Uncertainty in Population Projections for Statistically Less Developed Countries: A Discussion of Estimation Issues Using Data for Four South Asian Countries.

Extended abstract prepared for the 2010 meetings of the Population Association of America

Thomas McDevitt, U.S. Census Bureau

Population projections are inherently uncertain because the future is unknown and, particularly for the more distant future, largely unknowable. Even so, demographers are sometimes asked to assign confidence intervals to their projections to help users, often non-demographers, better understand the likely range of possible future population size, growth, and composition. A decade ago, Ron Lee (1999:156) began a discussion of probabilistic population forecasting with the statement "It is generally agreed that demographers have a responsibility to indicate how certain or uncertain their forecasts may be" and went on to propose probabilistic forecasting as a preferred strategy for assigning confidence to projections. Other demographers have disagreed, not with the value of admitting uncertainty, but with the argument that probabilistic projections are the best solution to the problem (Heilig 1996). In general, demographers recognize that uncertainty arises from measurement error, year-to-year variation in demographic events (statistical variance), and basic estimation decisions made by the analyst, which give rise to possible bias in estimated levels of projection parameters. Uncertainty also arises from the trend modeling process; unforeseen catastrophic events; policy and program impacts on fertility, mortality, and migration; and structural changes in society (such as rising female labor force participation, Lee (1999:158-159)). However, the literature on uncertainty in population projections has tended to focus on using the historical time series for a parameter to estimate the stochastic component of variation in the parameter – i.e., stochastic population forecasting – on the one hand and, on the other, on using the range of expert opinion about the likelihood of future levels in the components of change to calculate confidence intervals (Lee 1992, 1999; Lutz, Sanderson and Scherbov 1996, 1999). The distinction between, and relative importance of, the stochastic and bias elements of uncertainty has received much less attention (the 2009 PAA paper by Alkema et al. being an exception).

This poster focuses on the uncertainty in projected fertility for four South Asian countries – Pakistan, India, Bangladesh, and Burma – distinguishing variability attributable to year-to-year variation in estimates (variance) from estimation decision-based variation (bias). The work draws upon the international demographic estimates and projections of the U.S. Census Bureau and the United Nations Population Division.

Countries

Pakistan. Pakistan conducted censuses in 1972, 1981, and 1998. However, the country's demographic profile is best understood from data collected through a series of nationally representative household surveys. Estimated TFRs from demographic surveys uniformly indicate a slow decline in fertility in the country, at least up until the current time period. Direct and indirect estimates of TFR are not entirely consistent.

India. India conducts censuses every ten years, fields a Sample Registration System (SRS) to track fertility and mortality on an ongoing basis, and also conducts National Family Health Surveys (NFHS), providing additional demographic data for the country. However, estimated fertility from the SRS and NFHS are not entirely consistent; Retherford and Mishra (2001) have argued that the SRS may be providing somewhat better estimates of fertility than the NFHS.

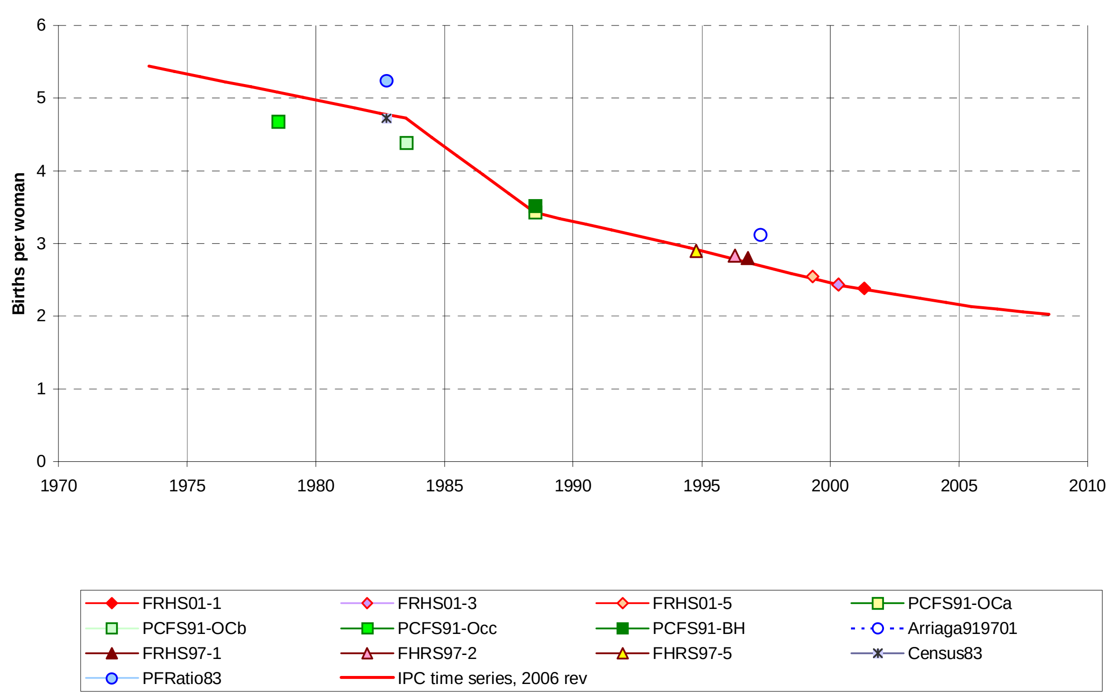
Bangladesh. Bangladesh conducted six demographic surveys over the period 1989-2007, beginning with the 1989 Bangladesh Fertility Survey and following with five Demographic and Health Surveys. Direct, birth history-based estimates and indirect, Brass P/F ratio and Arriaga estimates are available from these surveys. The demographer must choose from a slightly higher series of indirect estimates and a slightly lower set of direct estimates, both downward-trending.

Burma. Burma conducted censuses in 1973 and 1983. Some uncertainty is associated with the levels of fertility and mortality in the country from the last census to the present. In the case of fertility, this uncertainty stems from the availability of multiple measures from demographic surveys conducted in 1991, 1997, and 2001 and the lack of a recent census against which to compare projections using alternate fertility scenarios.

Illustration Using Burma

Both the probabilistic and "high"-"low" scenario approaches to describing uncertainty in projected fertility level involve the juxtaposition of (1) a set of historical "observations" or, in the United Nations' terminology, estimates and (2) a range of future possibilities. These are either a probabilistic-based range of future levels or a range of possible future values based on two or more alternative scenarios derived for analytical purposes.

Figure 1. Estimates of TFR, multiple methods and sources: Burma



In fact, the historical/estimation part of the time continuum is comprised of an underlying trend, normal variability in values around that trend, measurement error associated with individual data points, and the choice of estimation technique. The combination of these four sources of difference in estimated fertility level is illustrated in the first chart for Burma. This chart shows thirteen estimates of total fertility rate – direct estimates based on census questions about births in the 12 months prior to a census or from birth history data, indirect estimates derived using the own-children technique, and indirect estimates derived using Brass and Arriaga techniques – plus a trendline adopted by the Census Bureau in its 2006 revision for Burma.

Fertility estimates shown in Figure 1.

FRHS01-1, FRHS01-3 and FRHS01-5 are one-, three-, and 5-year average birth history-based estimates of TFR from the 2001 Fertility and Reproductive Health Survey.

PCFS91-Oca, PCFS91-Ocb, and PCFS91-Occ are three own-children estimates of TFR from the 1991 Population Changes and Fertility Survey (PCFS)

PCFS91-BH is a birth history-based estimate of TFR from the 1991 Population Changes and Fertility Survey

Arriaga919701 is the middle estimate of three TFRs calculated using Arriaga's ARFE-3 indirect estimation technique and data on cumulative and current fertility from the 1991 PCFS, the 1997 Fertility and Reproductive Health Survey, and the 2001 FRHS. Only the middle estimate is used.

FRHS97-1, FRHS97-2, and FHRS97-5 are one-, two-, and 5-year average birth history-based estimates of TFR from the 1997 Fertility and Reproductive Health Survey.

Census83 is a direct estimate of TFR from the 1983 census, using a question about births in a reference period preceding the census.

PFRatio83 is an indirect estimate of TFR using Brass' P/F ratio method and cumulative and current fertility from the 1983 census.

The IPC time series is from the 2006 revision, the current revision reflected in the Census Bureau's International Data Base (December 2009).

Separate logistic curves fitted to the direct, indirect (Brass/Arriaga), and indirect (own-children) point estimates are consistent in indicating falling TFR from the late 1970s or early 1980s to the early 2000s. Fitted curves with a lower asymptote of 2.0 are shown in Figure 2. However, for ten to twenty years after the latest available point estimate, projected fertility could be as high as that indicated by the indirect (Brass/Arriaga) series or as low as that indicated by the direct or own-children estimates. This spread in estimates represents uncertainty in projected population beyond the 2000-05 period. In addition, however, the spread of point estimates used to fit the curves indicates some additional, variance-based uncertainty from about 2005 onward.

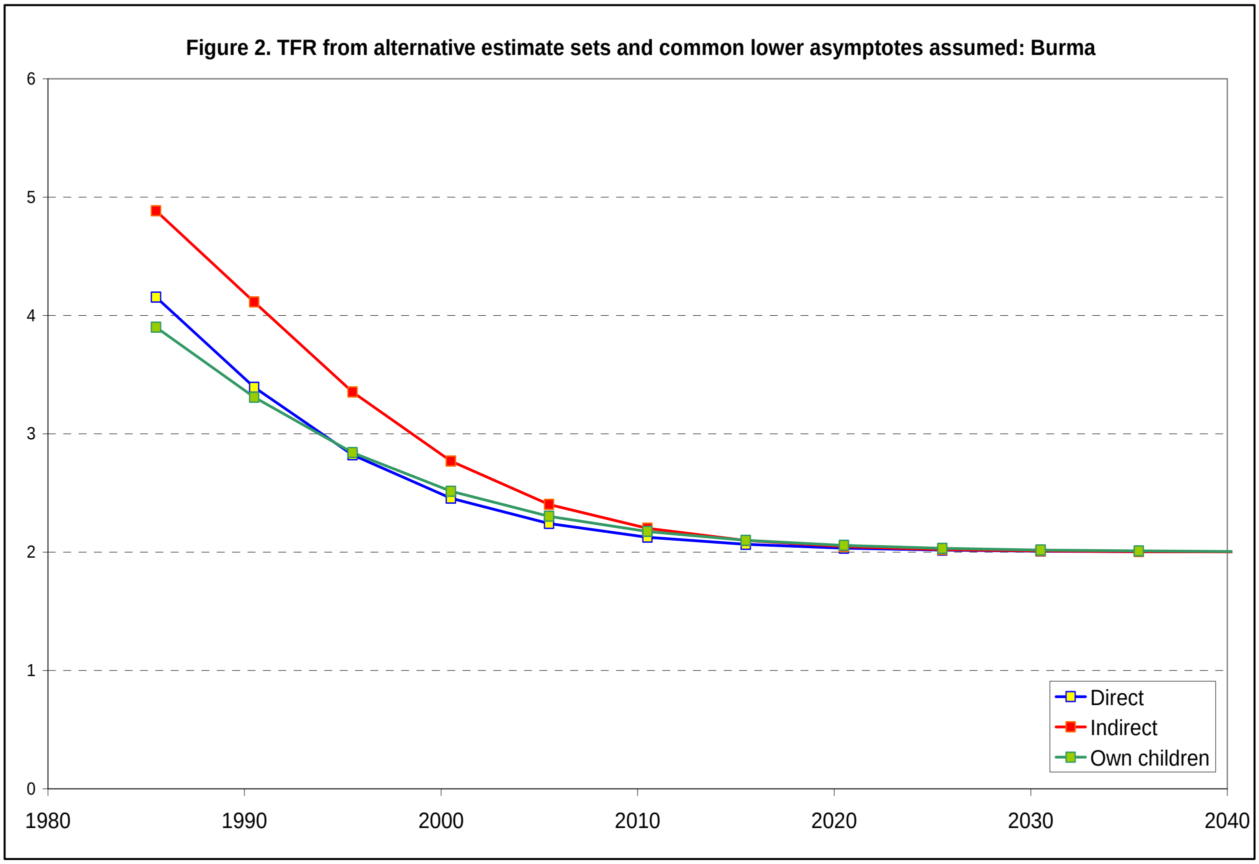


Figure 2. TFR from alternative estimate sets and common lower asymptotes assumed: Burma
6
5
4
3
2
1
0
1980
1990
2000
2010
2020
2030
2040
Direct
Indirect
Own children

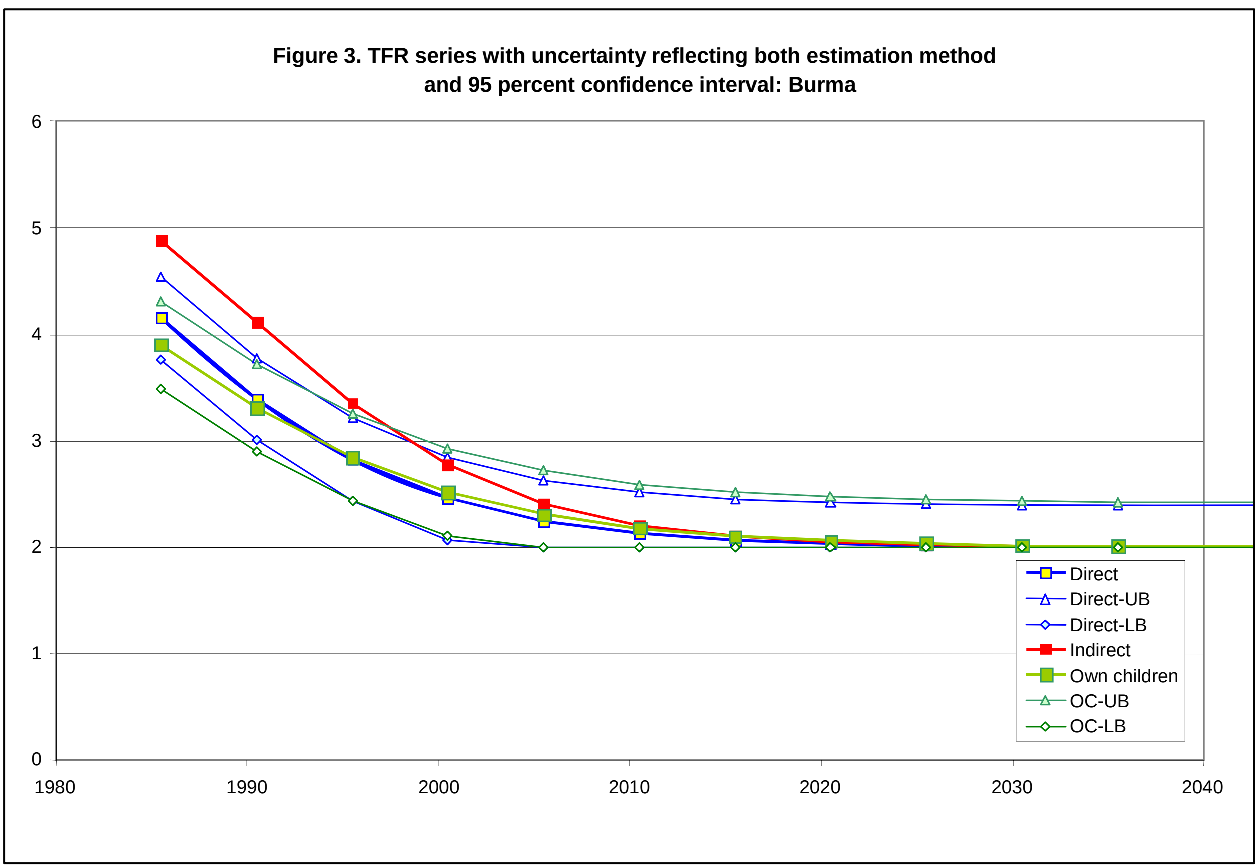


Figure 3. TFR series with uncertainty reflecting both estimation method
and 95 percent confidence interval: Burma
6
5
4
3
2
1
0
1980
1990
2000
2010
2020
2030
2040
Direct
Direct-UB
Direct-LB
Indirect
Own children
OC-UB
OC-LB

Figure 3 shows the three logistic time series with 95 percent confidence intervals based on the procedure for calculating variance proposed by Li, Lee and Tuljapurkar (2002). Over the intermediate term, from 2010 to 2020, method-specific uncertainty (measured as the difference between the indirect and own-children TFR estimates) accounts for a small part -- between 7 and 13 percent -- of the overall range of estimates shown (upper bound, 95 percent confidence interval for the indirect estimate less lower bound, direct and own-children-based TFR estimates).

Range in TFR estimates for Burma, calculated and compared with the UN Population Division range (high minus low)

Year	Constrained methods-based estimates range		Constrained overall range of estimates		Percentage of overall range accounted for by method	Year	UN variants range
2010	0.07	2.13 to 2.20	0.59	2.00 to 2.59	13		
2015	0.07	2.07 to 2.10	0.51	2.00 to 2.51	15	2013	0.50
2020	0.03	2.03 to 2.06	0.47	2.00 to 2.47	7	2018	0.80
2025	0.02	2.01 to 2.03	0.44	2.00 to 2.44	5	2023	1.00

The poster is expected to conclude:

- First, for statistically less developed countries the bias component of uncertainty may be a substantial component of uncertainty in future population size, fertility and mortality level, and projected population composition. Burma illustrates the case where the bias component is less important.
- Second, estimated uncertainty depends on the selection of fertility estimates used. Direct and indirect estimates will differ, but so will the variation in the estimate set derived using a specific class of estimation techniques. During a period of falling fertility, some indirect fertility estimation techniques that assume stable fertility levels will overstate fertility levels. Direct estimation, on the other hand, may understate fertility levels for reasons having to do with reference period error and displacement of births into the past by respondents. The direction of bias in estimates is recognized by demographers but students of forecasting should also recognize that restricting an estimate set to a single, consistent class of estimates forces the analyst to make a subjective judgment about which time series is most plausible, which carries the least bias.
- Uncertainty associated with fertility levels may be less than that implied by the spread between the United Nations Population Division's "high" and "low" fertility scenarios (a total difference of one birth) for longer projection intervals.

References

Alkema, Leontine, Adrian Raftery, Patrick Gerland, and Francois Pelletier. 2009. "Estimating the Total Fertility Rate from Multiple Imperfect Data Sources and Assessing Its Uncertainty." Paper presented at the 2009 annual meetings of the Population Association of America, April 30 – May 2, Detroit.

Heilig, Gerhard K. 1996. "World Population Prospects: Analyzing the 1996 UN Population Projections." IIASA LUC-Project WP-96-146. Available from the Internet at <http://www.iiasa.ac.at/Publications/Documents/WP-96-146.html/>.

Lee, Ronald D. 1992. "Stochastic demographic forecasting," ***International Journal of Forecasting***, 8:315-328.

_____1999. "Probabilistic Approaches to Population Forecasting," in Wolfgang Lutz, James W. Vaupel, and Dennis A. Ahlburg, eds., ***Frontiers of Population Forecasting***. Supplement to volume 24, ***Population and Development Review***, pp. 156-190.

Li, Nan, Ronald Lee and Shripad Tuljapurkar. 2002. "Using the Lee-Carter Method to Forecast Mortality for Populations with Limited Data." Unpublished paper, photocopy. Subsequently published in International Statistical Review, 2004, 72(1):19-36.

Lutz, Wolfgang, Warren C. Sanderson, and Sergei Scherbov. 1996. "Probabilistic Population Projections Based on Expert Opinion," in Lutz, W., ed., ***The Future Population of the World. What Can We Assume Today?*** Revised edition. London: Earthscan Publications.

_____. 1999. "Expert-based Probabilistic Population Projections," in Wolfgang Lutz, James W. Vaupel, and Dennis A. Ahlburg, eds., ***Frontiers of Population Forecasting***. Supplement to volume 24, ***Population and Development Review***, pp.139-155.

Retherford, Robert and Vinod K. Mishra. 2001. "Comparison of Fertility Trends estimated from India's Sample Registration System and India's First and Second National Family Health Surveys." Paper presented at the annual meeting of the Population Association of America, Washington D.C., March 29-31.