



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-03-052

Conditional Probabilistic Population Forecasting

Warren C. Sanderson (wsanderson@notes.cc.sunysb.edu)

Sergei Scherbov (sergei.scherbov@assoc.oeaw.ac.at)

Brian C. O'Neill (oneill@iiasa.ac.at)

Wolfgang Lutz (lutz@iiasa.ac.at)

Approved by

Leen Hordijk
Director

October 27, 2003

Contents

Introduction	1
An Introduction to the Methodology	1
Conditional Probabilistic Forecasting and Scenario Analysis.....	3
Conditional Probabilistic Forecasts with Future Jump-Off Dates.....	6
Concluding Thoughts	12
References	12

Abstract

Since policy makers often prefer to think in terms of alternative scenarios, the question has arisen as to whether it is possible to make conditional population forecasts in a probabilistic context. This paper shows that it is both possible and useful to make these forecasts. We do this with two different kinds of examples. The first is the probabilistic analog of deterministic scenario analysis. Conditional probabilistic scenario analysis is essential for policy makers because it allows them to answer “what if” type questions properly when outcomes are uncertain. The second is a new category that we call “future jump-off date forecasts”. Future jump-off date forecasts are valuable because they show policy makers the likelihood that crucial features of today’s forecasts will also be present in forecasts made in the future.

Acknowledgments

This paper has been submitted to the International Statistical Review for publication. An earlier version was presented at a workshop on “How to deal with uncertainty in population forecasting” held at the Vienna Institute of Demography, December 12-14, 2002, Vienna, Austria.

About the Authors

Warren C. Sanderson is a Professor in the Departments of Economics and History, State University of New York, Stony Brook, NY 11794-4384, USA.

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.

Brian C. O'Neill is a Research Scholar at IIASA and Assistant Professor (Research) at the Watson Institute for International Studies, Brown University, Providence, RI, USA.

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.

Conditional Probabilistic Population Forecasting

Warren C. Sanderson, Sergei Scherbov, Brian C. O'Neill, and Wolfgang Lutz

Introduction

The last decade and a half has witnessed rapid development in the area of probabilistic population forecasting (see Alho 1990, 1997; Alho and Spencer 1985; Keilman et al. 2002; Lee 1999; Lee and Tuljapurkar 1994; Lutz et al. 1996, 1997, 2001; Lutz and Scherbov 1998; and Pflaumer 1988, among others). A probabilistic forecast goes beyond a traditional deterministic one by providing an integrated estimate of the forecast's uncertainty, often a crucial quantity for decision makers. These forecasts give distributions of outcomes rather than single numbers resulting from alternative scenarios. Since policy makers often prefer to think in terms of alternative scenarios (for example, outcomes with and without a certain policy), the question has arisen as to whether it is possible to make conditional forecasts in a probabilistic context.

This paper answers that question by demonstrating how to obtain conditional probabilistic population forecasts. We do this with two different kinds of examples. The first is the probabilistic analog of deterministic scenarios and the second is a new category that we call "future jump-off date forecasts". Both are important for policy analysis.

Scenario analysis is essential for policy makers because it allows them to answer "what if" type questions. For example, they may want to know *what* the age structure of their country would be in fifty years *if* fertility were lower than in the official projections. Future jump-off date forecasts are valuable because they help in answering questions about the value of waiting to learn about how the future is unfolding. For example, a country may be deciding on whether to build up a retirement fund for its citizens. The decision could be made to raise taxes now or to wait ten years to improve its projections of future population aging. Future jump-off date forecasts allow us to assess how much uncertainty about the future is likely to be resolved by waiting.

In the following section, we briefly discuss the probabilistic forecasting methodology used in Lutz et al. (2001). It is the basis for the quantitative examples in the next two sections. We discuss the probabilistic counterpart of traditional scenario analysis, and then we present a first look at future jump-off date forecasts. The last section contains some concluding thoughts.

An Introduction to the Methodology

Creating population forecasts from an initial distribution of the population by age and sex and forecasts of total fertility rates (TFR), life expectancies at birth, and net

migration rates is a widely accepted procedure. Probabilistic population forecasts differ from deterministic forecasts in that they quantify the uncertainty of the course of future rates and therefore must specify future total fertility rates, life expectancies, and net migration rates as distributions and not as points. Distributions can also be used to quantify other uncertainties such as those relating to the base population size.

In order to generate the required distributions, Lutz et al. (2001) let v be the total fertility rate, the change in life expectancy at birth, or net migration to be forecasted for periods I through T and v_t be its forecasted value at time t . The forecasted value, v_t , can be expressed as the sum of two terms, its trend (mean) at time t , \bar{v}_t , and its deviation from the mean at time t , ε_t . In other words, $v_t = \bar{v}_t + \varepsilon_t$, where ε_t is the idiosyncratic noise. The \bar{v}_t were chosen based on the arguments given in Lutz et al. (1994, 1996) and updated based on subsequent information. The ε_t term is assumed to be a normally distributed random variable with mean zero and standard deviation $\sigma(\varepsilon_t)$. The $\sigma(\varepsilon_t)$ are also based on arguments from the same sources.

Because of the persistence of the factors represented by the ε_t , we would generally expect them to be autocorrelated. One of the most commonly used methods of specifying how the ε_t term evolves over time is the simple autoregressive formation (AR(1)), where $\varepsilon_t = \alpha \cdot \varepsilon_{t-1} + u_t$, where u_t is an independently distributed normal random variable with mean zero and standard deviation $\sigma(u)$. Another commonly used method is the moving average formation of order q , MA(q) where q is the number of lagged terms in the moving average. We use the following moving average specification:

$$\varepsilon_t = \sum_{i=0}^q \alpha_i \cdot u_{t-i}, \text{ where } u_{t-i} \text{ are independently distributed standard normal random variables. To ensure that the standard deviation of } \varepsilon_t \text{ is equal to its prespecified value, } \alpha_i = \frac{\sigma(\varepsilon_t)}{\sqrt{q+1}}.$$

The choice between AR(1) and MA(q) does not have to do with estimation, but rather with representation. Data do not exist that would allow the estimation of the parameters of either specification at the regional level used in Lutz et al. (2001). Neither is more theoretically correct than the other. Both are just approximations to a far more complex reality. When comparably parameterized, they produce very similar distributions of ε_t

The choice between the two, therefore, rests on which more accurately reflects arguments concerning the future. From our perspective, the moving average approach has the advantage that the $\sigma(\varepsilon_t)$ terms appear explicitly making it easier to translate ideas about the future into that specification.

The future levels of vital rates can be correlated in different ways. Most important are (a) the correlations between deviations from assumed average trends in fertility and mortality rates, (b) the autocorrelation of deviations within each series of vital rates, and (c) the correlations among the deviations from the average vital rate trends in different world regions. The forecasts of the world's population used in this paper assume: (1) a zero correlation between fertility and mortality deviations from their

trends within regions, (2) a 31 term moving average specification separately for fertility and mortality deviations, which implies an autocorrelation between deviations one year apart of around 0.96, and (3) cross-regional correlations of fertility and mortality deviations within each year of 0.7 and 0.9, respectively. This methodology is considerably different from the one used in Lutz et al. (1996, 1997), where piecewise linear paths for future vital rates were used.

Because of temporal and regional correlations, vital rates paths for all regions are determined simultaneously and then used to make population forecasts, which were aggregated to the world total. This process was repeated 2,000 times, generating a distribution of world population sizes for each year from 2001 to 2100.

Conditional Probabilistic Forecasting and Scenario Analysis

One important audience for probabilistic forecasts is the user community. Often when demographers want to communicate the importance of particular variables in their forecasts to members of this community, they use scenarios. In population forecasting, scenarios are typically clear “if...then” statements in which the implications of a certain set of assumptions on fertility, mortality and migration are being demonstrated (Lutz 1995). Such scenarios can illustrate the laws of population dynamics but do not give the user any information about the likelihood of the described path. For instance, an immediate replacement fertility scenario merely shows what would happen if fertility immediately jumped to the replacement level without saying that this is a likely or even plausible path. For policy makers who want to know what would be the long-term consequences of alternative fertility trends resulting from alternative policies, for example, such scenarios can nonetheless be useful guides. Conditional probabilistic forecasting is a way of posing and answering the same type of question within a probabilistic framework.

The first discussion of conditional probabilistic population forecasts, of which we are aware, appears in Alho (1997). Alho first turned the deterministic world population forecasts in Lutz et al. (1994) into a probabilistic one and computed the probability of the world’s population falling between the high and low scenarios. Next, Alho considered the case where the UN’s world population forecasts for 2025 could be regarded as a Lutz et al. (1994) forecast conditional on the success of family planning programs. Alho showed that if the probability of being between the UN’s high and low variants was 75 percent, then those programs would have to reduce the variance of the probabilistic version of the Lutz et al. (1994) forecasts by at least 42 percent. Alho regards this as “much too high to be credible” in light “of the past record of ineffectiveness of government interventions concerning fertility in the industrialized countries” (p. 83). He showed that if the reduction in the variance were less than 42 percent, then the probability content of the interval between the UN high and low variants must be less than 75 percent.

Alho (1997) is an example of taking known unconditional and conditional distributions and learning about the nature of the conditional distribution by studying the plausibility of the conditions needed to obtain it from the unconditional one. Here an example that is at the other end of the continuum is presented. Starting with unconditional distributions and conditions that are of interest to policy makers, the

example demonstrates how probabilistic forecasting can produce conditional distributions that are useful in scenario analysis. There are many possible intermediate cases as well, where information about some aspects of conditional distributions and some features of the conditions themselves are combined in order to investigate particular questions. An example of this can be found in O'Neill (2003).

The approach used here was developed in Sanderson et al. (forthcoming). An application on whether immigration can compensate for Europe's low fertility appears in Lutz and Scherbov (2002). The example begins with Figure 1, which shows the distribution of the world's population in 2050 conditional on average fertility and mortality levels for the world over the period 2000-2050. The x-axis is divided into three ranges labeled "low fertility," "medium fertility," and "high fertility." Low fertility includes all of the 2,000 simulated futures where the average total fertility rate in 2000-2050 was below 1.6. Medium fertility includes those paths where the average total fertility rate was between 1.6 and 1.8; and high fertility includes paths in which the average total fertility rate (over the whole projection period) was above 1.8.

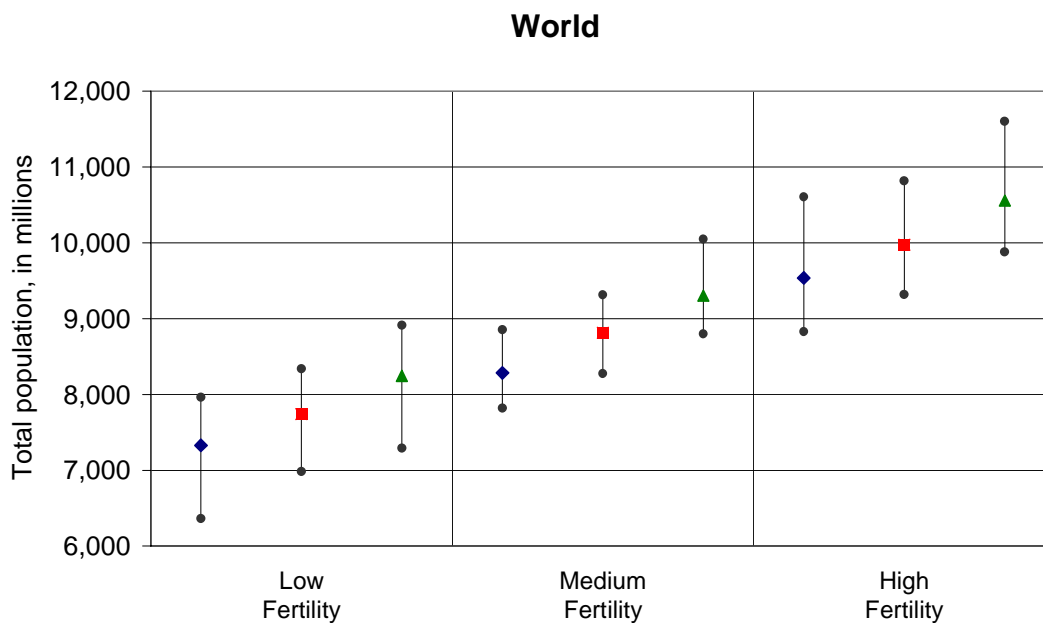


Figure 1. Median and interdecile ranges for the world population, conditional on three alternative fertility and mortality levels. The three lines within each category refer to the low (left), central (middle), and high (right) groups of life expectancy. Source: Authors' calculations.

Within each of the three panels there are three lines that have different symbols near their centers. The lines with the diamonds near their centers refer to paths where the global average life expectancy at birth was lower than 68 years. The lines with the dented squares refer to paths where the average life expectancy was between 68 and 71, and the lines with triangle shapes to paths with average life expectancies over 71. The aggregations of the total fertility rates and life expectancies at birth were chosen so that one-third of our paths was in each group. The symbols are placed at the medians of the distributions. The circles at the endpoints of the lines indicate the 80 percent prediction intervals.

Now we are in the position to answer some “what-if”-type questions. For example, what would be the effect on world population size in 2050 of high fertility trends versus low fertility trends over the coming decades combined with the medium range of uncertainty for future mortality? We can immediately read the answer off the figure. In the middle group, the median population of the world in 2050, if we experienced low fertility, would be around 7.7 billion people with the 80 percent prediction interval covering the range 7.0 to 8.3 billion people. If we experienced a high fertility world, the median population would be considerably higher, around 10.0 billion people, with a prediction interval between 9.0 and 10.9 billion people. The difference between the medians is 2.3 billion people, which is quite large considering that the median of the unconditional population distribution is 8.8 billion people. Clearly, the difference in fertility is very significant.

We can also read the figure to tell us about the influence of differences in life expectancies on future population size. We can do this easily by looking at the middle panel, labeled “medium fertility.” When life expectancies are in the low group, the median population size is 8.3 billion. When they are in the high group, the median population is 9.2 billion. Therefore, in 2050 the effect on population size of moving from low to high fertility, keeping life expectancy constant, is much larger than the effect of moving from low to high life expectancy, keeping fertility constant.

Figure 2 is similar to Figure 1, except that it deals with the proportion 60 years and above. As fertility increases, the proportion 60 and above decreases, but as life expectancy increases, the proportion gets larger. Let us consider the difference in the proportion due to having high fertility as opposed to low fertility, again assuming medium life expectancy. The median proportion is 25 percent when fertility is low and around 19 percent when it is high. Assuming medium fertility and varying mortality, we see that when mortality is low the proportion is below 20 percent, compared to 24 percent when mortality is high. Thus, the effects of fertility and mortality are more similar in determining the proportion 60 and above than they are in determining population size.

The two examples in this section show that in making the transition from deterministic to probabilistic forecasting, we do not have to give up on answering the kinds of “what-if” questions that users and policy makers so often pose.

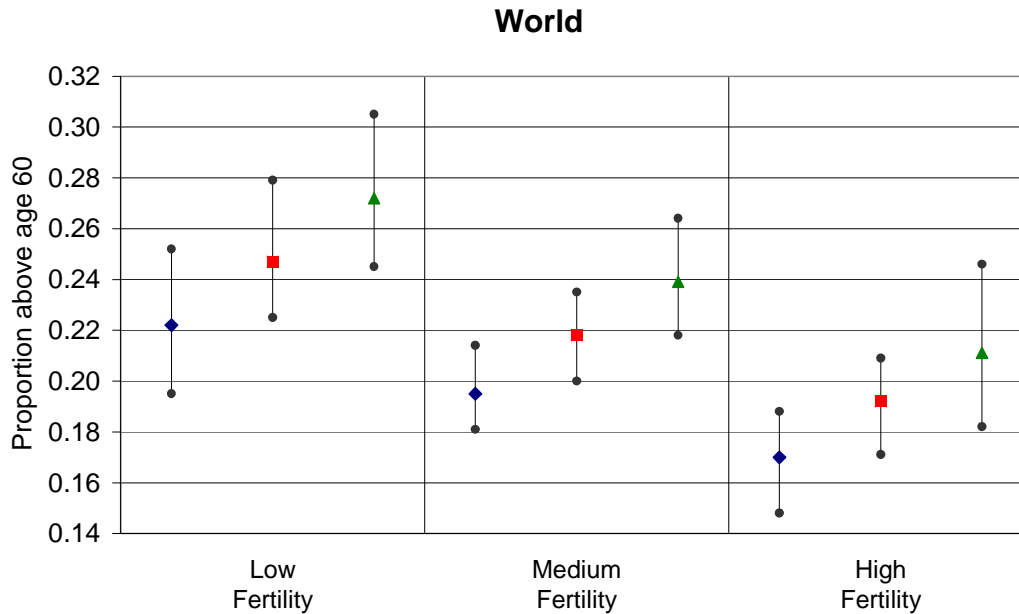


Figure 2. Median and interdecile ranges for the global proportion above age 60, conditional on three alternative fertility and mortality levels. The three lines within each category refer to the low (left), central (middle), and high (right) groups of life expectancy. Source: Authors' calculations.

Conditional Probabilistic Forecasts with Future Jump-Off Dates

In many policy areas, we come across the question: Should we act now or should we wait until we learn more? Waiting has a cost because it can foreclose certain policy options or make them more expensive. On the other hand, by waiting policy makers could possibly acquire important and relevant information, and avoid potentially unnecessary policy interventions. Since population is an important driver of many processes, it is valuable to know how much the demographic outlook might change if we wait.

For example, the question of whether to act now or wait to learn more is central to the debate over climate change policy. The climate change issue is characterized by both long timescales – today's emissions of greenhouse gases will affect climate for decades to centuries – and substantial uncertainties in climate impacts on society and costs of emissions reductions. Many argue that it would be beneficial to wait to learn more (and reduce uncertainties) before deciding whether, and how much, to reduce emissions. This strategy would avoid investments in emissions reductions that may turn out to be unnecessary. Others argue that reductions should begin now, because if climate change turns out to be serious, we would later regret not acting early. The question of how the potential for learning about various aspects of the problem affects today's optimal decision remains unresolved (Webster 2002). It is possible that learning about the outlook for future population growth could impact such decisions. Population is one factor affecting the outlook for future greenhouse gas emissions. If, by waiting a decade or two, we learn that population is likely to be much lower in the future than we currently expect, our outlook for future emissions will likely also be revised downward,

reducing the urgency of emissions reductions. If we learn that population is likely to be much higher, our outlook for emissions will also be higher, justifying more aggressive action to reduce emissions.

Probabilistic forecasts with future jump-off dates are constructed to help us learn about the value of waiting for more information. These forecasts are, of course, conditional on what happens between the beginning of the current forecast period and the future jump-off date. For example, imagine that it is the year 2000 and forecasts are made of the distribution of the size of the world's population in 2050. How different would the forecasted distribution of population sizes be in 2050 if the forecast were made in 2010 instead of 2000? We do not have to wait to 2010 to answer this question. The technique of making probabilistic forecasts with future jump-off dates allows us to think about this question now.

Projections in 2010 may differ from projections in 2000 because something is learned between now and then. At a minimum, the values of demographic variables like population size, fertility, mortality, and migration in that ten-year period will be observed. Other factors such as new policies, economic trends, or social conditions that are relevant to the outlook for future demographic rates will also be observed. It is possible as well that demographic theory will be improved through research, that new breakthroughs in health (or new epidemics of disease) will occur, or that new contraceptive technology will be developed. All of these types of learning could change the outlook for the future. Learning based on these other factors is not considered here. In the example below, learning is only based on the observation of demographic variables. While learning by observation is only one type of learning, it is likely to be an important one in population projections.

In this section, we take some small first steps toward understanding how this passive learning process takes place, so that users of forecasts are not surprised when forecasts change and so that policy makers can use probabilistic forecasts in the design of adaptive policies.

Let us imagine that it is now 2010 and all the relevant population information has been compiled and is available. Certainly it would be appropriate to make new forecasts, even if the methodology and assumptions that were originally used were completely correct. The forecasts based in 2010 would take into account what actually happened between 2000 and 2010. Without actually making new forecasts, the projections made in 2000 could be used to anticipate what new projections would look like.

In order to make this inquiry practical, a very simple approach will be used here. Instead of observing exact population characteristics, the assumption is made that only whether or not global population size is above or below the median of its distribution can be observed. There is nothing theoretically attractive in dividing the observations into only two groups in 2010, but it makes this introduction to passive demographic learning as simple as possible.

Table 1. Forecasted distributions of the world's population size beginning in 2000 and beginning in 2010. Source: Authors' calculations.

World	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: 2000 Jump-Off Date										
	Below 6	6 to 7	7 to 8	8 to 9	9 to 10	10 to 11	11 to 12	Above 12	Median	RIDR*
2000	0	100	0	0	0	0	0	0	6055	0
2010	0	85.05	14.95	0	0	0	0	0	6828	0.062
2020	0	8.05	81.45	10.5	0	0	0	0	7538	0.129
2030	0.1	3.05	41.85	47.35	7.5	0.15	0	0	8085	0.195
2040	0.15	3.75	24.85	40.25	25.15	5.2	0.65	0	8525	0.27
2050	0.5	5.05	18.95	30.95	26.8	13.45	3.3	1	8796	0.352
2060	1.55	7	17.45	25.75	22.35	16.25	6.45	3.2	8935	0.427
2070	4	8.35	16.9	21.2	19.9	15.05	8.5	6.1	8974	0.52
2080	6.8	10.1	16	18.45	18.45	12.55	9	8.65	8890	0.606
2090	10	12.75	14.85	17.9	15	12.25	6.45	10.8	8678	0.702
2100	14.25	14.05	14.45	16.5	12.9	10.45	6.85	10.55	8413	0.779

* Relative Interdecile Range (RIDR) is measured as the difference between the ninth decile and the first decile divided by the median.

World	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel B: 2010 Jump-Off Date										
	Below 6	6 to 7	7 to 8	8 to 9	9 to 10	10 to 11	11 to 12	Above 12	Median	RIDR*
2020/L	0	16.1	83.9	0	0	0	0	0	7268	0.088
2020/H	0	0	79	21	0	0	0	0	7787	0.081
2030/L	0.2	6.1	70.3	23.4	0	0	0	0	7704	0.147
2030/H	0	0	13.4	71.3	15	0.3	0	0	8486	0.139
2040/L	0.3	7.5	42.3	42.4	7.4	0.1	0	0	7996	0.224
2040/H	0	0	7.4	38.1	42.9	10.3	1.3	0	9083	0.216
2050/L	1	9.9	31.1	36.6	17.9	3.2	0.3	0	8152	0.294
2050/H	0	0.2	6.8	25.3	35.7	23.7	6.3	2	9521	0.279
2060/L	3.1	12.8	26.9	31.3	17.3	6.8	1.6	0.2	8256	0.389
2060/H	0	1.2	8	20.2	27.4	25.7	11.3	6.2	9760	0.352
2070/L	7.8	13.6	23.9	25.2	17.4	8.7	1.9	1.5	8213	0.485
2070/H	0.2	3.1	9.9	17.2	22.4	21.4	15.1	10.7	9891	0.438
2080/L	12.5	14.7	22.2	19.8	16.9	7.9	4.1	1.9	8045	0.568
2080/H	1.1	5.5	9.8	17.1	20	17.2	13.9	15.4	9816	0.536
2090/L	16.1	18	18.5	18.3	13.3	9.4	3.6	2.8	7888	0.641
2090/H	3.9	7.5	11.2	17.5	16.7	15.1	9.3	18.8	9638	0.647
2100/L	22	17.7	16.6	17.6	9.9	9	3.6	3.6	7652	0.716
2100/H	6.5	10.4	12.3	15.4	15.9	11.9	10.1	17.5	9328	0.734

* Relative Interdecile Range (RIDR) is measured as the difference between the ninth decile and the first decile divided by the median.

Table 1 consists of two panels. Panel A provides the distributions of future world population size expressed in intervals below 6 billion, 6 to 7 billion, 7 to 8 billion, and so on with the uppermost interval being above 12 billion. The numbers in the cells are the percentages of our 2,000 simulated future population paths. Median population sizes are in column 9. The tenth column contains an uncertainty measure, the relative interdecile range (RIDR) defined as the difference between the ninth decile and the first decile of the distribution divided by the median.

Panel B is based on a division of the 2010 distribution into population paths that were above the median in that year and those that were below it. There are 1,000 observations in each of these subgroups. There are two rows in Panel B for each decade following 2010, one labeled with an “L” and another with an “H.” The “L” rows are the population distributions at the indicated date for the observations that were below the median in 2010 and the “H” rows are from the paths that were above the median in 2010.

One disadvantage of this very simplified example is that the forecasts with jump-off dates in 2000 and 2010 are not exactly comparable. The vital rate paths used in the 2000 forecasts all start at their observed values, while the paths in forecasts that have the 2010 jump-off date have a distribution of starting values. One way of testing the plausibility of this example is to consider the uncertainty of forecasts of various durations based on a jump-off date of 2000 and a jump-off date of 2010. Holding duration constant, the example would be questionable if the uncertainties of N year ahead forecasts were very different depending on whether they were made in 2000 or 2010. When the jump-off date is the year 2000 and a forecast is made for 10 years into the future, the uncertainty measure in 2010 is 0.062, which can be read off the row in Panel A labeled 2010. In the case of a forecast made 10 years ahead based on being below the median in 2010, the uncertainty measure in 2020 is 0.088. This can be read off the row in Panel B labeled 2020/L.

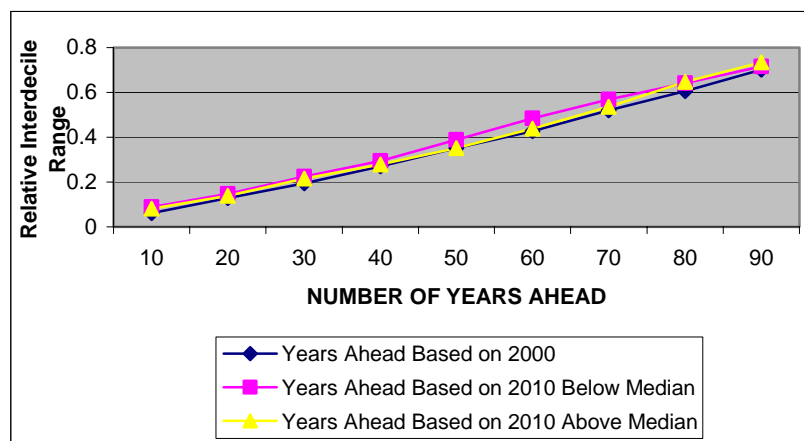


Figure 3. Comparison of the Relative Interdecile Ranges (RIDR) for forecasts made for 10 through 90 years ahead starting from 2000 and starting from an observation either above the median or below the median in 2010. Source: Authors’ calculations.

The uncertainty measures for 10- through 90-year ahead forecasts based on 2000 and the two sub-samples from 2010 are shown in Figure 3. The results from the two 2010 groups track those from 2000 quite well, but are always slightly higher than the uncertainty measures based on 2000. Figure 3 is what is expected given the construction of the example and it suggests that it is plausible to proceed.

The median population forecast for 2100 based on information up to 2000 is 8.414 billion people. After a 10-year wait, the median forecast for 2100, based on being above or below the median in 2010, would either be 7.652 billion or 9.328 billion. It would seem that if anyone were to predict 914 million more people in the world in 2100 from the perspective of 2010, the forecaster must have made a big mistake in 2000. Yet, this could well happen even if the methodology is probabilistically correct. A prediction of a 2100 population size in 2010 that is 762 million smaller than the one that was predicted in 2000 is also easily possible. These are substantial differences. Clearly, forecasts of the future will be different in 2010 than they are today. One interesting feature of probabilistic forecasting is that it can give us some idea about how much different future forecasts could be from current ones, and with what likelihood.

Population size in 2010 has such a persistent effect because of a number of factors. First, past population size influences future population size. Paths that yield large populations in 2010 will also yield large populations in 2100, even if population growth rates after 2010 are the same. Second, some populations are large in 2010 because they had high fertility rates. These high fertility rates alter the age structure of the population making it younger. Younger populations tend to grow more, other things being equal, a process that demographers call “population momentum.” Fertility and mortality themselves have persistence built into them. The persistence of fertility and mortality means that on paths where fertility was high and mortality was low, leading to relatively large populations in 2010, they are likely to remain high and low respectively for a while. The effects of the persistence of fertility and mortality over time are compounded by the relatively high interregional correlations of fertility and mortality, by the persistence caused by population momentum, and by the size effect itself.

It is crucially important that attention be given not only to the effects of the passage of time on the median forecast, but to the entire distribution of forecasted population sizes. Most of the differences in the distributions based on the paths above and below the median in 2010 are in the extremes (tails) of the distributions. For example, 6.5 percent of the paths that were above the median in 2010 resulted in populations of less than 6 billion in 2100, compared to 22.0 percent of the paths that were below the median in 2010. The difference at the high end of the distribution is even more striking. Over 17 percent of the paths that were above the median in 2010 ended the century with 12 billion people or more. In contrast only 3.6 percent of the paths that were below the median in 2010 did so.

This has been a very short and simplified presentation of the basic concepts of conditional probabilistic forecasting with future jump-off dates. It is meant only to be suggestive. This analysis of learning with the passage of time has illustrated how sensitive the long-term population outlook is to near-term trends. It can also help to understand why projections of population size in 2100 have changed so significantly over the past 10 years. We have simply learned a great deal over the past decade.

During this decade population growth has been lower than originally expected and this has significantly decreased our new long-term expectations.

Concluding Thoughts

Conditional probabilistic projections represent a way to combine the benefits of probabilistic projections, particularly the quantification of uncertainty, with the benefits of alternative scenarios, which give clear indications of the sensitivity of results to underlying assumptions. We have shown that the same kinds of conclusions about, for example, the relative importance of fertility and mortality trends to population size outcomes can be drawn using conditional probabilistic forecasts as can be drawn using alternative deterministic scenarios. An added benefit is that the conditional probabilistic forecasts provide an estimate of the likelihood of the underlying demographic conditions, as well as an estimate of their effect on outcomes. These projections can be extremely useful to both the research and policy communities. For instance, many analyses of the potential for long-term environmental change are based on the approach of considering a set of alternative future scenarios conditional on different sets of assumptions about future development. The scenario approach dominates as a response to deep uncertainty about the many socioeconomic, technological, and environmental factors that must be included in such analyses. Conditional probabilistic projections present a possible means of retaining some of the advantages of the probabilistic approach without discarding the benefits of conditional scenarios (see O'Neill 2003 for an example).

Probabilistic projections with future jump-off dates, which are conditional on how population characteristics evolve between now and the future jump-off date, present a new way to address an important set of research questions that are also policy relevant. They provide a means of anticipating how our forecasts might change in the future and how likely those changes appear to be at the moment. These kinds of analyses cannot be done deterministically. Because there can be costs and benefits to changes in the outlook for the future, these projections could have interesting new applications. For example in the climate change issue, the prospects of learning about technological costs, or about physical aspects of the climate system, have been incorporated into analyses of whether it is better to act now or to wait to learn more. However as far as we are aware, no such analysis – for climate change or any other issue – has been performed taking into account the prospects for learning about the outlook for population.

References

- Alho, J. 1997. Scenarios, uncertainty and conditional forecasts of the world population. *Journal of the Royal Statistical Society, Series A (Statistics in Society)* 160(1): 71-85.
- Alho, J. 1990. Stochastic methods in population forecasting. *International Journal of Forecasting* 6(4): 521-530.
- Alho, J. and B. Spencer. 1985. Uncertain population forecasting. *Journal of the American Statistical Association* 80(390): 306-314.

- Keilman, N., D. Pham, and A. Hetland. 2002. Why population forecasts should be probabilistic – Illustrated by the case of Norway. *Demographic Research* 6(15): 408-454. <http://www.demographic-research.org/>
- Lee, R.D. 1999. Probabilistic approaches to population forecasting. Pages 156-190 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.
- Lee, R. and S. Tuljapurkar. 1994. Stochastic population projections for the United States: Beyond high, medium and low. *Journal of the American Statistical Association* 89(428): 1175-1189.
- Lutz, W. 1995. Scenario Analysis in Population Projection. Working Paper WP-95-57. Laxenburg, Austria: IIASA.
- Lutz, W., C. Prinz, and J. Langgassner. 1994. The IIASA world population scenarios to 2030. Pages 391-422 in W. Lutz (ed.), *The Future Population of the World: What Can We Assume Today?* London: Earthscan.
- Lutz, W., W. Sanderson, and S. Scherbov. 1996. Probabilistic population projections based on expert opinion. Pages 397-428 in W. Lutz (ed.), *The Future Population of the World: What Can We Assume Today?* Revised edition. London: Earthscan.
- Lutz, W., W. Sanderson, and S. Scherbov. 1997. Doubling of world population unlikely. *Nature* 387(6635): 803-805.
- Lutz, W., W. Sanderson, and S. Scherbov. 2001. The end of world population growth. *Nature* 412: 543-545.
- Lutz, W. and S. Scherbov. 1998. An expert-based framework for probabilistic national population projections: The example of Austria. *European Journal of Population* 14: 1-14.
- Lutz, W. and S. Scherbov. 2002. Can Immigration Compensate for Europe's Low Fertility. Interim Report IR-02-052. Laxenburg, Austria: IIASA.
- O'Neill, B.C. 2003. Conditional Probabilistic Population Projections: An Application to Climate Change. Interim Report IR-03-051. Laxenburg, Austria: IIASA.
- Pflaumer, P. 1988. Confidence intervals for population projections based on Monte Carlo methods. *International Journal of Forecasting* 4: 135-142.
- Sanderson, W.C., S. Scherbov, W. Lutz, and B.C. O'Neill. Forthcoming. Applications of probabilistic population forecasting. In W. Lutz and W.C. Sanderson (eds.), *The End of World Population Growth in the 21st Century: New Challenges for Human Capital Formation and Sustainable Development*. London: Earthscan.
- Webster, M. 2002. The curious role of "learning" in climate policy: Should we wait for more data? *The Energy Journal* 23(2): 97-119.