International Institute for
Applied Systems Analysis
Schlossplatz 1
A-2361 Laxenburg, Austria

Tel: +43 2236 807 342
Fax: +43 2236 71313
E-mail: publications@iiasa.ac.at
Web: www.iiasa.ac.at

Interim Report **IR-03-014**

Toward Structural and Argument-Based Probabilistic Population Projections in Asia: Endogenizing the Education-Fertility Links

Wolfgang Lutz (lutz@iiasa.ac.at)
Sergei Scherbov (sergei.scherbov@oeaw.ac.at)

Approved by

Leen Hordijk (hordijk@iiasa.ac.at)
Director

April 7, 2003

Contents

Abstract

Among the different sources of uncertainty in population forecasting, uncertain changes in the structure of heterogeneous populations have received little attention so far, although they can have significant impacts. Here we focus on the effect of changes in the educational composition on the overall fertility of the population in the presence of strong fertility differentials. With data from India we show that alternative paths of future female enrolment in education result in significantly different total fertility rates (TFR) for the country over the coming decades, even assuming identical fertility trends within each education group. These results from multi-state population projections by education are then translated into a fully probabilistic population projection for India in which the results of alternative education scenarios are assumed to expand the uncertainty range of the future TFR in the total population.

This first attempt to endogenize structural change with respect to education – which is the greatest measurable source of fertility heterogeneity in Asia – has resulted from a larger exercise of the Asian MetaCentre for Population and Sustainable Development Analysis to collect empirical information, scientific arguments as well as personal, informed opinions about likely future population trends in Asia from a large number of population experts in the region. The paper summarizes this exercise, in which individual experts were also asked to give their views on the uncertainty ranges of future demographic rates. The paper also describes the errors of past population projections in the region and summarizes the main findings from in-depth interviews of these experts, resulting in the conclusion that education trends are a main determinant of future fertility, thus providing the reason for this attempt to endogenize the education-fertility link in probabilistic population forecasts.

Acknowledgments

This work has been conducted in the context of two workshops organized by the Asian MetaCentre for Population and Sustainable Development Analysis (Singapore) and funded by the Wellcome Trust (UK). The important substantive input of all the participants of these workshops is gratefully acknowledged.

About the Authors

Wolfgang Lutz is the Leader of the Population Project at IIASA, and Director of the Vienna Institute of Demography of the Austrian Academy of Sciences.

Sergei Scherbov is a Senior Scientist at the Vienna Institute of Demography of the Austrian Academy of Sciences, and a Senior Research Scholar at IIASA.

Toward Structural and Argument-Based Probabilistic Population Projections in Asia: Endogenizing the Education-Fertility Links

Wolfgang Lutz and Sergei Scherbov

Introduction

Asia's population, which comprises more than half of the world's population, is currently going through significant structural changes. While the population of Asia is still expected to grow from currently 3.5 billion to around 4.8 billion in 2050, for the second half of the century we expect the beginning of a population decline (see Lutz et al. forthcoming-A). Simultaneously Asia's population is expected to age dramatically, with the proportion above age 60 expected to increase from currently 9 percent to some 23 percent by the middle of the century and more than 35 percent toward the end of the century. But there are very significant regional differences in Asia both with respect to expected future population growth and with respect to the speed of aging. While in 2050 we expect almost 40 percent of the population of Japan to be above age 60, this proportion is projected to only increase to 18 percent in South Asia. Simultaneous to this significant demographic change Asia is experiencing very significant changes in terms of its human capital as measured by the educational composition of the population by age and sex. Due to significant educational efforts in many parts of Asia over the past decades, the young population tends to be much better educated than the older one. Since better-educated women in Asia have lower fertility, this has direct demographic consequences. The improvements have been particularly impressive in China, where it is estimated that within two decades China will have more working-age people with secondary and tertiary education than Europe and North America together (Lutz and Goujon 2001).

These future demographic trends, however, are highly uncertain. The projections cited above only refer to the best guess from today's perspective, or more precisely the median of an uncertainty distribution. It is hard to say precisely, how social, economic and even environmental changes in Asia will play together in determining the future trends in fertility, mortality and migration in different parts of the continent. On the other hand, we are not completely ignorant about the future size or structure of the population. Many of the people who will be alive in 2030 have already been born, and we know their cohort sizes. Also, we can assume with high probability that a country in the midst of its fertility transition will continue with its fertility decline until a low level is reached. How should we communicate this to policy makers, who want to get the best synthesis of what we can say about the future trends to base their planning on these forecasts? To tell them that we simply do not know will not serve them well, especially if we think we know more than nothing. It is a bit more informative to say, here is one scenario and here is another, but we do not know how likely they are. This still does not help them much in their planning. Giving them just one projection, however, and telling them that this is the way the future will look, is probably highly welcome by the planners but very dangerous, if there are any costs associated with a projection error. A false feeling of certainty can be very harmful and we should not create it, especially when we know better. The challenge for the population forecaster, therefore, is to try to be specific of what we think is more certain and what is less certain, and to what degree uncertainties differ. The most comprehensive and efficient, and therefore best way to do so is to produce probabilistic forecasts.

In the Asian context this task was taken on by the Asian MetaCentre for Population and Sustainable Development Analysis. The Asian MetaCentre is a group of population research institutes throughout Asia with headquarters at the National University of Singapore and a training branch at Chulalongkorn University in Bangkok (for more information see www.populationasia.org). A series of two workshops in 2001 and 2002 were dedicated to an exercise to identify the main drivers of demographic change in Asia and to translate these insights into probabilistic projections. This included the more traditional analysis of past time series, where available, the study of the errors of past projections and individual expert statements about the assumptions of the most likely trends in vital rates together with subjective quantitative assessments of the uncertainty ranges. This exercise also included new methods such as a rather large number of in-depth interviews of experts in which they were extensively asked about their reasons for making certain assumptions rather than others, and during which an attempt was made to assess about which assumptions the experts felt more confident or less confident. These interviews were conducted by an expert in cognitive science, who is an expert on experts, but not on population. Based on these findings the exercise then went one step further to try to explicitly model the one structural determinant that has been singled out as the single most important one, namely female education, and make the future fertility uncertainty dependent on future education in the context of probabilistic projections. To our knowledge such a structural approach to probabilistic projections has not been applied before.

This paper has four main parts. First, we briefly describe the *ex post* error analysis that has been carried out with reference to the 1975 and 1980 UN projections, in order to set the stage with a view to uncertainty. Next, we will show what happens when the individual expert assessments about trends and uncertainty are merged with respect to individual countries and directly translated into probabilistic projections. We then proceed to the meta-analysis of individual expert views and arguments that also draws on the in-depth interviews; from there we move toward defining a structural model. In the final section, we proceed to a probabilistic projection based on the results of the structural modeling.

The exercise was carried out in close collaboration with more than 50 Asian national population experts.[1] Since in the context of this short paper it is impossible to present these steps for all of the Asian countries for which the exercise was carried out, we had to be extremely selective. Instead of the possibly more consistent way of presenting all the steps for just one country, we chose the alternative option of illustrating different aspects of the exercise with data from different countries to present a better view of the Asian panorama and to better acknowledge the participation of experts from many countries in this exercise.

[1] A list of participants is available from the Asian MetaCentre's home page: www.populationasia.org

Analysis of Errors in Past Projections of Six Southeast Asian Countries

When we think about the uncertainty of future demographic trends and by how much our forecasts today may turn out to be wrong, it is always a useful first exercise to look at projections produced in the past and see how well they did. Different ways for doing such *ex post* error analyses have been discussed by Keilmann (1999) and Alho (1997). A recent report by the US National Research Council (2000) also summarizes work in this field and recommends *ex post* error analysis as an important part of any effort of probabilistic population projections. It seems to us that such an exercise is particularly useful when it not only looks at the mean percentage error of total population size or other output variables, but also at the specific assumptions that lead to these errors and if possible, even tries to find out why and with what justification the forecasters in the past made the assumptions the way they did (Lutz et al. 2003).

Since it was difficult to find a consistent set of nationally-produced population projections for a large number of Asian countries that were produced at least 20-30 years ago, it was decided to use the United Nations projections with base years in 1975 and 1980. These estimates and projections were then compared to the United Nations 1998 assessment (United Nations 1999). This analysis was carried out by the Asian MetaCentre and is fully documented in Khan (2003). Here we will only highlight selected findings with respect to the trends in the total fertility rate.

The upper panel of Table 1 compares the 1975 projections to the 1998 estimates; the lower panel compares the 1980 projections to the same 1998 estimates. We present the comparisons of both projections because there are significant differences that are indicative for what can go wrong with projections. To the right of each panel we find a decomposition of the total error into the error that is due to wrong estimates of the baseline data and the error that is due to wrong assumptions about the change in fertility over time.

For the 1975 projections it turns out that the baseline errors in four of the six countries were more significant than the errors due to wrong fertility change assumptions. For Thailand this error was particularly big; for the period 1975-80, the UN 1975 projections (which is actually the 1978 assessment published in 1980) give a TFR of 5.53 while now the 1975-80 level is only estimated at 4.25. In the Philippines there have been compensating errors in different directions; the 1975-80 level was overestimated by 0.87 children, but the decline was assumed too rapid by 0.74 children, so that the total error turned out to be very small. In other cases, such as Vietnam, there were errors in the same direction. A great underestimation of the decline by 1.54 children adds to a minor baseline error of 0.25 children producing a very significant total error of 1.8 children: Instead of a projected TFR of 4.4 it turned out to be 2.6. Averaging over all six countries, both the baseline error and the change error were 0.6 children. Because of partly compensating errors they do not add up to a total error of 1.2 but only 0.86.

Table 1. Total fertility rate. Source: Khan (2003, p. 18).

	1975 Projection			1998 Estimate (UN 1999)					
Country	1975-1980	1995-2000	Projected decline	1975-1980	1995-2000	Estimated decline	Total error	Base error	Change error
Indonesia	5.13	3.38	1.75	4.68	2.58	2.1	0.8	0.45	0.35
Malaysia	4.26	2.7	1.56	4.16	3.18	0.98	-0.48	0.1	-0.58
Philippines	5.83	3.75	2.08	4.96	3.62	1.34	0.13	0.87	-0.74
Singapore	2.47	2.1	0.37	1.87	1.68	0.19	0.42	0.6	-0.18
Thailand	5.53	3.28	2.25	4.25	1.74	2.51	1.54	1.28	0.26
Vietnam	5.84	4.39	1.45	5.59	2.6	2.99	1.79	0.25	1.54
						Mean absolute error:	0.860	0.592	0.608

	1980 Projection			1998 Estimate (UN 1999)					
Country	1975-1980	1995-2000	Projected decline	1975-1980	1995-2000	Estimated decline	Total error	Base error	Change error
Indonesia	4.81	2.46	2.35	4.68	2.58	2.1	-0.12	0.13	-0.25
Malaysia	5.03	2.46	2.57	4.16	3.18	0.98	-0.72	0.87	-1.59
Philippines	4.62	2.87	1.75	4.96	3.62	1.34	-0.75	-0.34	-0.41
Singapore	1.84	1.74	0.1	1.87	1.68	0.19	0.06	-0.03	0.09
Thailand	4.27	2.51	1.76	4.25	1.74	2.51	0.77	0.02	0.75
Vietnam	5.48	2.87	2.61	5.59	2.6	2.99	0.27	-0.11	0.38
						Mean absolute error:	0.448	0.250	0.578

For the 1980 projection (which is the 1982 assessment published in 1985) the baseline errors are significantly lower. Obviously, the demographic information systems in those countries have improved significantly over these five years. As a consequence the total error is only 0.45, almost half the level of the 1975 projections. The change error, on the other hand, is at almost the same level as above. In other words, while the quality of information about the baseline demographic data has greatly improved, the quality of the assumptions has not – despite the fact that the projection period was five years shorter. The errors again go in both directions. In Malaysia the fertility decline was greatly overestimated, while in neighboring Thailand, it was strongly underestimated.

What did we learn from this *ex post* error analysis? The first thing we learned was the importance of the baseline error, which is often underestimated by forecasters. On the other hand, we see a clearly improving trend over time. Since by any standard demographic estimates through vital registration, surveys and censuses in these countries are significantly better today than in the early 1980s, it does not seem to be meaningful to expect similar baseline errors for today's projections. As to the errors in the assumed fertility change the great difference even between neighboring countries shows that clearly more attention needs to be given to specific national social, cultural und economic features. The national experts from those countries participating in the exercise thought that part of the problem was that the projections were made in New York without such specific knowledge. This is, of course, easy to say in retrospect. The real challenge is how to include such in-depth knowledge about the specific national conditions in new projections. As we will see in the following, exclusively relying on national experts' opinions may also not be the optimal solution.

Directly Translating the Merged Expert Views into Probabilistic Projections for the Philippines

As part of the written enquiry among population experts in Asia about future demographic trends and their uncertainty, seven experts from the Philippines filled out questionnaires specifying uncertainty ranges for fertility, mortality and migration in their country and specifying their reasons for making these assumptions. In addition to these questionnaires, the same group of experts engaged in a group discussion with the task to specify a joint distribution.

Table 2 shows the ranges that the seven experts individually specified for the TFR in 2015-2020 and 2045-2050. As is evident from the table, several experts did not define symmetric uncertainty distributions. This is perfectly consistent with the task given to them, namely, to allocate 100 percentage points over a given range of TFR categories for the specified years. The questionnaire gave them several different examples of how this could be done, but the examples had nothing to do with demography in order not to influence their statements; the sample distributions were both symmetric and asymmetric.

Table 2. Fertility levels and uncertainty ranges defined by individual experts from the Philippines.

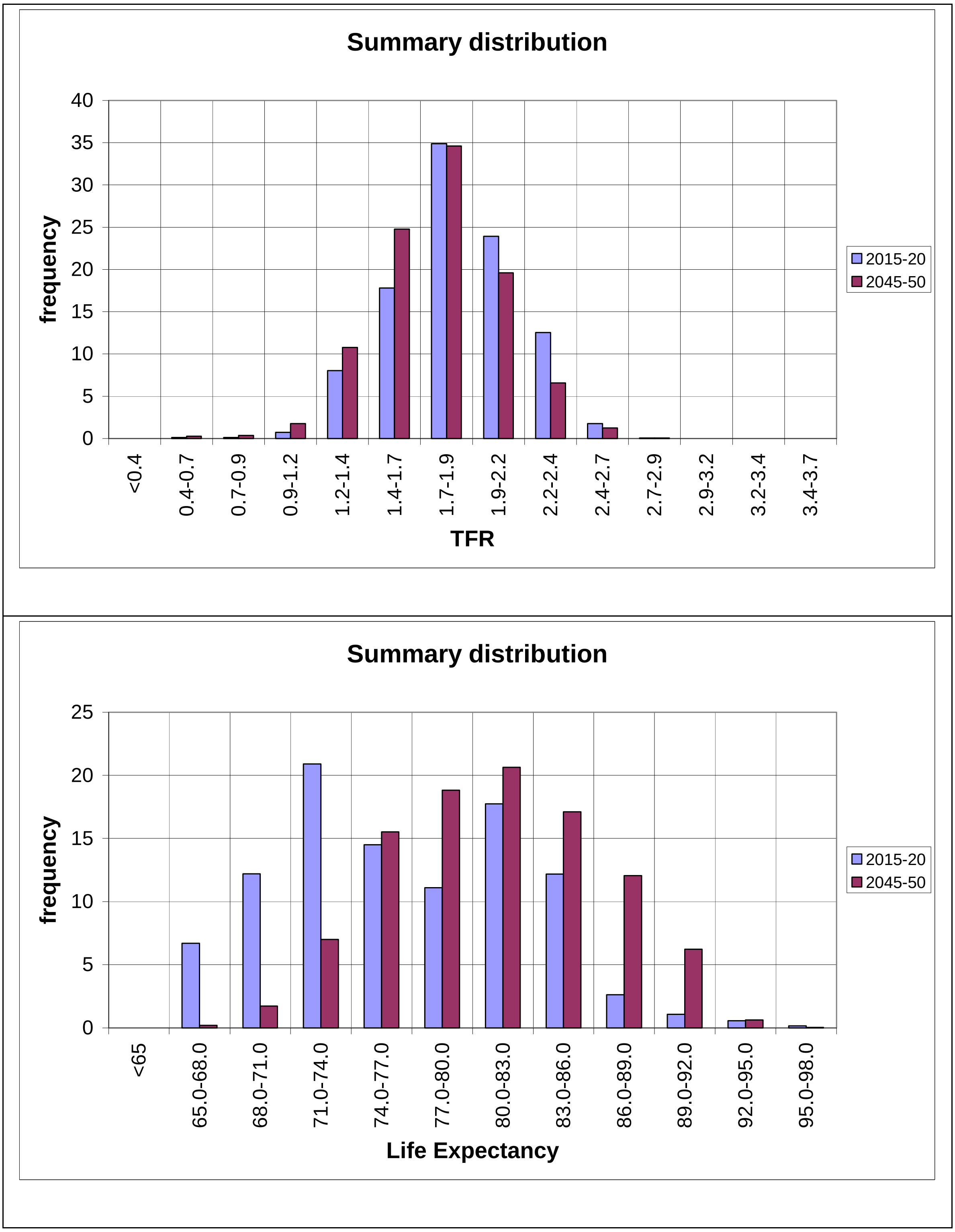
TFR in 2015-2020 Range	Expert number 1	2	3	4	5	6	Range	Expert number 7
<1.33	10	10						
1.33-1.67	10	85		2	10		1.3-1.8	90
1.67-2	80	5	90	25	80	20	1.8-2.3	10
2-2.33			10	68	10	70		
2.33-2.67				5		10		
2.67-3								

TFR in 2045-2050 Range	Expert number 1	2	3	4	5	6	Range	Expert number 7
<1.33	10	10					<1.3	20
1.33-1.67	10	75	90	4	10		1.3-1.8	60
1.67-2	80	15	10	65	90	40	1.8-2.3	20
2-2.33				30		50		
2.33-2.67				1		10		
2.67-3								

Figure 1 merges these seven independent distributions by giving equal weight to each expert. Since the six experts were of different academic standing and for some the substantive arguments listed for justifying their choice made more sense than for others (with this not being a direct function of academic seniority), one could have weighted the individual experts when producing the merged distribution. But since there are no easily accepted criteria for assigning different weights to individual experts, we leave this to future research and assume equal weights here.

Although the merged distributions shown in Figure 1 are more symmetric than the individual distributions, they do not perfectly resemble normal distributions, which have been the basis for most expert-defined uncertainty distributions so far (Lutz et al. 1999). Since there is no reason why the simulations that produce the probability distributions of output parameters need to be based on symmetric assumptions, here we take the expert-defined assumptions at face value and directly translate them into independent cohort component projection simulations, which also consider the mortality assumptions that were specified by the experts in a similar manner as fertility (Scherbov 2002). The only thing that could not be done with these non-normal distributions was to assume correlations between the distributions as done in Lutz et al. (2001). But since we assume independent fertility, mortality and migration trends for the Philippines, and since inter-country correlations are not an issue because we produce independent projections for each country, this is not a limiting factor here.

Figure 1. Merged distribution of the individual ranges given in Table 2 for fertility (above) and life expectancy (below).



The results of the group processes for defining uncertainty distributions are discussed in more detail in Scherbov (2002). The distributions agreed by the groups tend to be significantly different from the merged distribution of the individual experts. In the case of the Philippines the group consensus has a lower range of uncertainty. This is a process frequently described in the literature on group dynamics and attributable either to the dominance of one person or a group desire for uniformity with individuals uneasy to challenge the majority opinion. For this reason we took the merged distribution of individual statements as the basis for the projections presented below.

Figure 2 and Table 3 give the results of 1,000 independent cohort component projection simulations based on the fertility distributions as shown in Table 2 and the corresponding individual mortality distributions. The stochastic process assumed includes annual fluctuations in fertility and mortality as described in Lutz et al. (2001). The results show that the total population of the Philippines is likely to increase significantly over the coming 20 years from around 77 million to well above 100 million. For 2050 the 80 percent uncertainty range is 109-145 million, the median 127 million. Table 3 also shows that simultaneous to this significant growth of approximately 65 percent, the population of the Philippines will rapidly age. The proportion aged 65+ will likely increase by a factor of four from currently 4 percent to 17 percent. The 80 percent uncertainty range in 2050 is 14-20 percent, which means that an increase by at least a factor of three is almost certain.

Figure 2. Probabilistic projection of total population size in the Philippines.

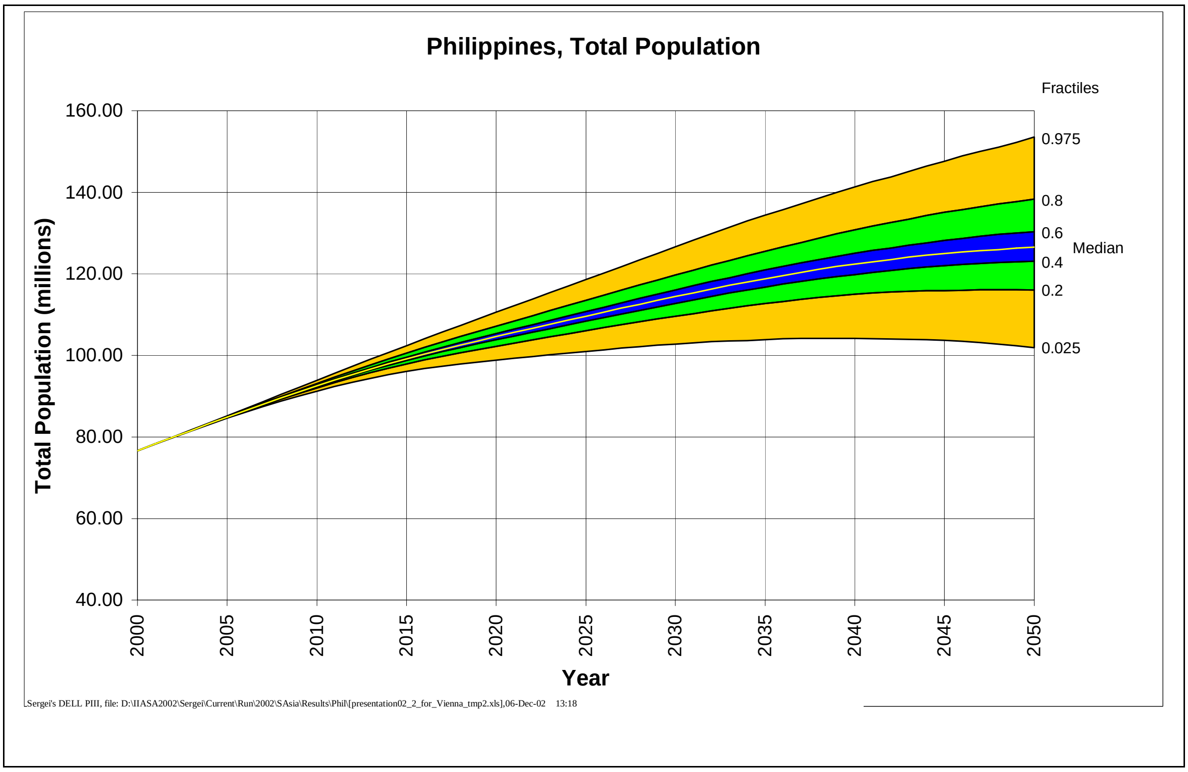


Table 3. Results of probabilistic projections for the Philippines for selected demographic variables.

	2000	2010	2020	2030	2040	2050
Total population (in millions)	76.618 (76.612-76.625)	92.521 (91.649-93.410)	104.586 (100.580-108.451)	114.460 (106.346-122.495)	122.416 (110.021-135.151)	126.587 (109.488-145.146)
Proportion below age 15	0.363 (0.363-0.363)	0.318 (0.313-0.323)	0.252 (0.230-0.272)	0.205 (0.172-0.235)	0.192 (0.160-0.223)	0.170 (0.136-0.205)
Proportion 15-65	0.599 (0.599-0.599)	0.636 (0.631-0.641)	0.683 (0.663-0.703)	0.699 (0.673-0.730)	0.680 (0.656-0.708)	0.662 (0.637-0.688)
Proportion above age 65	0.038 (0.038-0.038)	0.046 (0.045-0.048)	0.065 (0.060-0.073)	0.093 (0.082-0.110)	0.125 (0.106-0.151)	0.166 (0.138-0.200)
Old age dependency ratio (65+ / 15-65)	0.064 (0.064-0.064)	0.073 (0.071-0.075)	0.095 (0.087-0.107)	0.132 (0.116-0.158)	0.182 (0.155-0.224)	0.250 (0.206-0.310)
Support ratio	15.656 (15.652-15.660)	13.732 (13.271-14.065)	10.560 (9.322-11.516)	7.568 (6.317-8.652)	5.487 (4.468-6.440)	4.004 (3.230-4.863)

Meta-Analysis of Expert Views and the Transition to a Structural Model Considering Education

The above-described projections directly translate what a group of national experts said were their best knowledge-based assumptions for the future demographic trends in the Philippines. As international experts designing and analyzing this exercise in defining uncertainty ranges of future fertility and mortality, we are certainly less knowledgeable than the national experts about the specific conditions and trends in the countries concerned. But we also see some problems that tell us that to simply take the distributions defined by the group of national experts and directly translate them into projections, should not be the end of the story. More can and should be done.

Taking a bird's eye view as meta-experts, there are two apparent concerns with simply taking the numerical assumptions as given by the national experts. First, the merging of the individual distributions turned out to be quite different from the uncertainty distribution subsequently defined through a group process. Second, these numerical assumptions neither reflect the value of the arguments provided by the individual experts nor the analysis of the large array of issues raised in the in-depth interviews with the experts. After reading through the transcribed interviews (which do not disclose the person's identity) and looking at their analysis by an experimental psychologist (the analysis of these interviews will be described elsewhere) the specific quantitative statements made by the experts certainly look less stable and reliable. It also showed that some national experts were very hesitant to make such statements, while others were extremely confident in what they specified. But this could be more a feature of the personality and his/her character than a reflection of the solidity of the knowledge base.

How can one further proceed from this sobering assessment of the cognitive base on which individual experts based their statements about future trends and their uncertainty ranges? In the context of scientific work, the only way out of this is to take the usual scientific approach to try to distinguish between the message and the substantive arguments on the one hand, and the person who makes these points on the other. Although in practice these two things can never fully be separated, there are methods of scientific discourse and peer review that can evaluate arguments independently of the person bringing them up. This issue is extensively discussed in another paper (Saariluoma 2002). Since the substantive arguments brought up by the experts in explaining and projecting demographic trends covered a very wide spectrum (ranging from the impact of electrification in rural villages on fertility to specific government heath policies) it turned out to be impossible to carry out the broad meta-level evaluation of specific arguments in the context of the international literature on these topics, as had originally been planned. Instead we chose to single out the one argument that would feature most prominently in most of the expert interviews. The choice was not difficult because in addition to the rather diffuse reference to all kinds of government policies, female education clearly stood out as the single most important factor mentioned. All of the national experts from the different countries in Asia involved in the exercise almost consistently mentioned that they thought that the improving level of female education has been and will be the main driver of fertility decline.

The line of argumentation in these interviews seems clear: The combination of great educational fertility differentials in Asia (more highly educated women have significantly lower fertility) with the fact that younger women are and will be more educated than the older women, greatly contributes to fertility decline. However, as clear and convincing as the argument seems at first, it becomes a bit more complex when thought through in specific quantitative terms.

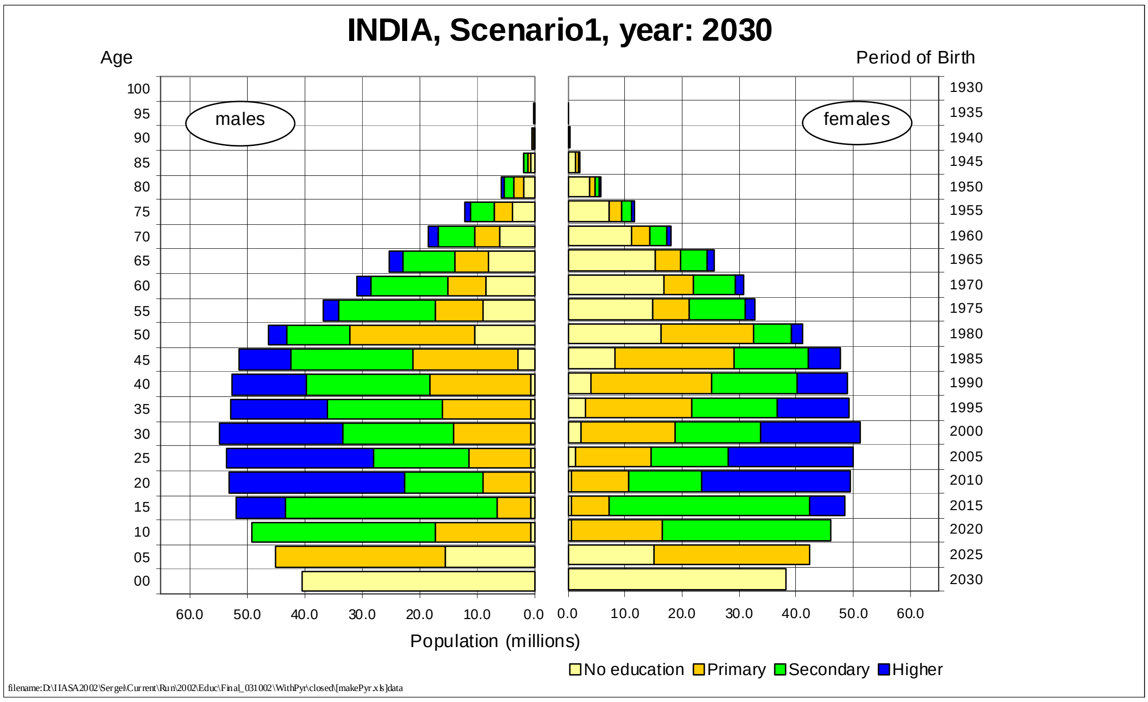
As an example of this exercise of trying to operationalize the more qualitative expert view that the education of women greatly matters for fertility, we chose the world's two most populous countries: India and China. Table 4 gives the most recently available educational fertility differentials for women in the two countries. The table gives the TFR for four categories of women: those without any formal education and those with some primary, secondary and tertiary education. The precise definition and discussion of these categories and educational fertility differentials in general are given in Lutz and Goujon (2001). The table clearly and impressively shows how in these two major countries, higher education is associated with lower fertility. The total TFR gives the fertility of the total population with weights of the different educational groups corresponding to the current educational composition in India and China.

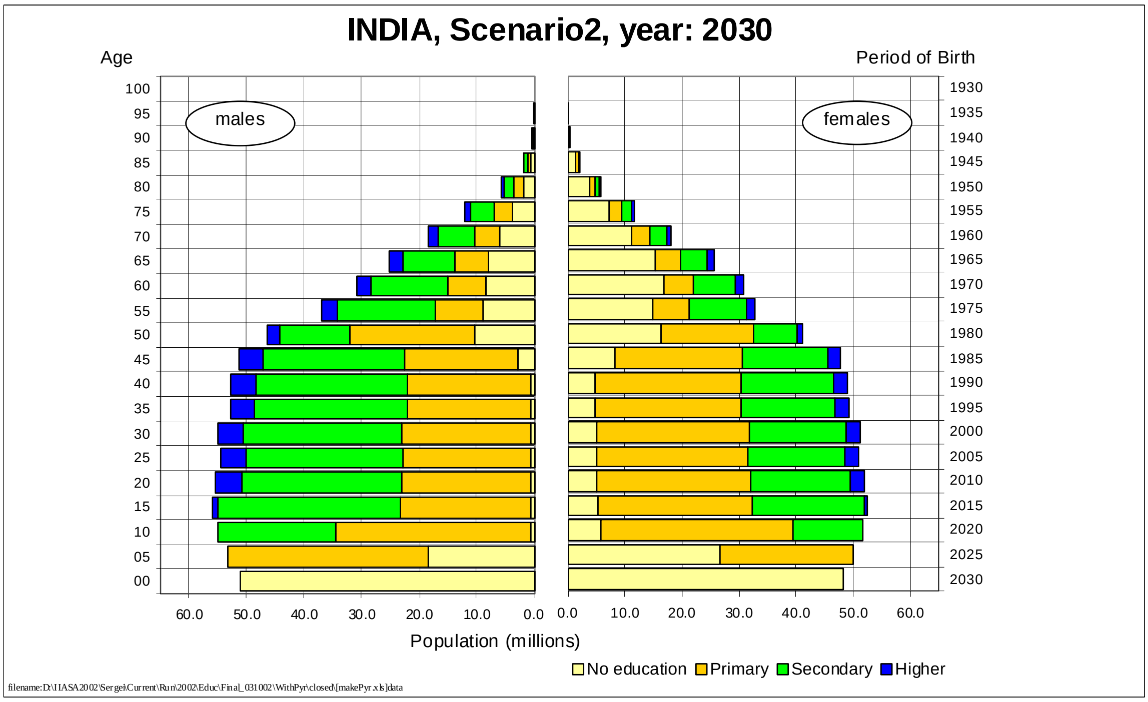
Table 4. Education-specific and total TFR India and China.

	India 2000	China 2000
No education	3.78	2.43
Some primary education	2.89	2.14
Some secondary education	2.36	1.63
Some tertiary education	1.96	1.08

Figure 3 gives the age-, sex- and education-specific population pyramids for India as projected for 2030 under constant enrolment rates (Scenario 2) and ICPD goals (Scenario 1 – see definition below). The pyramids clearly show that the younger cohorts of women in India are better educated than the older ones, although the gender gap is still significant. Under the ICPD scenario the gender gap disappears for the younger segment of the population.

Figure 3. Education-specific age pyramids for India under the two alternative education scenarios in 2003.





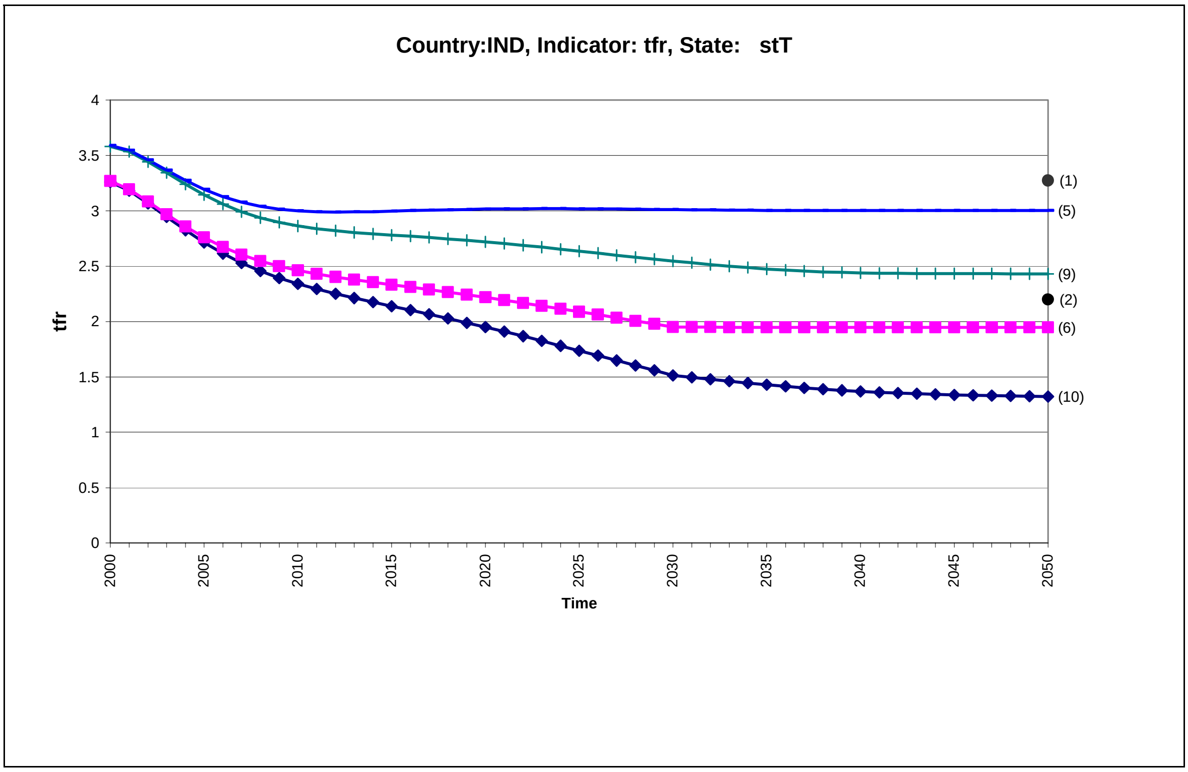
When we think about the future of fertility in India, we must differentiate between two different effects: (a) the change in the educational composition of the population, and (b) the fertility trends within each educational group. Empirical analysis shows that both have been changing in the past in most countries and there is reason to assume that they will continue to change in India, which is still in the midst of the process of demographic transition. For the educational composition, further improvement is a near certainty because it is already pre-programmed into the age structure: The younger, better-educated cohorts will inevitably become older and replace the older, less-educated cohorts. Only the future education of the young cohorts is uncertain at this point, depending on future school enrolment rates at different levels. Policies can make a difference here. We capture this difference through two extreme scenarios on girls' education in India, one scenario in which all enrolment rates stay constant in the future (in a way the most pessimistic one) and another in which India manages to implement the ambitious education goals as defined in the 1994 Cairo World Population Conference (ICPD). This highly optimistic scenario assumes the elimination of the gender gap in primary and secondary education by 2005-10, 90 percent net primary enrolment by 2010-15 and secondary enrolment of 75 percent by 2025-30, as well as an increase in transition to tertiary education by 5 percentage points until 2025-30. Trends between 2000 and the target year are based on linear interpolation.

Table 5 defines 12 scenarios that result from the cross-classification of different future trends in the educational structure of the population and different trends in education-specific fertility rates. The first four scenarios that keep the educational structure of the population frozen are of a purely hypothetical nature because, as has been mentioned above, it is already embedded in the age structure that over time the younger, more educated age groups will move the age scale and improve the average education of the female population of reproductive age. But it is still important to talk about this hypothetical case of the educational composition by age remaining constant in its current form because it serves as a point of reference in the minds of experts thinking about this issue. For this reason Figure 4 shows the aggregate level TFRs resulting from Scenarios 1 and 2 in 2050 as black dots to the right of the figure. If the educational composition remains frozen and the education-specific fertility remains constant (Scenario 1), then clearly the aggregate TFR remains constant over time. Scenario 2 gives the case in which the educational composition remains frozen, but education-specific fertility rates decline as discussed below. In this case the aggregate TFR declines by about one child, which is the fertility trend effect that is completely free of the effect of the changing educational structure.

Table 5. Definition of 12 scenarios combining different possible trends in the educational structure with different assumptions about education-specific fertility trends.

	Educational Structure of the Population		
Fertility trends for educational categories	Structure constant (purely hypothetical)	Enrolment rates constant	ICPD goals for enrolment
All fertility rates constant	1	5	9
Rates reach Chinese level by 2030	2	6	10
China +0.5 by 2030	3	7	11
China –0.5 by 2030	4	8	12

Figure 4. Selected scenarios for combining education assumptions with education-specific fertility trends.



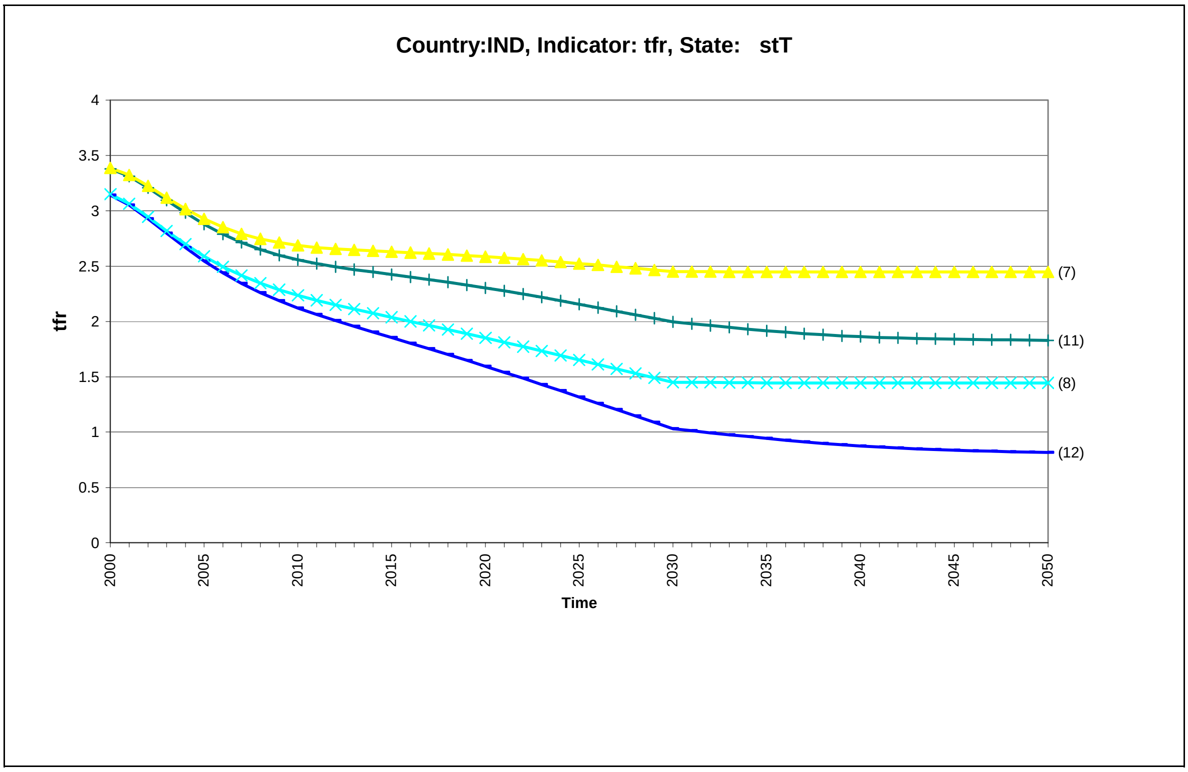
During the course of the Asian MetaCentre exercise, we had quite some discussions about what should be assumed in terms of education-specific fertility trends. Experts did not provide us with their views on this because their assumptions always included the compositional changes. We considered three different options for dealing with these issues: (1) assuming proportional fertility changes in all educational categories, (2) assuming convergence of all educational fertility trends to one target level or (3) assuming that education-specific fertilities will move to the observed levels of another country that is already further advanced in the process of fertility decline. Option (1) is not meaningful as a general rule because in some countries the fertility of university graduates is already so low that no further declines will be expected, although the fertility of less-educated women is expected to decline. In a country like Singapore, where the aggregate TFR is expected to increase, the fertility of the lowest educational group is still above 4.0 und certainly not expected to increase. Option (2) is also not meaningful in this context, because if we have complete convergence, by definition the changes in the educational composition do not affect aggregate fertility. For this reason we chose Option (3), which is consistent with the frequent demographic practice to think in terms of analogies, as is more generally done in the context of the demographic transition. Specifically, for the case of India presented here, we assumed that in 2030, India will have education-specific fertility rates comparable to those of China today. A qualitative discussion of this assumption concluded that this is a meaningful way of handling the issue.

Figure 4 also gives four lines for the cross-classification of the two assumptions for education-specific fertility (constant versus linear change to the Chinese pattern) with the two education assumptions (constant enrolment versus the ICPD Scenario). When comparing Scenarios 5 and 9 and 6 and 10, respectively, we see that the changing educational structure makes a significant difference even with identical education-specific fertility trends. By 2050 this difference accounts for more than half a child, i.e., more than one-third of the level of fertility as given by Scenario 10. If we also consider the hypothetical case of a frozen education structure (Scenario 2) the difference due to differential educational structures becomes almost one child. This clearly illustrates that the experts have made a valid and quantitatively important point when they suggested that the changing educational structure would be a major force towards lower fertility.

How should these insights be translated into probabilistic fertility assumptions? The first thing that is unclear in this context is whether experts, when they stressed the effect of the changing educational composition, had one of the two extremes – the constant enrolment scenario or the ICPD scenario – or something in between in mind. A qualitative discussion with some of the experts at the second seminar indicated that this uncertainty about future school enrolment should be assumed to be part of the total fertility uncertainty. As to the uncertainty of education-specific fertility trends, it was assumed that the uncertainty range considered here would be half a child up and down, as compared to the mean trend (Scenarios 6 and 10), which is the linear move from the Indian 2000 rates to the Chinese 2000 rates by 2030. This is consistent with what was assumed in Lutz et al. (2001). There the 80 percent range was assumed to be plus/minus one child if the TFR was above 3.0, and plus/minus 0.5 if it was below 2.0, with linear interpolation in between. Since this was also supposed to include the education uncertainty in addition to the education-specific fertility uncertainty, these assumptions are roughly consistent.

Figure 5 gives the four scenarios that combine the plus/minus 0.5 children in education-specific fertility assumptions with the two extreme education scenarios. Following the logic outlined above the appearing range between Scenarios 7 and 12 is then taken to represent the 90 percent uncertainty interval of a normal distribution representing India's fertility uncertainty in any given year between 2000 and 2050.

Figure 5. Selected scenarios combining education-specific fertility trends, which are 0.5 children higher and lower than the mean with different future education trends.



New Probabilistic Projections for India, including Education Uncertainty

A final step was to translate these new fertility uncertainty ranges, which include the education uncertainty in addition to the uncertainty of education-specific fertility trends, into full probabilistic population projections. Here we present the results for India. While the fertility distribution is based on the range between Scenarios 7 and 12 in Figure 5 as discussed above, the stochastic mortality assumptions are taken from the projections given in Lutz et al. (2001). For simplicity a closed population is assumed here. The stochastic model chosen assumes annual fluctuations in birth and death rates following the model that is described in Lutz et al. (forthcoming-B).

Figure 6 presents the fractiles of the distribution of India's future population size resulting from 1,000 simulations. As with all such projections the uncertainty range opens up over time. But the graph also shows that with very high probability, India's population will increase from presently 1 billion to more than 1.2 billion by 2020. The median shows a stabilization of India's population size after 2035 at a level of roughly 1.4 billion. But the uncertainty range becomes very broad beyond 2030. While in more than 20 percent of the simulations, India's population starts to decline after 2030, in another 20 percent of the cases, it shows very significant further growth, with the upper end of the 95 percent interval reaching 1.7 billion by 2050. In the black bracket in Figure 6, one sees the total population numbers that result from Scenarios 6 and 10 as discussed above, i.e., the central fertility decline assumption combined with the two different educational scenarios. The difference is very significant and in the order of 0.2 billion. In other words, these calculations imply that with otherwise identical assumptions, an India that will follow the ICPD education goals will have 200 million people less than an India with constant school enrolment ratios.

Figure 6. Resulting distribution of total population size in India.

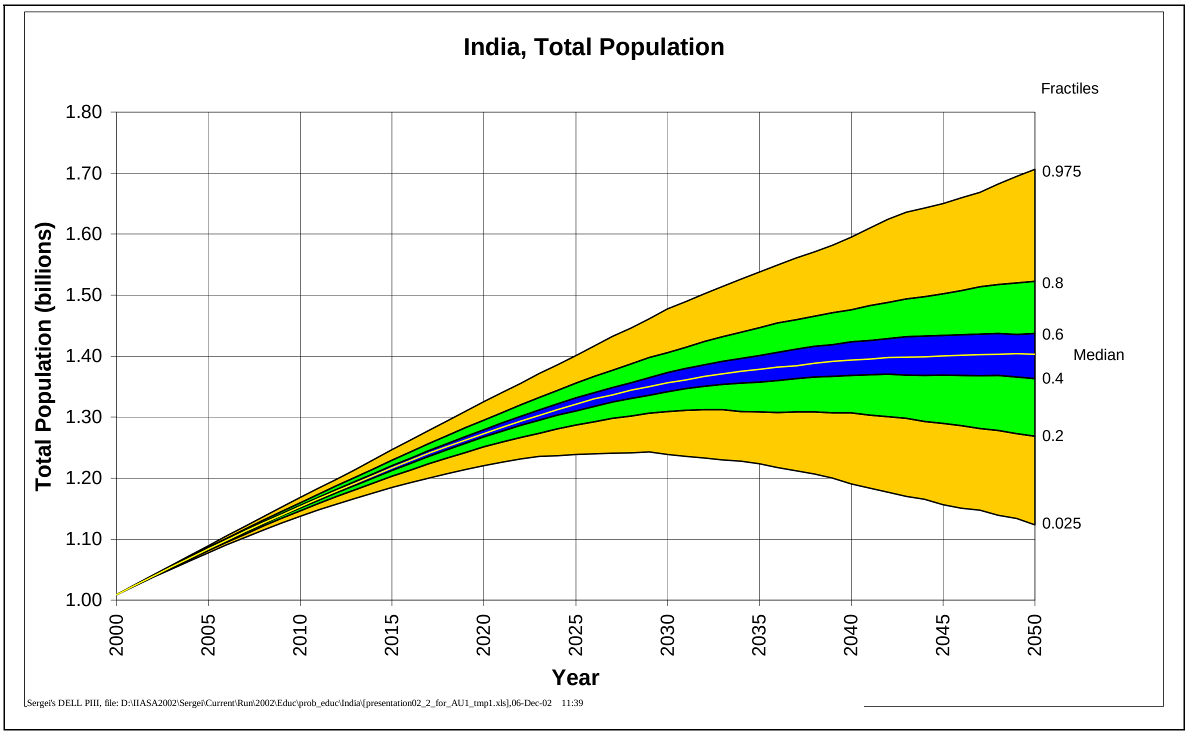
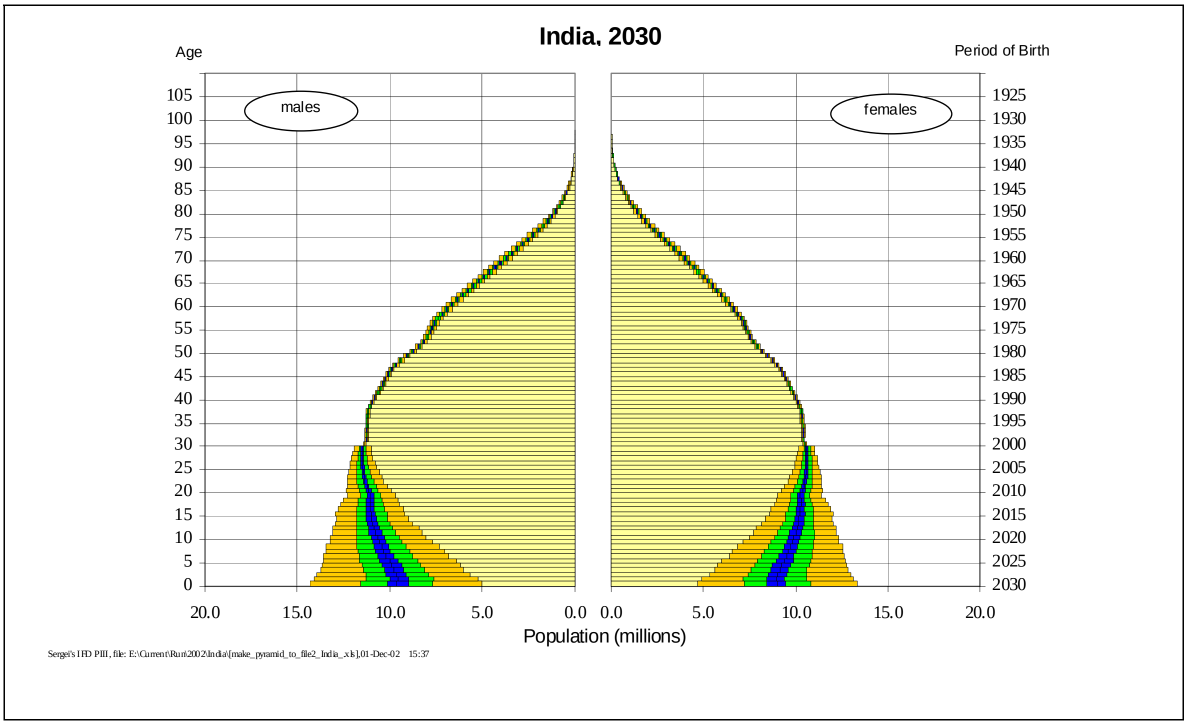


Table 6 gives the numerical results of these new probabilistic population projections for India. The values given refer to the median, with the 80 percent intervals given in parentheses. This shows that despite an expected further population growth of around 40 percent over the coming five decades, India's population will also get significantly older. The proportion above age 65 will increase from currently only 5 percent to around 14 percent with the 80 percent uncertainty range going from 0.12 to 0.16. The proportion of children below age 15 will likely decrease to about half its level, from currently 0.35 to only 0.17. For the children, the 80 percent uncertainty interval is much larger, ranging from 0.11 to 0.21. The fact that the uncertainty ranges differ greatly by age is most clearly shown by the probabilistic age pyramid in Figure 7. Unlike such pyramids for industrialized countries, where there is also visible uncertainty at a very old age due to the uncertainty about the path of future old age mortality, this is not yet visible in India because the population is still much younger and life expectancy is still much lower. Hence the giant share of the uncertainty of India's future population is due to fertility uncertainty.

Table 6. Medians and 80 percent ranges (in parentheses) for selected projection output parameters in India.

	2000	2025	2050
Total population	1.009 (1.009-1.009)	1.321 (1.268-1.374)	1.403 (1.208-1.603)
Proportion below age 15	0.335 (0.335-0.335)	0.238 (0.212-0.260)	0.167 (0.113-0.213)
Proportion 15-65	0.615 (0.615-0.615)	0.691 (0.671-0.714)	0.697 (0.660-0.733)
Proportion above age 65	0.050 (0.050-0.050)	0.071 (0.068-0.075)	0.138 (0.117-0.161)
Old age dependency ratio (65+ / 15-65)	0.081 (0.081-0.081)	0.103 (0.099-0.107)	0.198 (0.172-0.228)
Support ratio	12.395 (12.395-12.395)	9.690 (9.332-10.053)	5.060 (4.388-5.811)

Figure 7. Probabilistic population pyramid for India in 2030.



Conclusions

The paramount importance of the specific path of the future fertility trend in India on the total population size is an *ex post* justification for spending a considerable amount of time and effort to capture some of the structural determinants of future fertility trends in India. Whether or not there will be 200 million more Indians in 2050 is not a trivial issue. As we have shown, this will depend significantly on the future educational efforts in India.

The analysis presented here has only been a first step in the direction of trying to incorporate this very important structural dimension of population uncertainty into population projections. Much more attention should be given to this and other important structural drivers of fertility decline in developing countries in the future. Here we focused on female education as probably the most important observable source of fertility heterogeneity. Urbanization would be the next logical candidate for such analysis. But it should probably be cross-classified with education (Cao and Lutz forthcoming).

In conclusion we can say that the explicit consideration of some of the most important drivers of changes in demographic rates can make a major difference in the way we see the population evolve in the future. This is particularly true in the context of probabilistic population projections, where the uncertainty about the evolution of the structure of heterogeneous populations is added to the uncertainty about demographic trends within each sub-population. This aspect is quantitatively more important in heterogeneous populations, such as India, than in the more homogeneous ones.

How to deal with this issue in the context of specific populations to be forecasted, cannot be determined purely by statistical models. It requires deep substantive analysis and inevitably a degree of expert judgment. But we should be careful to base our assumptions on explicit arguments that are open to the usual instruments of scientific review and evaluation.

References

Alho, J.M. 1997. Scenarios, uncertainty and conditional forecasts of the world population. *Journal of the Royal Statistical Society, Series A* 160(1): 71-85.

Cao, Gui-ying and Wolfgang Lutz. Forthcoming. China's future urban and rural population by level of education. In Wolfgang Lutz and Warren Sanderson (eds.), *The End of World Population Growth, Human Capital and Sustainable Development in the 21st Century*. Draft book manuscript. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Keilmann, Nico. 1999. How accurate are the United Nations world population projections? Pages 15-41 in W. Lutz, J.W. Vaupel and D.A. Ahlburg (eds.), *Frontiers of Population Forecasting*. A Supplement to Vol. 24, 1998, *Population and Development Review*. New York: The Population Council.

Khan, H.T. Abdullah. 2003. A Comparative Analysis of the Accuracy of the United Nations' Population Projections for Six Selected Asian Countries. IR-03-015. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Lutz, Wolfgang and Anne Goujon. 2001. The world's changing human capital stock: Multi-state population projections by educational attainment. *Population and Development Review* 27(2): 323-339.

Lutz, Wolfgang, Warren C. Sanderson and Sergei Scherbov. 1999. Expert-based probabilistic population projections. Pages 139-155 in Wolfgang Lutz, James W. Vaupel and Dennis A. Ahlburg (eds.), *Frontiers of Population Forecasting*. A Supplement to Vol. 24, 1998, *Population and Development Review*. New York: The Population Council.

Lutz, Wolfgang, Warren Sanderson and Sergei Scherbov. 2001. The end of world population growth. *Nature* 412: 543–545.

Lutz, Wolfgang, Warren Sanderson and Sergei Scherbov. Forthcoming-A. The end of population growth in Asia. In *Journal of Population Research*, a Special Issue on Fertility Decline in Asia.

Lutz, Wolfgang, Warren Sanderson and Sergei Scherbov. Forthcoming-B. The end of world population growth. In Wolfgang Lutz and Warren Sanderson (eds.), *The End of World Population Growth, Human Capital and Sustainable Development in the 21st Century*. Draft book manuscript. Laxenburg, Austria: International Institute for Applied Systems Analysis.

Lutz, Wolfgang, Sergei Scherbov and Alexander Hanika. 2003. 'Vienna: A city beyond aging' revisited and revised. Pages 181-195 in Gustav Feichtinger (ed.), *Vienna*

Yearbook of Population Research 2003. Vienna, Austria: Vienna Institute of Demography of the Austrian Academy of Sciences.

National Research Council. 2000. *Beyond Six Billion: Forecasting the World's Population*. Panel on Population Projections. John Bongaarts and Rodolfo A. Bulatao (eds.). Committee on Population, Commission on Behavioral and Social Sciences and Education. Washington, D.C.: National Academy Press.

Saariluoma, Pertti. 2002. The meta-science of expert knowledge in population forecasting. Paper presented at the seminar "How to deal with uncertainty in population forecasting," held at the Institute of Demography, 12-14 December 2002, Vienna, Austria.

Scherbov, Sergei. 2002. Probabilistic population projections based on expert opinion: Methodological and practical issues with application to South Asian countries. Paper presented at the Session on Population Projections at the IUSSP Asian Regional Population Conference, Bangkok, Thailand, June 2002.

United Nations. 1980. *Selected Demographic Indicators by Country 1950-2000: Demographic Estimates and Projections as Assessed in 1978*. New York: United Nations.

United Nations. 1985. *World Population Prospects: Estimates and Projections as Assessed in 1982*. New York: United Nations.

United Nations. 1999. *World Population Prospects: The 1998 Revision, Vol. II: Sex and Age*. New York: United Nations.