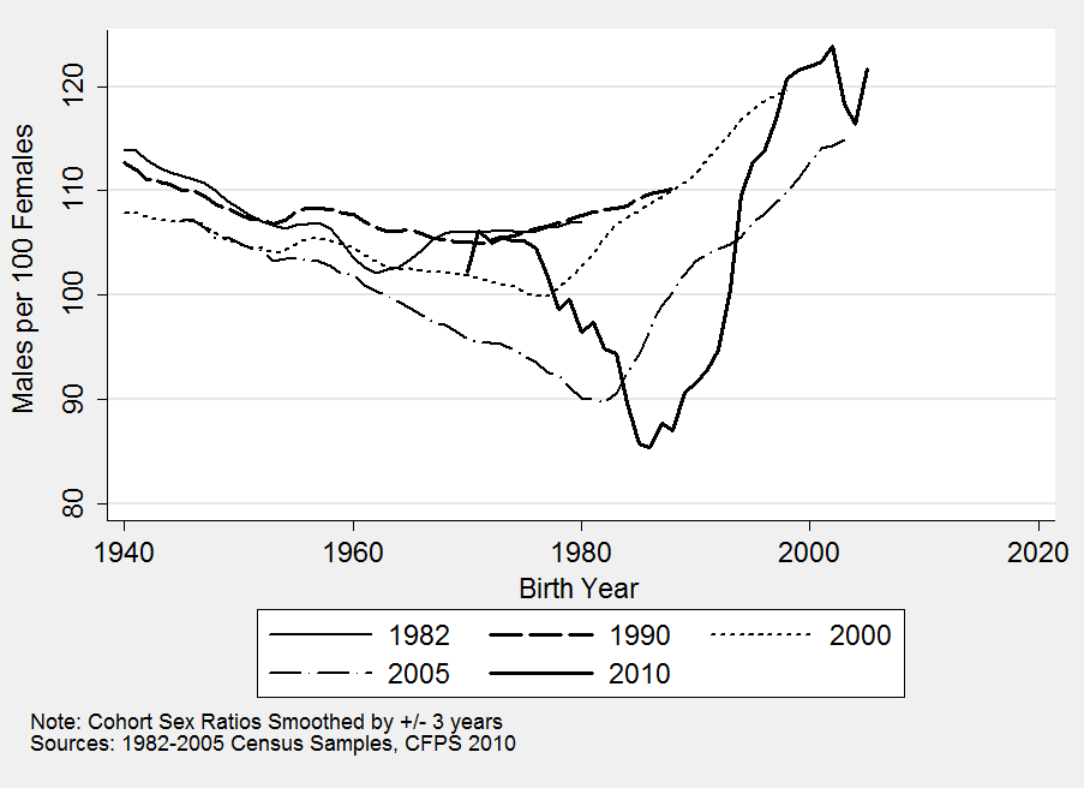


Figure 2: Cohort Mean Years of Education, by Census Year

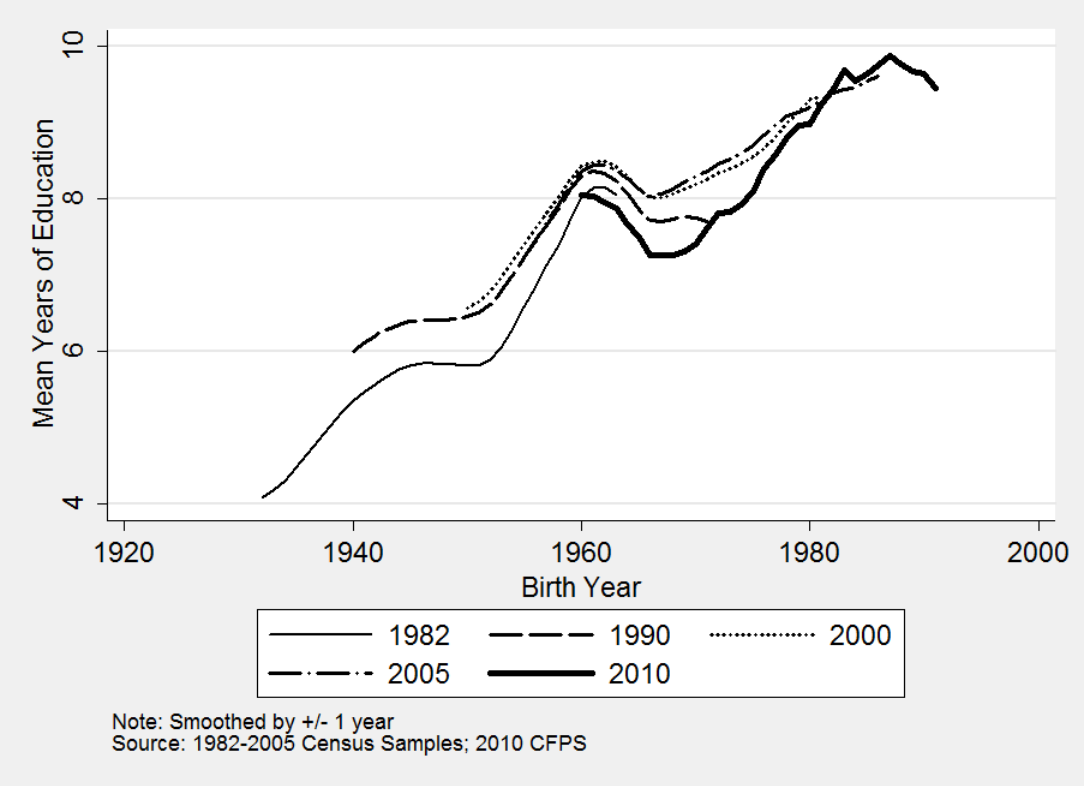


Figure 3a: Cohort Sex Ratio Differences by Region, 1982-2000

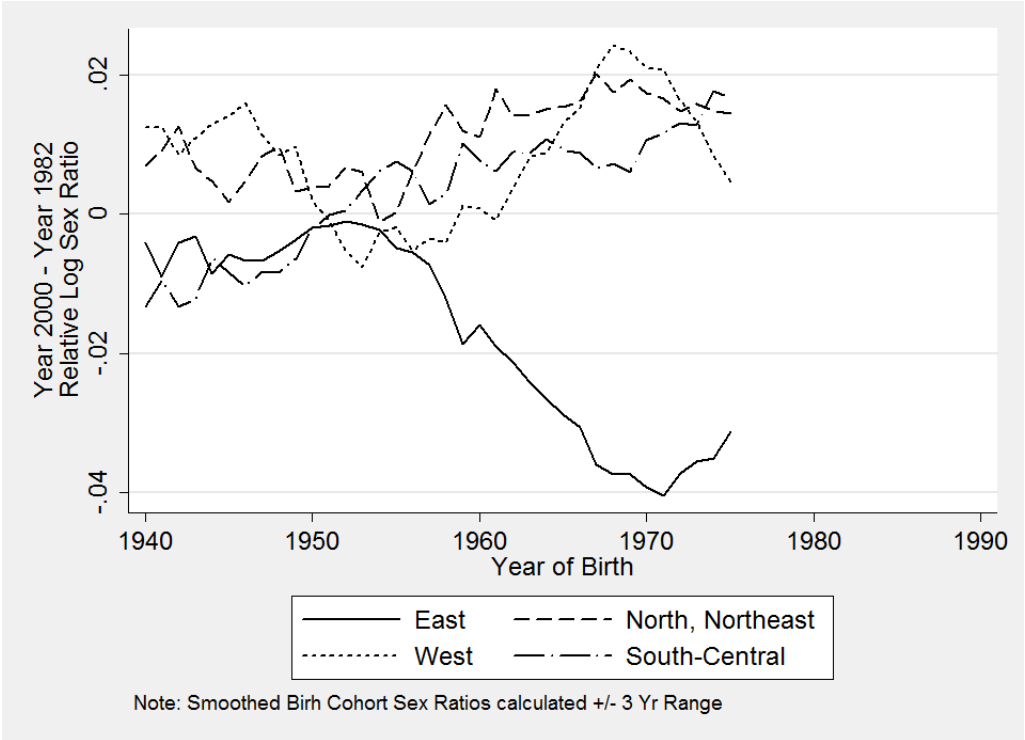


Figure 3b: Migration Patterns

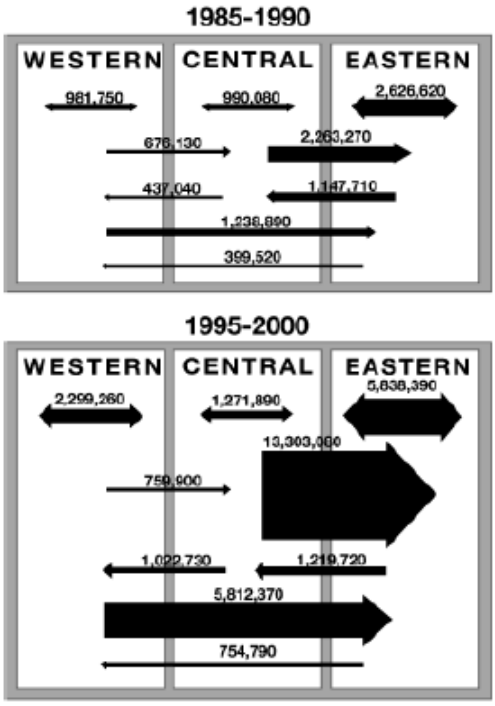


Figure 3 *Interprovincial migration within and among the three regions.*
Sources: National Bureau of Statistics (2002); State Statistical Bureau (1992).

Source: Fan (2005)

Figure 4a: Adjusted Cohort Sex Ratios

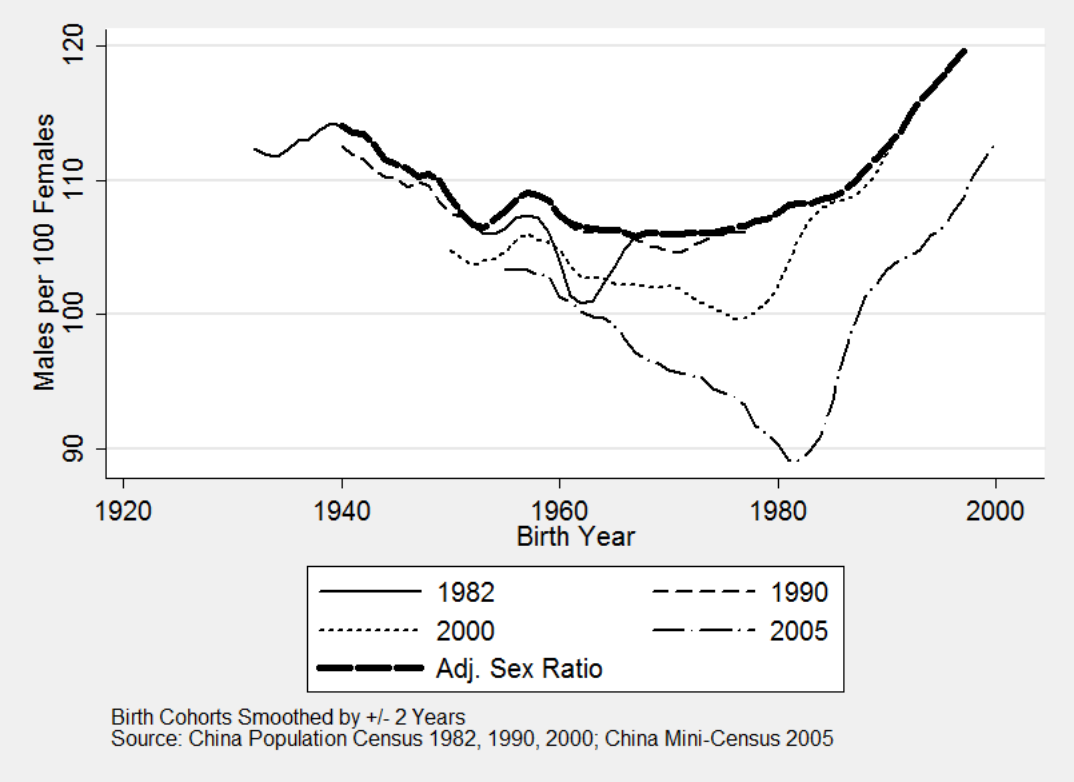


Figure 4b: Adjusted Cohort Sex Ratios, by Region

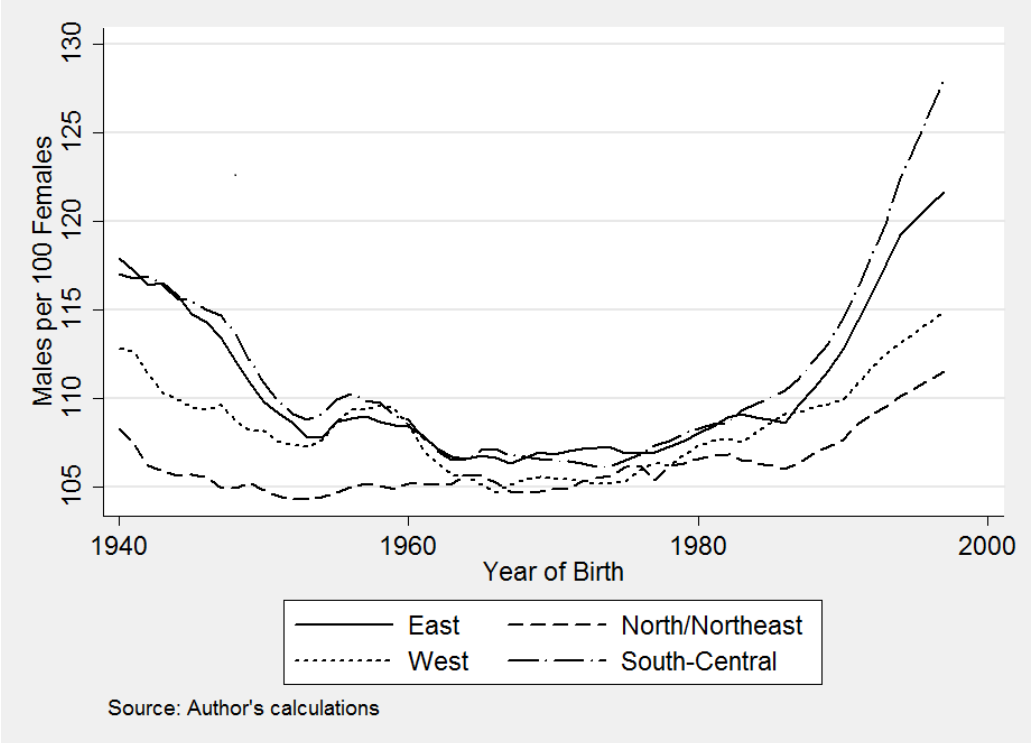


Figure 4c: Adjusted Cohort Years of Education, by Region

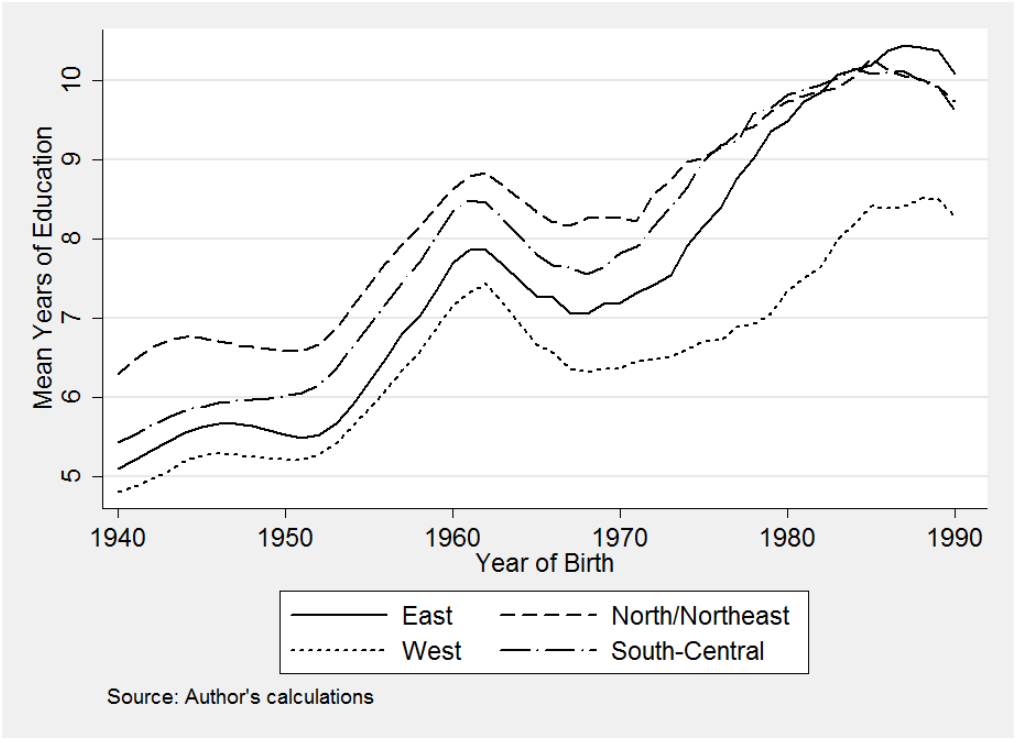
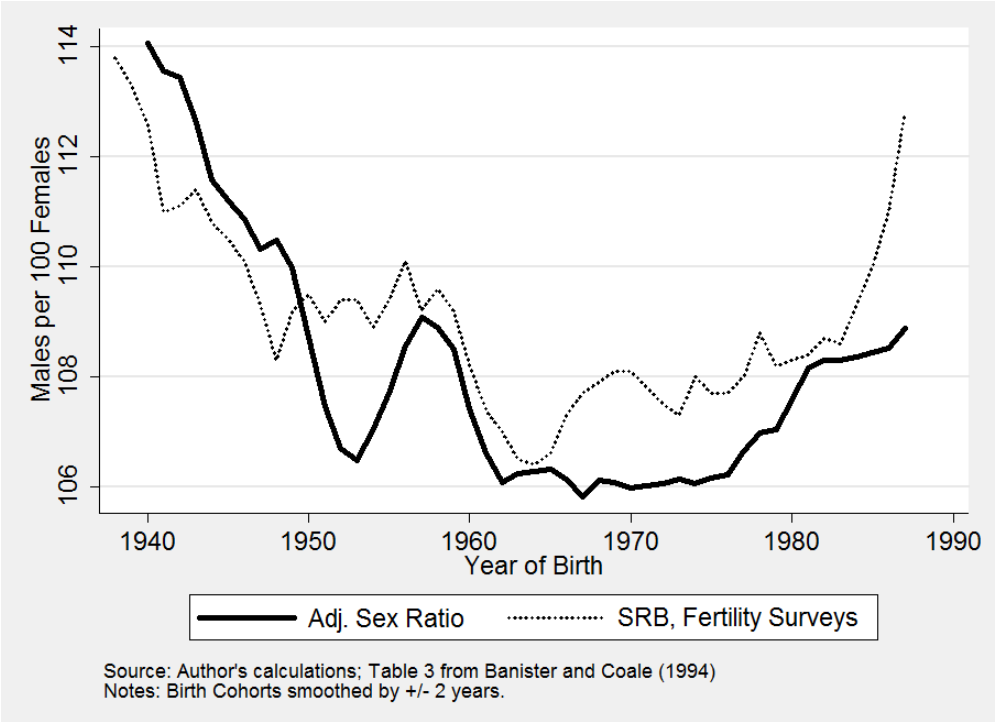


Figure 5: Cohort Sex Ratios at Birth, Coale and Banister (1994)



9 Tables

Table 1: Estimates of Relative Under-sampling

Census Year	Males (% Points)		Lower Educated*	
	Census Samples	Pub. Tabulations	Years of Education	Percentage Points**
1990	0.20	0.49	0.49	18.8
Birth Year Range	*1950-1985*		*1930-1971*	
2000	2.91	1.28	0.55	14.0
Birth Year Range	*1960-1995*		*1940-1981*	
2005	11.62	7.94	0.55	11.1
Birth Year Range	*1965-1997*		*1945-1986*	

** Census samples are analyzed*

*** Percentage Points calculation assumes that all individuals missing have 3 years of education.*

A Appendix Tables

Table B.1: Migrant Industry Characteristics

Name	Share of Migrant Employment	Sex Ratio
Manufacturing	35%	100
Agriculture	16%	37
Wholesale and retail trade, food and beverage industry	13%	109
Health, Phys. Ed., Social Welfare, Scientific Research, Technical Services	7%	77
Construction	6%	658
Social Services	5%	91
State organs, Mass organizations	5%	225
Commercial Brokers and Agents	4%	84
Transport, Warehousing, Postal, Telecom	4%	313
Finance and Insurance	2%	129
Mining	2%	495
Electricity, gas and water production and supply	1%	193
Information and Consultancy Services, Other Services	1%	185
Geological Prospecting and Water Conservancy	0%	270

Source: China Population Census 2000

Notes: Migrant is defined as anyone who moved to the current location within the last 5 years.

B Appendix Figures

Figure C.1: Cohort Sex Ratios, Published Tabulations from the Census

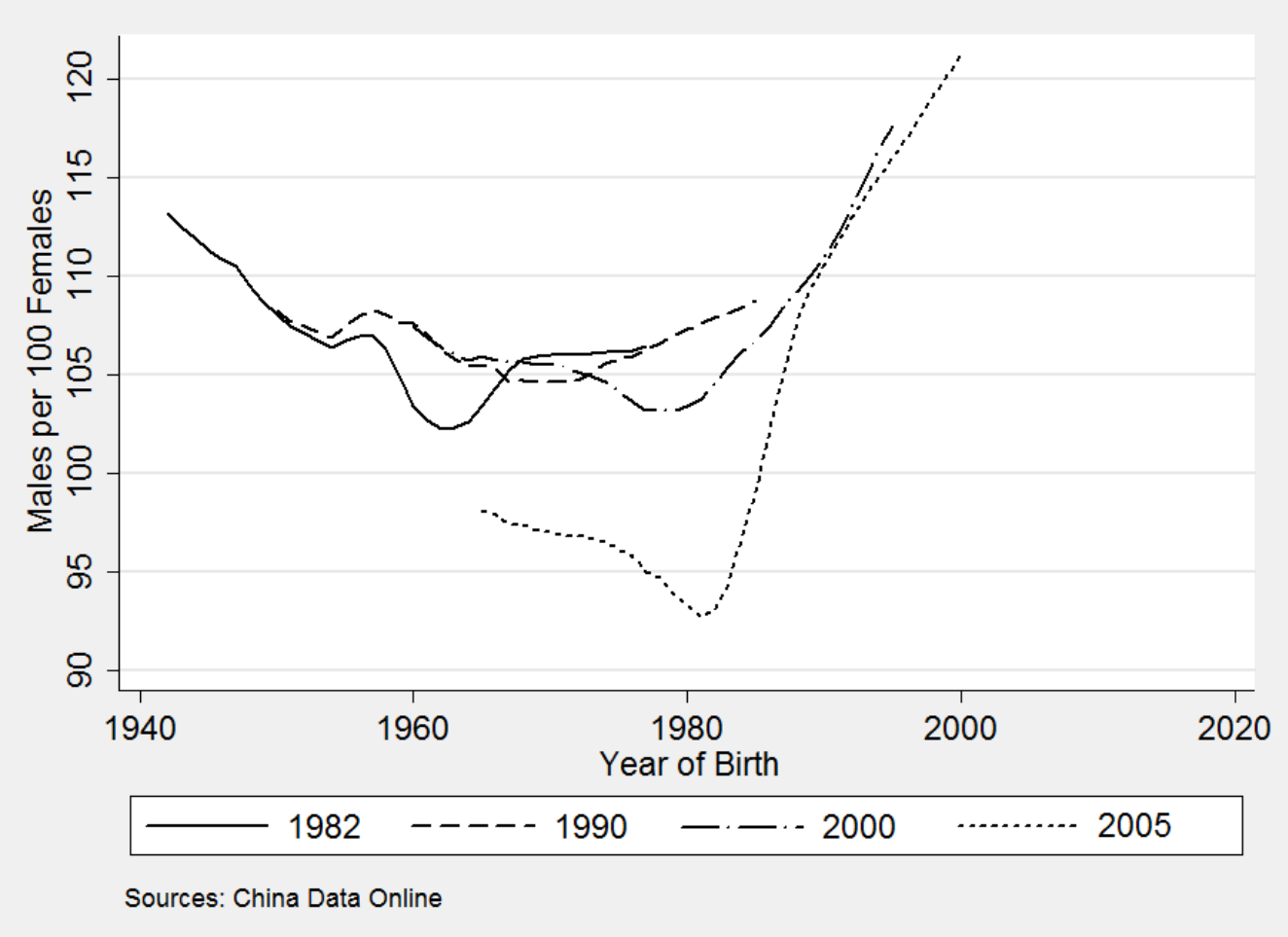
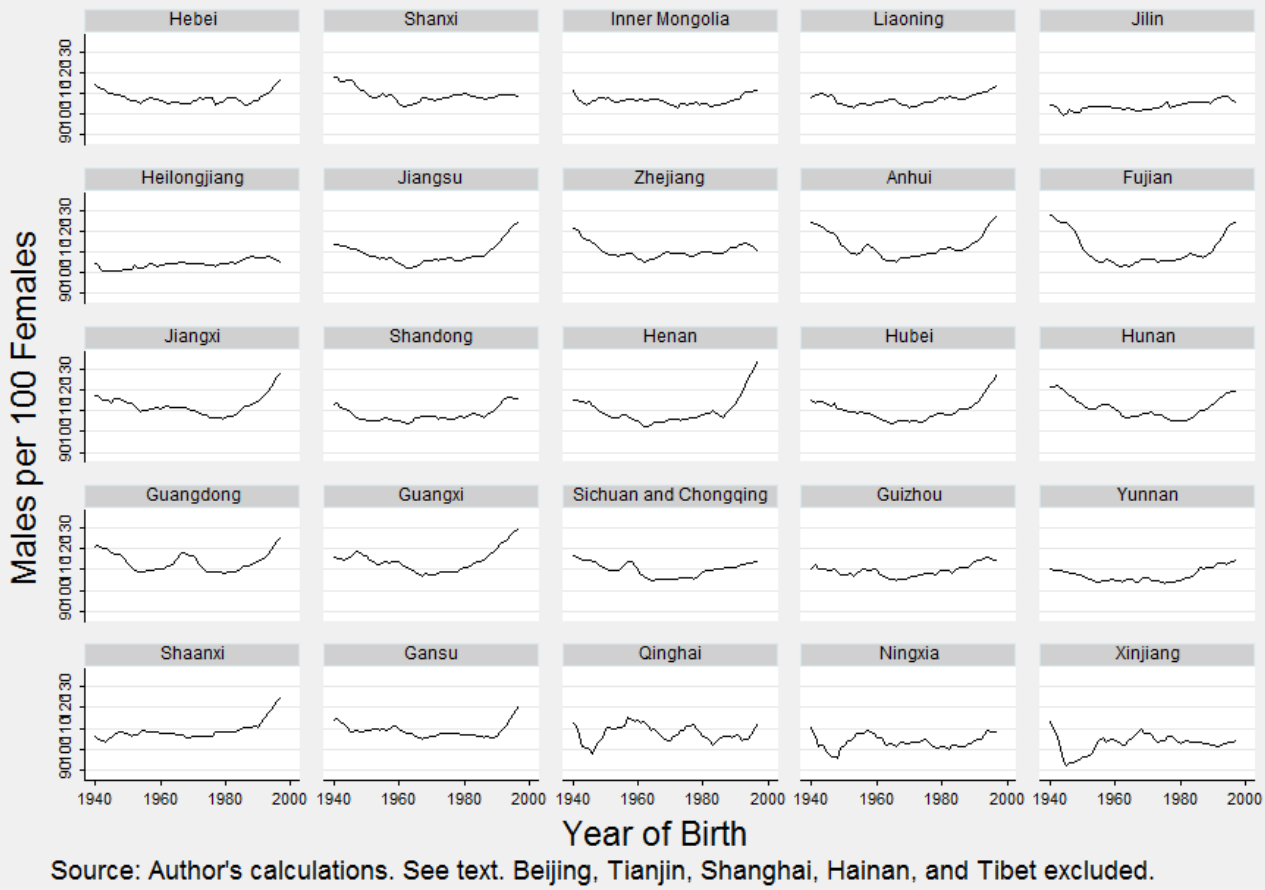
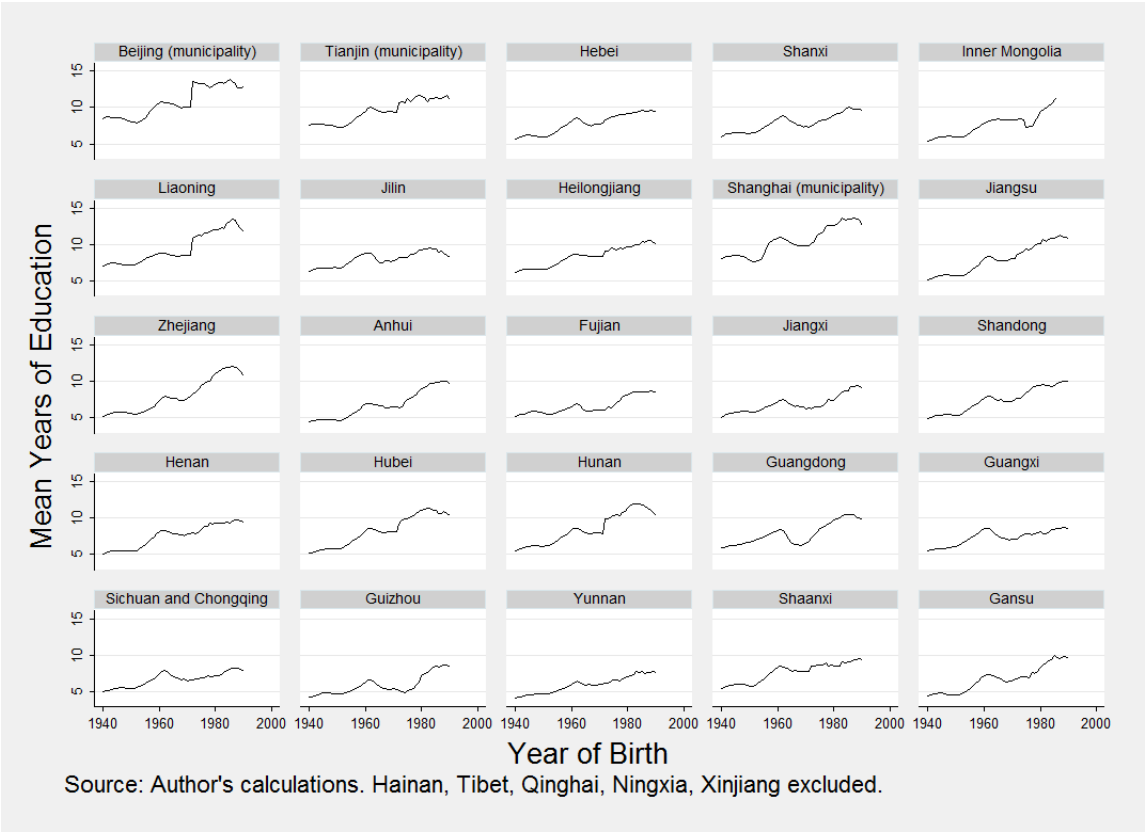


Figure C.2: Adjusted Cohort Sex Ratios by Source Province



Source: Author's calculations. See text. Beijing, Tianjin, Shanghai, Hainan, and Tibet excluded.

Figure C.3: Adjusted Years of Education by Source Province



Source: Author's calculations. Hainan, Tibet, Qinghai, Ningxia, Xinjiang excluded.

C Appendix

C.1 Calculating Implied Personnel Changes

The following section provides more detail behind the calculations of implied and actual military personnel changes in the 1995-2005 period.

The actual change in military personnel is estimated to be 45,000 using figures from the World Bank. (http://data.worldbank.org/indicator/MS.MIL.TOTL.P1)

The implied change in military personnel between the 2000 and 2005 Census is calculated by estimating the "area" between the 2000 and 2005 measurements of birth cohorts. By visual inspection, the area for the 1965-1995 birth cohorts is a discrepancy of about 10 males per 100 females.

Then, an estimate of the total number of females in the 1965-1995 birth year range is determined to be about 353 million by taking cohort estimates from the 1982 and 2010 Census Tabulations. 10% of that figure is about 35.3 million. This magnitude is larger than the actual size of the military.

C.2 The China Population Census

Banister (1984) describes the 1982 Population Census in detail. She notes that the Census was conducted with meticulous care; in fact, the Census was originally slated for 1980, but was implemented in 1982 for a careful pretest of the original questionnaire, adjustments, and extensive pilot censuses nationwide. This particular Census improved upon the 1953 and 1964 Census (the results of which were shrouded in secrecy until the late 1970's) by using definitions of variables like residence, occupation, and employment that were much more consistent with international practice than was the case in the previous Censuses. She also finds that in terms of age-sex population structure the rudimentary censuses of 1953 and 1964 almost matched the quality of the 1982 Census.

The article also finds surprising consistency between the population totals from China's 1982 Census and the permanent population registration system because of known or suspected errors in the registers. Errors in the population registers are a potentially serious problem as the deletion of the names of deceased persons may be delayed because of the family's desire to retain the rations of the deceased. Persons who move to an urban area without government approval may be refused permanent registration in the city, but may stay there for years as temporary residents with their location of permanent registration still in the countryside or nowhere. And, more seriously, as cadres in the 1970's and 1980's have enforced fertility limits through the One-Child policies, many localities have reportedly

refused permanent registration to children whose birth constituted a violation of government restrictions, as cadres were evaluated on population control, and frequently delayed the registration even of children whose birth was authorized.

The article finds some systematic male undercount. Specifically, by analyzing survival ratios of cohorts, the article finds that males with the age range of 15-25 were undercounted in both 1953 and 1964. She finds that the under-reporting of military personnel is likely the cause of this. She also finds that military personnel are missing from the 1982 data. The article also notes a possible overcount of both sexes at ages 49 and 52 in the 1982 Census. It also notes that because only males are undercounted in any of the three censuses, this implies that the actual sex ratio of China's population could have been higher than reported by the 1953 and 1964 censuses.

This section, drawing heavily from Lavely (2001), describes the sampling procedures of the the 2000 China Population Census.

Census sampling procedures are explicitly linked with the household registration (*hukou*) bureaucracy. Theoretically, these registers reflect the official knowledge about residents of the local administrative area and should reflect every change in the composition of a household, whether through births, deaths, marriages, or migration.

In the months before a census, the household registers are updated and verified in a process called "rectification". By the time of the census, the household registers should reflect all information available to local officials about households in the enumeration area. The enumerator then draws upon the household register to construct a list of households and individuals in the enumeration district. This listing includes the name and address of the household head, the number of registered household members, the number of births and deaths in the household in the past year, and the number of registered persons who are absent for less than and for more than six months. Thus at the time of the household interview, the enumerator already has in hand a list of basic information that can be checked against that provided by the household informant. Census procedures also encourage cross-checking between sources, and emphasize that while the household interview is mandatory, it is not to be used as the sole source of information. Thus, it is theoretically possible that for the Chinese Census data to be more accurate than the information provided by the Census respondents themselves.

Lavely (2001) notes that household registration work has been seriously distorted since the 1980's by the presence conflicting incentives from policies such as the "One-Child" fertility policies, and the fact that migration is technically illegal.

Register errors accumulate between censuses. The article notes that in the 2000 Census, pre-census rectification revealed that Wenzhou city had 50,000 deceased persons still on the

registers, and Chongqing Municipality had 130,000. These errors were corrected, but it is far easier to purge registers of the dead than it is to add the living. Children born outside the birth plan are often, despite central government regular regulations, excluded from the register, at least until substantial fines are paid. In Chongqing, rectification work uncovered 68,000 cases, but more could go undetected because both citizens and officials have reason to hide them.

The article also notes that parents conceal excess births to avoid fines, while officials keep them off the books because cadre job evaluations are based primarily on birth planning performance. Census officials tried to remove these disincentives by directing that officials be granted amnesty for previous birth planning falsification as long as they were truthfully reported in the census. But the census is a passing event, while birth planning concerns are perennial; it would only be natural for local officials, fearing a "squaring of accounts after the autumn harvest", to err on the side of discretion.

Lastly, the floating population is not easy to count in the best of circumstances, analogous in difficulty to counting undocumented aliens or the homeless in the U.S. While census workers in the more recent Census waves went to unusual efforts to enumerate urban migrants, including an extensive pre-enumeration of sojourners sleeping outdoors and in public officials, local officials, for various reasons, tried to minimize the population of their administrative domain. Low rates of population growth reflect success in population control, and small population totals boost per capita measures of income and productivity. Thus even as census workers valiantly sought out the floating population by scouring construction sites and searching under bridges. local officials in some places subjected migrants to "census fees", "security fees", and "temporary residence fees" in an attempt to discourage them from being counted.

Indeed, the 2000 Census was dogged by rumors of vast undercounts. Some provinces, such as Henan, Hunan, and Shaanxi, were said in newspaper accounts to have counted millions fewer people than expected. Lavely (2001) warns that users of the data should proceed with caution until data have been more carefully evaluated.

Overall, Lavely (2001) concludes that of the 1982, 1990, and 2000 Censuses, the 1982 wave is the most reliable.

As for the 2010 Census, Wu and He (2015) provides a more detailed description of the questionnaire and the sampling procedures. The articles notes that several measures were adopted to improve the enumeration accuracy of this Census relative to those in 1990, 2000, and 2005, namely, 1) a new enumeration method to account for differences in de jure and de factor residence 2) the use of more detailed ArcGIS mapping technology to map Census tracts, and 3) the inclusion of the foreigner population. The article also emphasizes the

particularly large administrative effort put into the 2010 Census, as it mobilized 6.5 million enumerators if at least junior high education to visit 400 million households within a period of 40-60 days. The article also documents the quality control procedures. As a result of these efforts, the 2010 Census was able to determine a relative undercount in the age 0-9 group in the 2000 Census, and a postenumeration check estimated an underenumeration rate of 0.12 percent, far below the 1.81 percent reported for the 2000 Census.