

## India's Population overtakes China's in April 2023: A debatable conclusion

Padmavathi Srinivasan<sup>1\*</sup> and K Srinivasan<sup>2</sup>

### Abstract

In April of 2023, the United Nations released the World Population Prospects 2022 (WPP), based on actual and adjusted data till 2022 and projecting the future size and distributions of population until 2100 for all the countries. This is a regular annual feature of the United Nations Population Division, but this year in particular, the released data became sensational news because India's population would overtake China's in April 2023. By this time, India's population would be 1428.63 million, 2.9 million more than that of China's which has been seeing declines for over a decade. However, this article argues that the data sets for India and China have a number of known and unknown errors of large magnitudes, especially for China. Official demographic data from China are surrounded by controversies and provoke debates over their quality and reliability. However, despite these concerns, the UN's estimates adjust the Indian data for potential census undercounts but have taken China's at their face value and the UN projections for the country worked out from the officially published figures. The types of errors, and their magnitude, in the population sizes of the base year, fertility and mortality estimates and assumptions are discussed, and it is concluded that based on India's official data, the country's population is set to overtake that of China's only in the year 2026 and the data quality for both countries need to be adjusted differently before making any valid comparisons.

**Key words:** demographic data; UN population projections; India's population overtakes China; quality of population projections; census data accuracy; errors in population projections; reliability of population data in China.

### Introduction

In April 2023, the headlines everywhere screamed that India's population has overtaken that of China, as per the estimations made by the United Nations, and India now is the world's most populous nation, displacing the position held by China from time immemorial. The United Nations annual publication, World Population Prospects 2022, released in April 2023, gave the projected figures of all the countries based on data on population size and distributions, and fertility and mortality trends of the countries available until 2022, and projected them into the future until 2100. These figures indicated that by

1<sup>st</sup> of July 2023, India's population would be 1428.628 million, about 2.9 million people more than that of China making India the most populous nation in the world. The impression created by the sensationalism of the prediction is that India has won a reputed marathon of population numbers to beat China by a close margin, an alleged feat that unfortunately also won plenty of mockery and taunts in the western media (e.g. cartoon by Der Spiegel, April 22, 2023). Closer home, debates rage on what this could mean for India's social and economic development and growth and future place in the world.

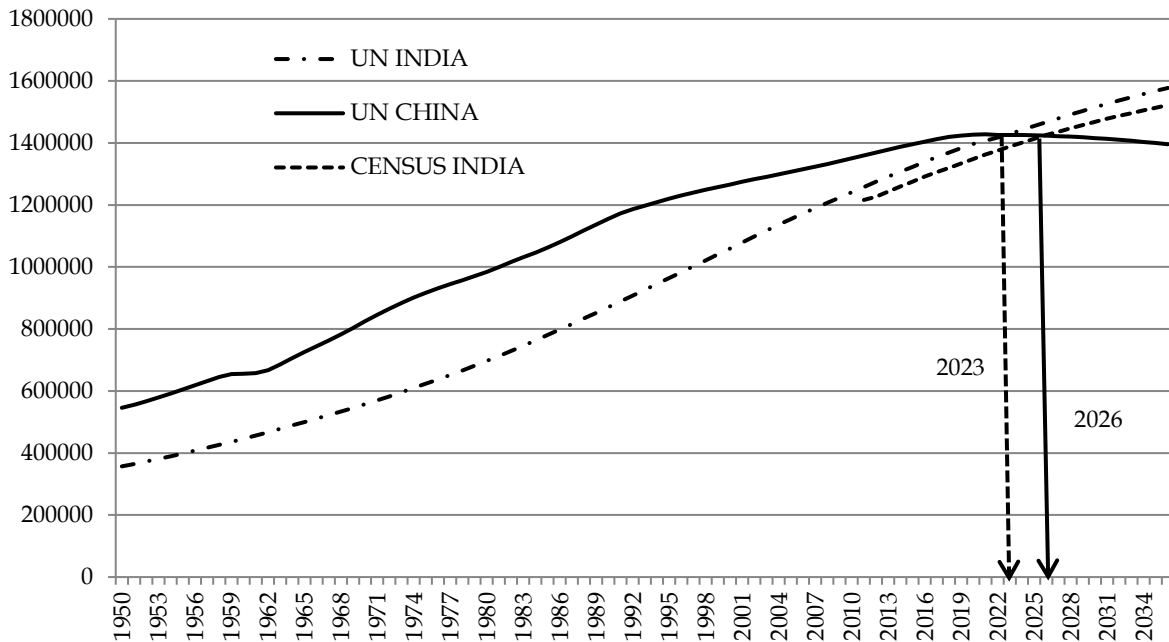
---

\* Corresponding Author

<sup>1</sup>Padmavathi Srinivasan is an independent research consultant in Demography/Population Studies, and holds an MBI and PhD. Email id: [psrinivasan.srinivasan@gmail.com](mailto:psrinivasan.srinivasan@gmail.com).

<sup>2</sup> K Srinivasan is a demographer and the recipient of the 2023 Laureate from the International Union for the Scientific Study of Population (IUSSP). Email id: [ksrini\\_02@yahoo.com](mailto:ksrini_02@yahoo.com).

**Figure 1** UN population estimates for China and India compared with the Census of India's figures



Note: The dashed vertical line indicates when India's population will overtake that of China's, as projected by the UN. The bold vertical line indicates the year when India's population, as projected by the Census of India, may overtake that of China.

All this fanfare however begs a deeper question as to whether the two countries are demographically comparable at all, given the enormous challenges in estimating the current levels of population size and distributions, fertility and mortality levels, and future trends using valid mathematical models to extrapolate to the future. This article argues that the demographic predictions are not based on hard indisputable facts but use data of varying quality from both the countries, and that there are biases in the complex exercise to forecast global trends for the sake of comparisons and finally that there are ambiguities that prevail on the data on population size, and fertility and mortality levels of both the countries. Specifically, I look at the most familiar demographic indicators, the population size, total fertility rate (TFR), and life expectancy at birth (LE) to assess the validity of the projections and predictions made by the United Nations, using data published in the World

Population Prospects for the years from 1950 to 2022 (the year when the most recent UN demographic data for all countries is available).

### Population Size

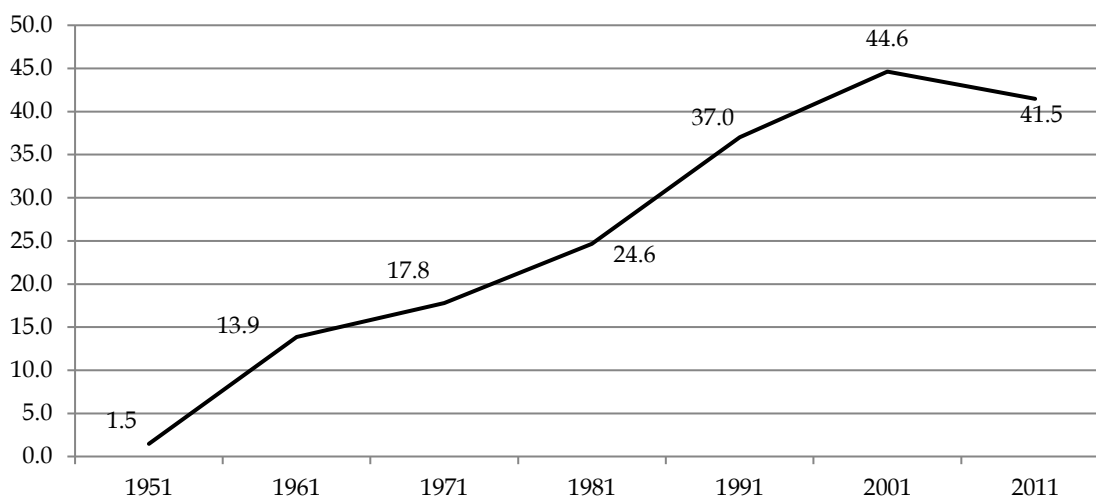
The UN's estimates of the population in India have been consistently on the higher side than official figures published by the Registrar General of India (RGI) since 1951 (Table1).

The gap between the UN population estimates for India and India's own official figures, though significant in all the census years from 1951, widens substantially in 1981 and hits a peak of 44.6 million in 2001 (Figure 2) and dips slightly to around 41.5 million in the year 2011 when India undertook its last population census. The 2021 census in India was postponed because of the pandemic and it is likely to be scheduled for the end of this year.

**Table 1** Population estimates (in 1000s) by the United Nations and official census count by the Registrar General of India (RGI)

India's Census Year	Census count (in '000) as of 1 <sup>st</sup> July*	WPP estimate as on 1 <sup>st</sup> July (in '000)	Excess of UN Pop over Census (in'000)
1951	363,453.86	364,922.36	1,468.50
1961	442,490.27	456,351.88	13,861.61
1971	552,201.81	569,999.18	17,797.37
1981	688,221.87	712,869.30	24,647.43
1991	851,942.72	888,941.76	36,999.04
2001	1,034,341.9	1,078,970.91	44,629.01
2011	1,216,164	1,257,621.19	41,457.19

Source: World Population Prospects 2022 published in 2023. United Nations. \*Census of India data published figures pertain to 1<sup>st</sup> of March of the year and have been extrapolated to 1<sup>st</sup> of July based on the intercensal growth rates to make it comparable to the date of WPP.

**Figure 2** Difference between UN population estimates for India and Census of India's official figures (population in million)

On the other hand, the UN's population estimates for China do not deviate severely from the Chinese census figures. While the UN's estimates 15 million fewer persons than the official count in 1951, when China's first census was undertaken, the UN estimates increase for the successive census years, although they are nowhere as high as those for India (Figure 3). The gap is the widest in 1990, with 22 million more people projected by the UN. For 2020, when China conducted its seventh census, the UN estimates differ from the census figures by only 14.9 million, which is surprising

considering that China undertook its census between November and December 2020 when the Covid-19 outbreak was peaking (Akimov, Gemueva, & Semenova, 2021). A census conducted at the time of utter chaos and disorder, monumental as the one brought on by the novel pandemic Covid-19, would have been extremely challenging, if not near impossible, with respect to logistics, population coverage and quality of responses. One can further assume that compounding this challenge is the sudden and steep rise in the number of deaths from Covid-19 and quick cremations of the

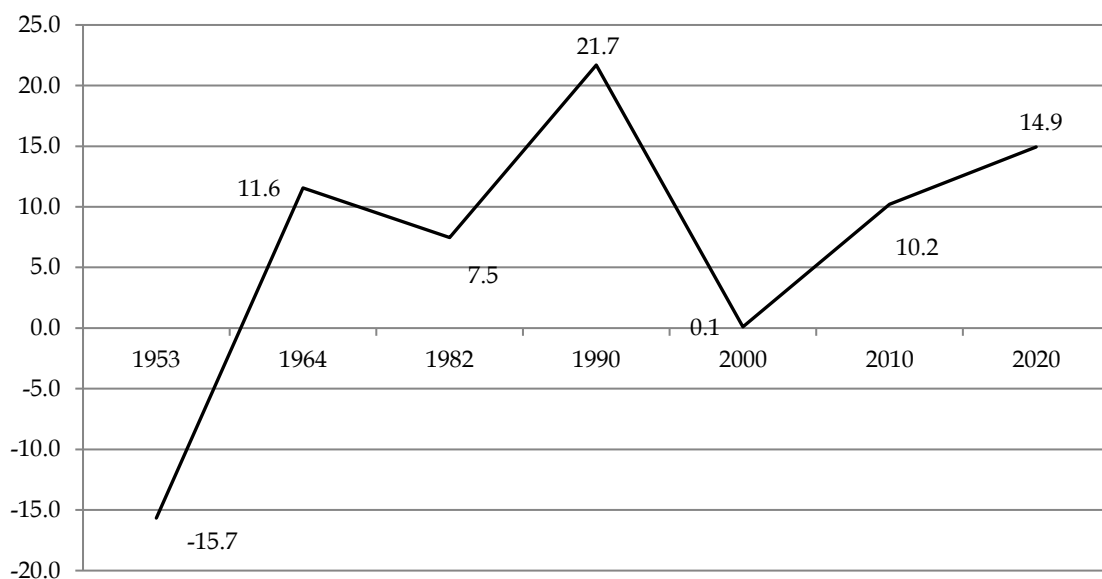
infected dead because of the highly contagious nature of the disease and little understanding of the transmissibility of the virus, which would have made it particularly challenging to do a thorough head count. Yet, the world would learn that China had successfully completed its census that year and the data released some six months later and the UN Population Division used that data.

death toll from the nation-wide famine that occurred in 1958-1961 (Ashton, Hill, Piazza and Zeitz, 1984) since the world would come to know of the extent of devastation from the famine only in the 1980s when demographic data from the census and surveys were made available to public at large. Thus, what happens in China with regard to deaths and population size cannot be taken as true in their face value.

Way back in 1985, the UN projected that China's population in 2020 would be 1,460 million, whereas China's population from its 2020 census was 1,411.78 million, making the UN's projection not that far off the mark from the official figure, even though in 1985 the world outside of China was oblivious to the demographic impact of China's draconian one-child policy on the country's future population growth. Similarly, the UN's 1965 and the 1975 estimated populations for China for the late 1950s and early 1960s do not reflect the enormous

In the case of India, the difference between the UN estimates and Indian official figures is largely attributed to the correction applied by the United Nations for census undercount as enumerated by India from the Post Enumeration Survey (PES) conducted after each census. The post enumeration survey on a sample of households is conducted by the Registrar General of India within a month or two after the census, in order to determine whether there were any omissions or duplications in the census count and the extent of coverage error.

**Figure 3** Difference between UN population estimates for China and census of China's official figures (in million)



Source: UN's World Population Prospects 2022. China's census data is taken from Akimov, A.V., Gemueva, K.A. & Semenova, N.K. (2021). The Seventh Population Census in the PRC: Results and Prospects of the Country's Demographic Development.

After the first PES was conducted following the 1951 census, it was found that there was a net undercount of about 2 percent, leading to a debate within the country as to whether the official census figure should be adjusted for the undercount. It was decided then that no correction would be applied to the official figures for the following reasons: 1) Indian census is essentially a de facto census (or an extended defacto) in which the household member present at any place at the time of sunrise on the 1<sup>st</sup> of March of the census year is counted at that place. The de facto place of residence is ascertained by visiting the household or institution twice: The first visit is within two weeks before the census (the enumeration round) to essentially record the number of persons in the household who are Resident Present (RP), Resident Absent (RA) and Visitor (V). 2) This is followed by a visit one week after 1<sup>st</sup> of March to identify the persons who were actually present in the house at the time of sunrise on the 1<sup>st</sup> March and determine the count of household members<sup>3</sup>. When the post enumeration survey on a sample of household is typically done after a month or two, it is difficult for the household respondent to recall whether or not a household member or visitor was actually present in the house at the time of sunrise on the 1<sup>st</sup> of March. Even if the survey is done within a short time after the census, the recall bias is likely to remain high. The net omission rates for the Indian census from 1981 to 2011 based on post enumeration surveys range from a low of 17.6 per 1000 persons in 1991 to a high of 22.9 per 1000 persons in 2011 (Census of India, 2011). These rates do not indicate severe undercount for a population as large as that

of India. The errors of recall lapse and sampling errors of the correction factors needed seem to override their appropriateness in making adjustments in the census data. Since the census post enumeration survey suffers from the sampling and non-sampling errors as does any sample survey, the post-enumeration surveys are generally used as guidelines for future censuses to improve their coverage, and very few rarely do countries employ the PES to adjust their census counts (United Nations, 2010). Despite the known limitations of the PES, it has been the standard practice of the UN to adjust the population counts for India in their publications, with the figures for India always adjusted upwardly in all the six censuses. As Srinivasan (2017) observes, although there is likely to be some discrepancy between the UN population estimates and India's census counts because of the differences in the months when the data are considered, a four-month difference alone cannot account for the UN's substantial overestimation of India's population after adjusting for census undercount.

### **Total Fertility Rate (TFR)**

The fertility estimations for India as a whole are based on the TFR estimations for each state available from the Sample Registration System (SRS) which is a dual recording system on a large sample of households from each state and in recent years to give estimates for each district and are used in the state projections. The TFR estimations made by the Registrar General of India from 1971 to 2020 and the UN estimates for India and China from 1950 to 2040 are shown below

---

<sup>3</sup> Based on the confirmation from the two visits, the person count is RP+ V if both were present in the household or institution at the time of sunrise on the 1<sup>st</sup> of March.

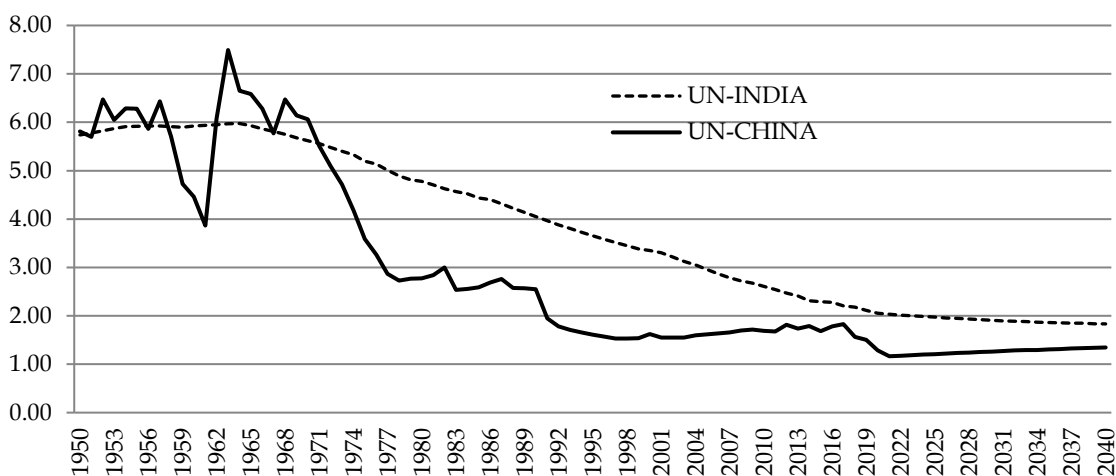
(Figure 4). The UN estimates for India seem to more or less be congruent with India's official estimates. However, in the case of China, the plethora of articles published by demographers in the West between 1990s and 2000s (e.g., Johansson & Nygren, 1991; Greenhalgh, & Li, 1995; Coale & Banister, 1996; Cai & Lavelly, 2003) on the extent of 'missing' female children in China and imbalances in the sex ratios because of the government's one-child policy adopted in 1970 points to the murkiness of China's official data. The one-child policy disallowed couples from having a second child, except in rural areas if the first born was a female, and couples defying the policy would either not report the second birth or, if the second baby was also a female, adopt out the child illegally. For much of their growing years, these 'undocumented' children were kept under the radar of the authorities by the parents for the fear of inviting state penalties and repercussions. As a result, these 'invisible' children were ineligible or unable to receive government benefits and support that all registered citizens are entitled to receive, including access to education and health care. Without the legal status, these

children as adults found it difficult to find work, move within and outside the country, and register their marriage, and any child of theirs was also unregistered (South China Morning Post, 2016). Efforts by demographers to account for these missing populations by analyzing subsequent Chinese censuses continued even several decades after the implementation of the one child policy. It is likely that this group who came of working age during China's peak of frenzied economic growth from the late 1990s formed a large proportion of the silent non-contractual and undocumented labor force that contributed to country's rapid economic development.

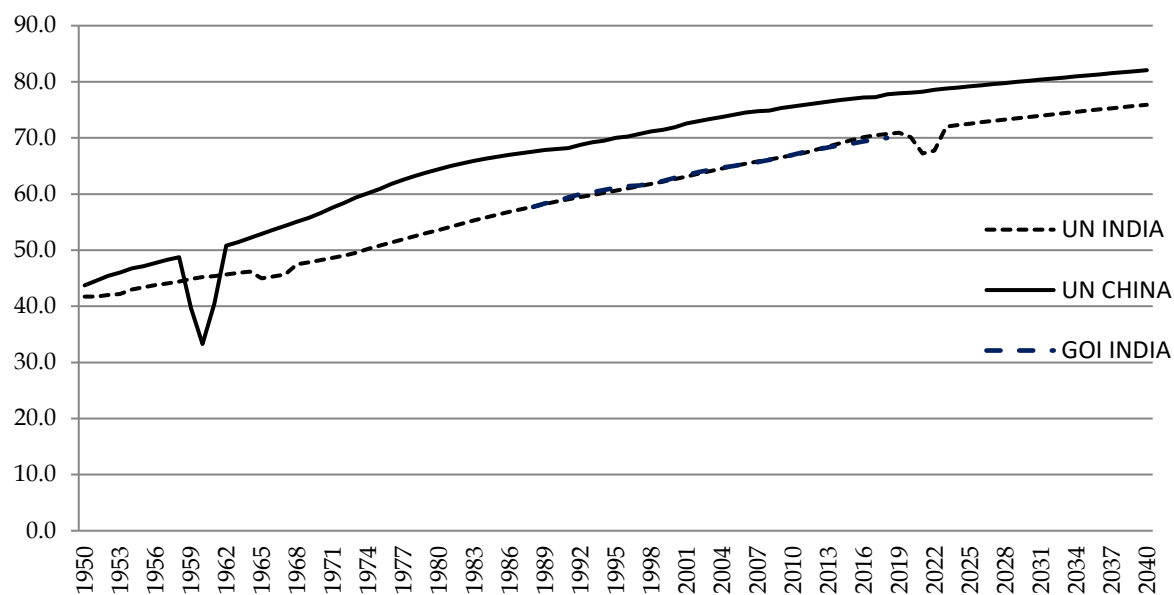
**Life expectancy at birth**

The most glaring difference in the United Nations projections is in the assessment of mortality rates, specifically the life expectancy at birth, a single yet strong index that describes the prevailing mortality conditions in a country. The United Nation's estimations show a sharp drop in the life expectancy at birth for India by three years between 2020 and 2021, the period of Covid-19 pandemic (Figure 5).

**Figure 4** Total Fertility Rate - Government of India's estimates and the United Nations estimates for India and China



**Figure 5** United Nation's estimates for expectation life at birth for India and China, 1950 to 2024, and India's official estimates from 1970 to 2020



Data: (i) World Population Prospects 2022, published in 2023.

(ii) Sample Registration System (SRS) based Abridged Life Tables 2016-2020, Office of the Registrar General & Census Commissioner, Government of India, released 2022. The figures pertaining to a particular year is the 5-year average.

This mortality adjustment for India has been done even though the annual number of deaths for the period 2020-2022 for the country as a whole are yet to be fully collated from the death rates compiled by the concerned statistical departments in each state in India. We know however from a preliminary examination of the registered deaths from the civil registration system for all the districts in Kerala (which has near-100 percent coverage) that the crude death rate for the state remained more or less unaffected between 2019 and 2020 and rose from around 7 per 1000 population in the year 2020 to only 8 per 1000 population in 2021, the period when the Delta variant, the most contagious and lethal of the variants of the Covid-19 virus, surfaced in the country.

By contrast, it would seem that the actual number of Covid-19 related deaths in China have not been evaluated sufficiently by the UN and appropriate adjustments made

factoring in the excess deaths from the pandemic in estimating the life expectancy in China for the pandemic period. China has not made public, at least not known outside of the country, the extent of casualties suffered during the pandemic, and the source of data which the UN relies on for its projections are based mostly on China's official mortality rates and trends from the pre-Covid era. The possibility of the demographic figures published officially by China are being either understated or overstated and this needs to be investigated by demographers when making mortality projections. For example, it was alleged (e.g, TIME, 2020) that during the Covid-19 pandemic the number of deaths in China were far more than what the government would have the world believe. These allegations are trifle hard to dismiss, because when looking back on China's modern history, the severity of the famine that plagued the entire country from the late

1950s to early 1960s as a result of Chairman Mao Zedong's aggressive *Great Leap Forward* economic policies was not known to the world until the 1980s, when China's census and demographic surveys in the 1980s became available to international demographers. Using these data, it was estimated that the famine might have caused at least 30 to 40 million deaths (Ashton, Hill, Piazza & Zeitz, 1984; Yushi, 2014). As Smil (1999) observes, 'beyond a narrow circle of China experts, the famine has also been virtually ignored by Western scholars and politicians' (p 1620, para 5).

### Conclusion

It is no small feat to work out the population size, age-sex distributions, and fertility and mortality levels of a country over a period of time in the recent past with minimal errors for any developing country, make realistic assumptions on the future trends in mortality, fertility and migration and carry out projections for the future years up to 2100, and that too for 200 countries as the United Nations Population Division (UNPD) has done year after year since the 1950s, their latest being the World Population Prospects for 2022 released in April 2023. The UNPD's efforts and contributions in this regard are to be commended. However, the UN's disregard for the differential quality of data among the developing countries, even after making the necessary adjustments for each country, for making valid cross country comparisons can be called into question, It is well known among statisticians that when a any parameter of variable X with an error of 'E1' is corrected by addition or multiplication by another variable Y (a correction factor) with error 'E2', the corrected variable has an error more than

the sum of the two errors, depending on the method of correction adopted. Every adjustment for data quality generates its own error compounding with the inherent error in the data and in the adjustment factor.

In India and other democratic countries in the developing world with open governments, surveys and studies can be undertaken independently with funding from both the governmental and non-government agencies and the government data can be liberally subject to scrutiny by non-governmental agencies for coverage and omissions. Corrections on the published data are possible, of course, subject to known or to be computed quantum of errors. On the other hand, all published demographic, health, social and economic data in totalitarian countries as China come only from the government sources and independent validations are not permitted. Sadly, China's official and demographic data have always been subject to intense scrutiny by international demographers and scientists (e.g. Retherford et al, 2005; Zhang & Zhao, 2006) and have been found to be loaded with large errors that are more of purposeful biases than random errors found in democratic countries, and surrounded by controversies for possible manipulations, suppression, and revision of true figures often to suit political narratives surface (Lowsen, 2022). Recently, a Chinese scientist, Yi Fuxian, working at the University of Wisconsin at Madison made a startling revelation that the true population count in China from the 2020 census is likely to be considerably lower than the number officially released by the Chinese government.<sup>4</sup> By his reasoning, India's

---

<sup>4</sup>see i) China's Demographic Manipulation by Yi Fuxian published in Project Syndicate on August 5, 2021. <https://www.project-syndicate.org/commentary/china-2020-census-inflates-population-figures-downplays-demographic-challenge-by-yi-fuxian-2021-08>.



population (based on United Nation's projections) would have overtaken that of China's much earlier than the year 2023. On the other hand, if you take the official population figures for India as done for China for the year 2023, India's population is still lower than that of China and the country will overtake China only by 2026. While Yi Fuxian's assertions may or may not be correct, it would be presumptuous on our part to expect that the projections and predictions made for a country will reflect the reality without a sound and reliable population related data with standard errors attached to them. Even careful mathematical modeling for estimation of Covid-19 deaths between 2020 and 2022 have failed to get at the truth even in developed countries. To illustrate, a recent evaluation of the mathematical model used by the WHO to estimate the excess deaths from Covid-19 around the world (Verma, Nath & Khan, 2023) shows that the predictive ability of the model was successful for only a handful of countries while for the remaining countries, including India, the model failed to be a reliable and useful tool.

A demographically and culturally most diverse and complex country such as India has enormous challenges collecting, collating and analyzing data from the different corners of the country. It would be naïve on a demographer's part to expect accurate data at all times, even in developed countries. Still, India does its best on data collection and public scrutiny; it has a rigorous census data collection and verification system, and importantly an open system willing to collaborate with non-

governmental and international agencies in the process of data collection, analyses and dissemination, and importantly, sharing of information. In addition to the nation-wide decadal census as a source of demographic data, there are the post-enumeration surveys, on-going sample registration system to obtain independent and reliable estimates of birth, death and infant mortality rates, in addition to the civil registration system, by states and even districts. There are also a number of national level surveys such as the National Family Health Surveys (NFHS), with the sixth round currently underway, and the National Sample Surveys to serve as independent checks on all demographic and economic parameters. Any demographic estimates for China, on the other hand, will have to, at best, be treated as crude approximations with unknown biases 'because they rely entirely on base population and vital statistic figures given by the Chinese government. The population figures provided by the censuses in China are known to have large reporting errors that are detected only after many years, often after more than a decade. As discussed earlier, the *Great Chinese Famine* of 1958 to 1961 resulted in an extreme loss of life (death estimates given by various scholars and journalists range from 16 million to more than 30 million) and led to an interruption in the lives for millions of survivors. Such a huge loss of life came to the lime light only after a decade. In fairness, one can postulate that the actual death toll in China from the Covid-19 pandemic will come to surface only a decade or two from now, leaving the international agencies responsible for data collection and

---

ii) South China Morning Post. China's population numbers are almost certainly inflated to hide the harmful legacy of its family planning policy, by Yi Fuxian. July 20, 2019 <https://www.scmp.com/comment/opinion/article/3018829/chinas-population-numbers-are-almost-certainly-inflated-hide>

dissemination scrambling to revise the figures for China retrospectively. In conclusion, it does seem that the population figures of India and China are not at all comparable. It is akin to comparing apples and oranges, and therefore, this procedure has to be revisited by the UN's Population Division.

## Reference

- Akimov, A.V., Gemueva, K.A. & Semenova, N.K. (2021). The Seventh Population Census in the PRC: Results and Prospects of the Country's Demographic Development. *Her. Russ. Acad. Sci.* **91**, 724-735.  
<https://doi.org/10.1134/S1019331621060083>.
- Ashton, B., Hill, K., Piazza, A., & Zeitz, R. (1984). Famine in China, 1958-61. *Population and Development Review*, *10*(4), 613-645.  
<https://doi.org/10.2307/1973284>.
- Cai, Y., & Lavelly, W. (2003). China's Missing Girls: Numerical Estimates and Effects on Population Growth. *China Review*, *3*(2), 13-29.  
<http://www.jstor.org/stable/23461902>.
- Coale, A. J., & Banister, J. (1996). Five Decades of Missing Females in China. *Proceedings of the American Philosophical Society*, *140*(4), 421-450.  
<http://www.jstor.org/stable/987286>.
- Feeney, G., Wang, F., Zhou, M., & Xiao, B. (1989). Recent Fertility Dynamics in China: Results from the 1987 One Percent Population Survey. *Population and Development Review*, *15*(2), 297-322  
<https://doi.org/10.2307/1973706>.
- Government of India. (2014). Census of India 2011: Report on post enumeration survey.
- Greenhalgh, S., & Li, J. (1995). Engendering Reproductive Policy and Practice in Peasant China: For a Feminist Demography of Reproduction. *Signs*, *20*(3), 601-641  
<http://www.jstor.org/stable/3174835>.
- Johansson, S., & Nygren, O. (1991). The Missing Girls of China: A New Demographic Account. *Population and Development Review*, *17*(1), 35-51.  
<https://doi.org/10.2307/1972351>.
- Retherford, R. D., Choe, M. K., Chen, J., Xiru, L., & Hongyan, C. (2005). How Far Has Fertility in China Really Declined? *Population and Development Review*, *31*(1), 57-84  
<http://www.jstor.org/stable/3401438>.
- Smil V. (1999). China's great famine: 40 years later. *BMJ (Clinical research ed.)*, *319*(7225), 1619-1621  
<https://doi.org/10.1136/bmj.319.7225.1619>.
- Srinivasan, K. (2017). Population concerns in India: Shifting trends, policies, and programs. Sage publications. New Delhi.
- Thomson Reuters Foundation. (2016). The battle of China's invisible children, victims of the one-child policy, to recover lost years. South Morning China Post. Published on 26 Dec, 2016  
<https://scmp.com/lifestyle/families/article/2056885/battle-chinas-invisible-children-victims-one-child-policy-recover>.
- TIME. (2020) Report of Urns Stacked at Wuhan Funeral Homes Raises Questions About the Real Coronavirus Death Toll in China, APRIL 3  
<https://time.com/5811222/wuhan-coronavirus-death-toll/>.
- United Nations. (2010). Post enumeration surveys: Operational guidelines. Technical report. Statistical Division. United Nations Secretariat. New York.
- Verma, V., Nath, D. C., & Khan, H. T. A. (2023). Evaluation of concordance in estimation of excess mortality due to COVID-19 pandemic. *Journal of evaluation in clinical practice*, *10.1111/jep.13866*  
<https://doi.org/10.1111/jep.13866>.
- Yushi, M. (2014). Lessons from China's Great Famine. *Cato Journal*, Vol. 34, No. 3 (Fall 2014).
- Zhang, G., & Zhao, Z. (2006). Reexamining China's Fertility Puzzle: Data Collection and Quality over the Last Two Decades. *Population and Development Review*, *32*(2), 293-321.  
<http://www.jstor.org/stable/20058875>