Reprint

How to deal with uncertainty in population forecasting?

Guest Editors: Wolfgang Lutz and Joshua Goldstein
How to Deal with Uncertainty in Population Forecasting?

Wolfgang Lutz and Joshua R. Goldstein
Guest Editors

RR-04-009
October 2004
How to Deal with Uncertainty in Population Forecasting?

Wolfgang Lutz
*International Institute for Applied Systems Analysis, Laxenburg, Austria*

Joshua R. Goldstein
*Princeton University, Wallace Hall, Princeton, NJ, USA*

Guest Editors

RR-04-009
October 2004

Research Reports, which record research conducted at IIASA, are independently reviewed before publication. Views or opinions expressed herein do not necessarily represent those of the Institute, its National Member Organizations, or other organizations supporting the work.


All rights reserved. No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopy, recording, or any information storage or retrieval system, without permission in writing from the copyright holder.
# Table of Contents

**Introduction: How to Deal with Uncertainty in Population Forecasting?**
Wolfgang Lutz and Joshua R. Goldstein

1

**Time Series Based Errors and Empirical Errors in Fertility Forecasts in the Nordic Countries**
Nico Keilman and Dinh Quang Pham

5

**Using the Lee–Carter Method to Forecast Mortality for Populations with Limited Data**
Nan Li, Ronald Lee and Shripad Tuljapurkar

19

**Mortality Forecasting and Trend Shifts: an Application of the Lee–Carter Model to Swedish Mortality Data**
Hans Lundström and Jan Qvist

37

**Toward a New Model for Probabilistic Household Forecasts**
Jiang Leiwcn and Brian C. O’Neill

51

**Assumptions on Fertility in Stochastic Population Forecasts**
Maarten Alders and Joop de Beer

65

**Probabilistic Population Projections for India with Explicit Consideration of the Education-Fertility Link**
Wolfgang Lutz and Sergei Scherbov

81

**Simpler Probabilistic Population Forecasts: Making Scenarios Work**
Joshua R. Goldstein

93

**Conditional Probabilistic Population Forecasting**
Warren C. Sanderson, Sergei Scherbov, Brian C. O’Neill and Wolfgang Lutz

157

**Conditional Probabilistic Population Projections: An Application to Climate Change**
Brian C. O’Neill

167
Random Scenario Forecasts Versus Stochastic Forecasts
Shripad Tuljapurkar, Ronald D. Lee and Qi Li

Developing Official Stochastic Population Forecasts at the US Census Bureau
John F. Long and Frederick W. Hollmann
Introduction: How to Deal with Uncertainty in Population Forecasting?

Wolfgang Lutz\textsuperscript{1} and Joshua R. Goldstein\textsuperscript{2}

\textsuperscript{1}International Institute for Applied Systems Analysis, Laxenburg, Austria. E-mail: lutz@iiasa.ac.at
\textsuperscript{2}Princeton University, Wallace Hall, Princeton, NJ, USA. E-mail: josh@princeton.edu

Demographers can no more be held responsible for inaccuracy in forecasting population 20 years ahead than geologists, meteorologists, or economists when they fail to announce earthquakes, cold winters, or depressions 20 years ahead. What we can be held responsible for is warning one another and our public what the error of our estimates is likely to be.—Nathan Keyfitz (1981)

In the pursuit of our daily life, be it at the individual level or as a society, we are constantly guided by expectations about the future. Typically, these expectations are based on the assumption that the future is going to be more or less the same as what we currently experience or that there are clearly predictable regularities such as day and night and the change of seasons.

In some instances, however, we know with near certainty that things will change fundamentally. In the field of demography, it is virtually certain that European societies will get older and the proportion of the population above age 60 will increase significantly and the proportion below age 20 will shrink. This change that we expect over the coming decades is already embedded in the current age structure of the population. Only extremely unlikely events such as a new disease killing large proportions of the elderly population while leaving the younger unaffected could change this aging trend in the near to mid-term future. This nearly certain population aging will bring fundamental changes to the functioning of our societies and will have significant impacts on individual life course planning, family networks, pension systems and macro-economic development alike. In terms of pensions in a pay-as-you go system younger people today cannot count on the same contributions/benefits ratio as they observe with people retiring today.

Even in case of this highly predictable trend towards population aging the exact extent of aging is rather uncertain, especially in the longer run. Will the share of the population over age 60 in Western Europe increase from currently 20 percent to 29 percent or 43 percent by 2050 (the 80 percent prediction interval according to Lutz \textit{et al.}, 2001). This is a very significant difference and the answer will largely depend on the still uncertain degree of future increases in life expectancy but also on the even more uncertain future trends in fertility and migration.

There is uncertainty in all three components of demographic change (fertility, mortality and migration). How should a forecaster deal with this uncertainty in a statistically consistent manner that is both informative to the users and feasible for the producing agencies? This is the topic of this set of eleven papers that are published in two blocks in this and the next issue of the \textit{International Statistical Review}.

Over the past years a rapidly increasing body of literature has dealt with the issue of uncertainty in population forecasting. The \textit{International Journal of Forecasting} published a special issue in 1992 (Ahlburg & Land); \textit{Population and Development Review} published a special supplement in 1999 (Lutz \textit{et al.}); and most recently the National Research Council (2000) has dealt extensively with
uncertainty in its volume Beyond Six Billion. This is because population forecasts are important for a large community of users, forming the basis of social, economic, and environmental planning and policy making. The medium variant projections, typically considered to be the most likely forecasts, that have been produced by national and international agencies have played a useful role and have generally achieved impressive accuracy. Yet, it is increasingly recognized that the way these projections deal with the issue of uncertainty is unsatisfactory.

The current practice of providing “high” and “low” variants to communicate uncertainty around the medium projection suffer from several drawbacks. The most important are: (a) In many cases, variants only address fertility uncertainty, ignoring mortality and migration uncertainty; (b) The variants approach is unspecific about the probability range covered by the “high” and “low” variants; (c) The variants are probabilistically inconsistent when aggregating over countries or regions because the chances of extreme outcomes in many countries or regions at once are portrayed as being the same as an extreme outcome in a single country or region; and (d) The variants typically do not allow for temporal fluctuations such as baby booms and busts that can produce bulges in age structure.

Considerable scientific progress has been made in the field of probabilistic population forecasting, including the analysis of past projection errors, the use of expert knowledge and substantive arguments, and the development of stochastic models of fluctuating demographic rates. Several national statistical agencies, such as Norway, the Netherlands, Finland and Austria, have already published fully probabilistic forecasts and more, such as the United States, are planning to do so.

This set of eleven papers covers many of key issues currently discussed in this field. Without discussing the individual papers in this introduction we will only briefly mention some of the key questions the papers address. We end each item with citations to the relevant papers in this collection.

1. Past projections errors: The fact that population projections have been carried out for a long time can be used to compare the past projections to actual trends (ex post error analysis). This can give an important piece of information as to what can go wrong in population projections. For the future this information can serve as a yard stick of what one would assume to be the minimum error if one wants to be on the safe side, while it is of course problematic to assume that the future errors will be exactly a replication of the past errors. A more difficult challenge is to address the deeper reasons of why forecasters in the past have made certain erroneous assumptions and how we can learn from this for our new forecasts. (Keilman & Pham).

2. Trends in the components: Demographic forecasts typically treat uncertainty in each component of population change—fertility, mortality, and migration—separately. An active area of research has been to develop models for projecting each of these components. In mortality, the dominant approach in forecasting has been to rely on formal models extrapolating historical trends. In fertility and migration, where the patterns of determination have been more complex, the role of judgment has been greater. The challenge remains two-fold: on the one hand to incorporate substantive knowledge into formal models, and at the other extreme to formalize the use of substantive knowledge. This also holds for projections that go beyond age and sex through considering, for instance, household size. (Leiwen & O’Neill; Lundström & Qvist; Li et al.; and Alders & De Beer).

3. The role of experts in defining uncertainty: Experts play a key role in all population forecasting but their tasks depend on the chosen approach. In some instances the role of the expert is limited to choosing the model, including some of its key parameters and the reference data from which the estimates should be drawn. In other instances experts do also make assumptions about the likely future level of the demographic components (such as the average future fertility level) or limit the range of future values (minimum or maximum fertility) while using time series data to define the fluctuations within this range. In still other cases the experts also define the variance. Experts can also go further and include into their considerations some of the structural drivers of future fertility trends such as changes in the educational composition.
4. **Definition of the temporal process:** Traditional population projections have assumed smooth, monotonic paths for future trends in all three components but most importantly in fertility. This is also mostly the case for high and low fertility variants which assume (piecewise) linear trends to different target fertility levels. Fully stochastic projections on the other hand incorporate annual fluctuations in all three components which seem to better reflect the pattern seen in the real world. An interesting research question is to what degree piece-wise linear scenarios—following the tradition of most statistical agencies—can be seen as a stylized approximation to the results of fully stochastic models with annual fluctuations. There are currently two schools of thought: on the one hand, the two methods cannot produce perfectly identical results and when the variances of the underlying processes differ can produce vastly different results. On the other hand, it does appear that for some purposes the scenario approach can produce results that are close to the fully stochastic approach. Further work on this issue will be important as statistical agencies try to improve their description of uncertainty in forecasts and will influence whether they take a fully stochastic approach or transition incrementally from scenarios. (Tuljapurkar *et al.*; Goldstein).

5. **Conditional probabilistic forecasts:** While probabilistic population projections give a comprehensive description of the uncertainty range they do not provide users with information about the consequences of specific alternative fertility, mortality or migration trends. Especially, in the case of policy analysis, decision makers often want to know what will be the consequences of altering a demographic component by some degree. For this reason, some producers of probabilistic forecasts have also separately produced if-then scenario projections. The concept of conditional probabilistic projections makes this separation unnecessary because it allows the analyst to look at the range of uncertainty in the results conditioning on some specific subset of the full range of stochastic trends. (Sanderson *et al.*; O’Neill).

6. **Implementation by Statistical Agencies:** Official agencies must not only produce forecasts that are scientifically defensible but must also assure that their descriptions of uncertainty can be understood by the broad and heterogeneous set of users. As probabilistic forecasts become more common, agencies can begin to learn from each other’s experience, but for now each producer is left to innovate on their own. The failures and successes of these endeavors if communicated appropriately, are likely to provide useful lessons for other countries around the world. (Long & Hollmann).

Not all demographers and statistical agencies are enthusiastically embracing probabilistic population projections for a number of substantive and institutional concerns. On the substantive side one concern raised is that the statement of precise uncertainty ranges (e.g. 95 percent intervals) conveys to the users a misleading sense of precision, as if one would have more detailed information than one actually has. In this context it is important to clearly tell the users that the stated uncertainty ranges should not be seen as precise objective probabilities but rather as indicative ranges depending on the specific model and parameter assumptions made according to the best judgement of the producers. A related substantive concern expressed (Lesthaeghe, 2002) is that probabilistic population forecasts tend to be “too mechanistic” and disregard most substantive scientific knowledge about the determinants of future mortality, fertility, and migration. This reservation does not concern all approaches to probabilistic projections equally. It applies more to models that are exclusively based on the extrapolation of past time series and the application of errors observed with passed projections to estimate future errors than it does to approaches based on substantive argumentation of experts about future trends and the relevant sources of uncertainty. Since this set of papers includes contributions choosing these different approaches the reader can make his/her own judgement about the relevance of this criticism for the approaches concerned.

There are also two reservations of a more institutional nature. Some people claim that probabilistic
forecasts are too difficult and too complex to be understood and they point to the fact that most users of forecasts do not even consider the usual high and low variants and would be less likely to use full distributions. In response one can say that clearly for a large group of users there is for good and understandable reasons only interest in one forecast that is considered the most likely one. For the smaller group of users that in their applications do have to worry about the possibility of deviations from this most likely path, however, one can also make the point that they often do not use the high and low variants because they do not really understand what these variants stand for. And because of the above described problems and inconsistencies of the high and low variants there is no way to understand what they stand for. Instead, it seems to us that people interested in uncertainty should either use completely probability-free scenarios for sensitivity testing or fully probabilistic projections for a more comprehensive view of the uncertainty involved.

A final institutional concern is that probabilistic population projections are too difficult to be implemented by statistical agencies that do have a lack of skilled manpower in this field. This is a genuine concern for many countries and international agencies alike. The solution to this problem can come from two sides: on the one hand demographers can try to develop simpler approaches to probabilistic forecasting that as much as possible build on established expertise and procedures in statistical agencies and on the other hand those agencies can involve in continued training in order to be able to apply these increasingly powerful and useful new ways of dealing with uncertainty in population forecasting.

The papers published in these two issues of the International Statistical Review have originally been presented at a seminar held in Vienna in December 2002 (sponsored by the Vienna Institute of Demography and the International Institute for Applied Systems Analysis, IIASA), and have been extensively refereed. The participants in this meeting also worked on a consensus statement from which some of the above mentioned points are taken and which concludes with the sentence: “We believe that the quantification of uncertainty will enhance the usefulness of population projections and make the work of forecasting agencies an even more valuable product for planners, policy-makers, scientists, and the public around the world.”

References
Ordering Information

Orders must include the publication number and should be sent to the Publications Department, International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria.

Telephone: +43 2236 807
Telefax: +43 2236 71313
E-mail: publications@iiasa.ac.at

A full list of IIASA publications is available at www.iiasa.ac.at/Publications