UK national population projections in perspective: How successful compared to those in other European countries?

Nico Keilman
Department of Economics
University of Oslo

Compared to population forecasts of other European countries, those made in the United Kingdom during the past 30 years had somewhat larger forecast errors for fertility and smaller errors for mortality. Migration forecasts in the UK were about as accurate as the European average. After controlling for various effects such as relative data volatility both at the time a projection is made and during the period of the projection, there is no indication that recent forecasts in European countries have been more accurate than older ones. Hence population forecasts are intrinsically uncertain, and a forecast for the UK in the form of probability distributions is presented.

Introduction

A recent article in this journal has analysed the accuracy of national population projections made for the United Kingdom over the last fifty years (Shaw 2007[1]). The author concluded that the total UK population has been projected reasonably accurately, but that this is largely a chance result of compensating errors in the assumptions of fertility, mortality and net migration. Fertility was overpredicted, mortality assumptions in the projections were too pessimistic (that is, life expectancies were too low), and net migration to the UK was underpredicted. In terms of the age structure, the largest errors were for the very young and the very old, while projections of the working age population have been quite accurate. Fertility and mortality errors have reduced in more recent projections, while migration errors have grown.

There is a growing literature, in which national population projections are evaluated *ex-post facto* against observed statistics (for instance Preston 1974[2]; Calot and Chesnais 1978[3]; Inoue and Yu 1979[4]; Keyfitz 1981[5]; Stoto 1983[6]; Pflaumer 1988[7]; Shaw 1994[8]; Keilman 1997[9], 2000[10], 2001[11]; National Research Council 2000[12]; Keilman and Pham 2004[13]). These studies have shown, among other things, that projection accuracy is better for short than for long forecast durations, and that it is better for large than for small populations. They also demonstrated that there are considerable differences in accuracy within a projection, that is, when broken down by region or other components. Because of the extrapolative nature of population projections, fluctuations in observed fertility, mortality and migration time-series are associated with large projection errors. Finally, poor data quality tends to go together with poor projection performance. This relationship is stronger for mortality than for fertility, and stronger for short-term than for long-term projections.

Shaw's article contributes to this literature. The aim of the current article is to put his results into international perspective. How do the findings for the UK compare with those in other countries? I shall report from a research project, which evaluated the accuracy of national population projections prepared by the statistical agencies of 14 European countries. The countries are Austria, Belgium, Denmark, Finland, France, Germany[14], Italy, Luxembourg, Netherlands, Norway, Portugal, Sweden, Switzerland and the United Kingdom. The projections were made during 1950–2001, and I have evaluated projection results for the years 1950–2002 against actual data. Projection results were assembled from published and unpublished sources for the 14 countries. Actual data for fertility, mortality, and migration are available from Council of Europe publications (CoE, 2002[15]). In some cases these data have been supplemented by other sources. Details are available in Keilman and Pham (2004[13]). The recently created data base for UK projections (Shaw, 2007[1]) was not available when the project was finished. Also, I shall refer briefly to results reported for a few non-European countries: Australia, Canada, Japan, and the United States.

Box one

Terminology

In this article, I look at how the projections have performed as predictors of future population change. I have analysed the error in the projections of future total fertility rate, life expectancy at birth and total net migration according to forecast duration. Two main related measures are considered, the mean error and the mean absolute error.

The projection error is calculated to be the projected value of a variable minus its actual value as currently estimated. (These 'actual' values may, of course, in some cases have been revised or be subject to further revision, for example, following Censuses.) A positive error is, therefore, an overprojection, that is, the projected value exceeded the actual value and a negative error is an underprojection.

The forecast duration is the difference between the base year of the projection and the calendar year for which the particular variable is projected. For example, the 1981-based projection of the total population at the year 2001 has a forecast duration of 20 years. For each projection error and forecast duration, we have a series of observations. So if we are considering the accuracy of projections of the total population 20 years ahead, we may have a 1950-based projection for 1970, a 1951-based projection for 1971 and so on. The actual number of observations obviously depends on the frequency with which projections were carried out in the respective countries. It also depends on forecast duration: we have fewer observations for long forecast durations than for short ones.

The mean error (ME) is the average of the projection errors for a particular set of observations and the mean absolute error (MAE) is the average of the projection errors irrespective of sign. For example, suppose for a particular variable in a particular country, that we only have observations from the 1951-based, 1961-based and 1971-based projections for the projection error 30 years ahead. If these errors were +10 per cent, +5 per cent and –3 per cent respectively, then the mean error is 4 per cent and the mean absolute error is 6 per cent. The mean error gives us a measure of bias; it tells us that, on average, this variable was overprojected by 4 per cent. The mean absolute error gives us a measure of precision; it tells us that, on average, the difference between the projected value and the actual value was 6 per cent. In this article I only present these mean errors where I have a minimum of ten observations.

I shall use the terms 'projection', 'forecast', and 'prediction' interchangeably. Although demographers interpret them slightly differently, most users do not, and hence I consider them as equivalents.

The results must be interpreted with care. My intention is not to argue that differences in accuracy between countries necessarily imply that forecasters in some country were more skilled than those in other countries. Large or small errors may simply indicate the ease or difficulty of carrying out a successful forecast. The degree of success may depend on a number of factors, such as the quality of the data that are available, the volatility in the actual trends, policy measures unexpectedly introduced by authorities, sudden political events that have an impact on demographic developments, etc.

I shall present summary statistics for the errors in projection assumptions for fertility, mortality, and net migration. In many cases, I use the Mean Error (ME) and the Mean Absolute Error (MAE). The ME reflects the bias in the assumed values, while the MAE indicates their precision. This terminology (see Box one for further details) is the same as that used by Shaw in his earlier article on UK projections. For fertility, I use the Total Fertility Rate (TFR), for mortality the life expectancy at birth of men and women, and for international migration the level of net migration.

Fertility

Figure 1 shows the precision of the fertility assumptions by plotting the mean absolute errors (MAE) in the TFR for the 14 European countries. For each country, the mean is computed across several forecast rounds, controlling for forecast duration. Each line represents one country. For long forecast durations there were fewer observations than for short ones. Therefore, in order to ensure observations are robust, I plotted in this and the following graphs, for each country, only those mean values that were based upon at least ten observations. The marked solid line gives the MAE value for the pooled set of errors, that is, the errors for all countries and forecast rounds combined, including those observations that were left out from the country-specific means. The latter curve is based on 295 observations for one year ahead and 203, 94, and 16 observations at durations of 10, 20, and 30 years ahead, respectively. For the UK I have observed errors for projections made during the period 1970–2000 (results for earlier projections are now available from the data base for UK projections (Shaw, 2007[1]) but, as noted above, this was not available when the analyses for this article were done).

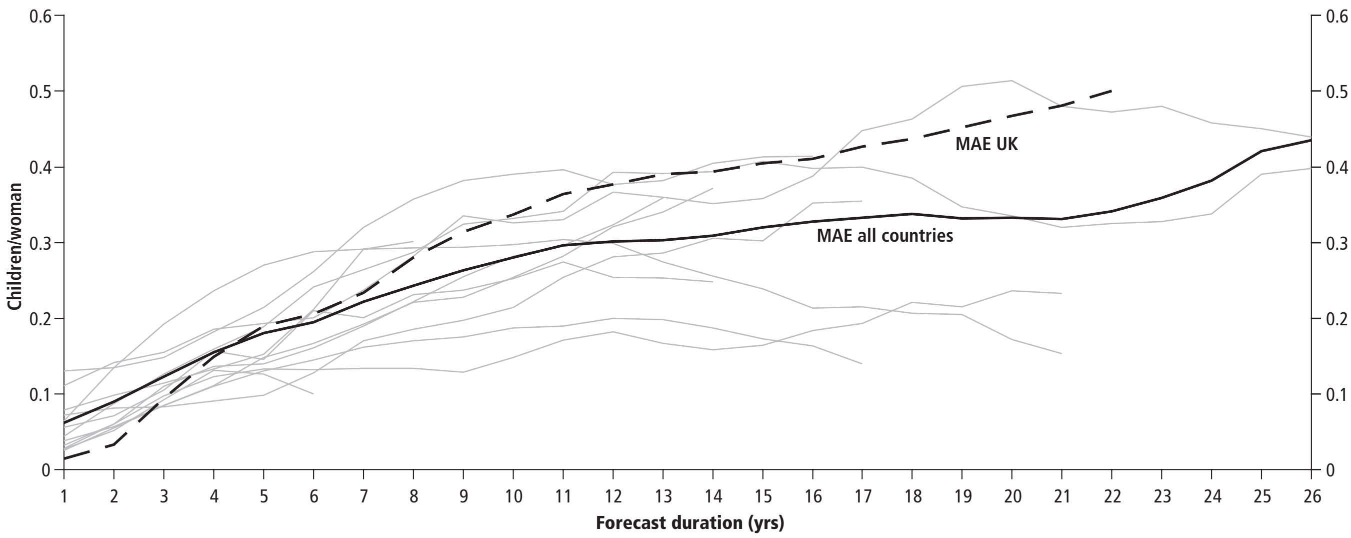
The pattern that emerges is that of slowly increasing errors. Short-term forecasts are more precise than long-term forecasts, because the likelihood that the factors which influence fertility may change is obviously likely to increase with increasing forecast duration. Across all 14 countries and all forecasts, the mean absolute error increases from 0.06 children per woman in the first year of the forecast, to 0.3 for a forecast horizon of 15 years ahead, and 0.4 children per woman 25 years ahead. Although the patterns for the individual countries vary greatly around the mean of the pooled errors, they are roughly consistent with the overall picture. The figure illustrates that, except for the first few years of the projection, MAEs for the UK are above the average. This indicates that TFR assumptions for the UK might have been more difficult to formulate than on average for the 14 countries.

Figure 2 shows that TFR forecasts after World War II in the 14 countries were too high overall (that is, across all countries and all forecast rounds). The mean error (ME) in the TFR is negligible in the first forecast year, but it grows regularly to a little over 0.4 children per woman 25 years ahead. This pattern reflects the well known fact that fertility was overpredicted in projections made in the late 1960s and the 1970s, a period when actual birth numbers fell rapidly throughout Europe. Indeed the mean error is only slightly lower than the mean absolute error in Figure 1, reflecting the fact that almost all errors were positive, that is, TFRs were consistently overpredicted. Two countries, Finland and Denmark, differ from the international pattern, in that they underpredicted their TFR by 0.1 children per woman 15–20 years ahead. The reason is that the observation period for these two countries starts when fertility had already reached a low level: 1972 in Finland and 1974 in Denmark. For short durations, the mean errors for UK projections are close to the international average, but in the medium and long run they are about 0.2 children per woman higher.

Mortality

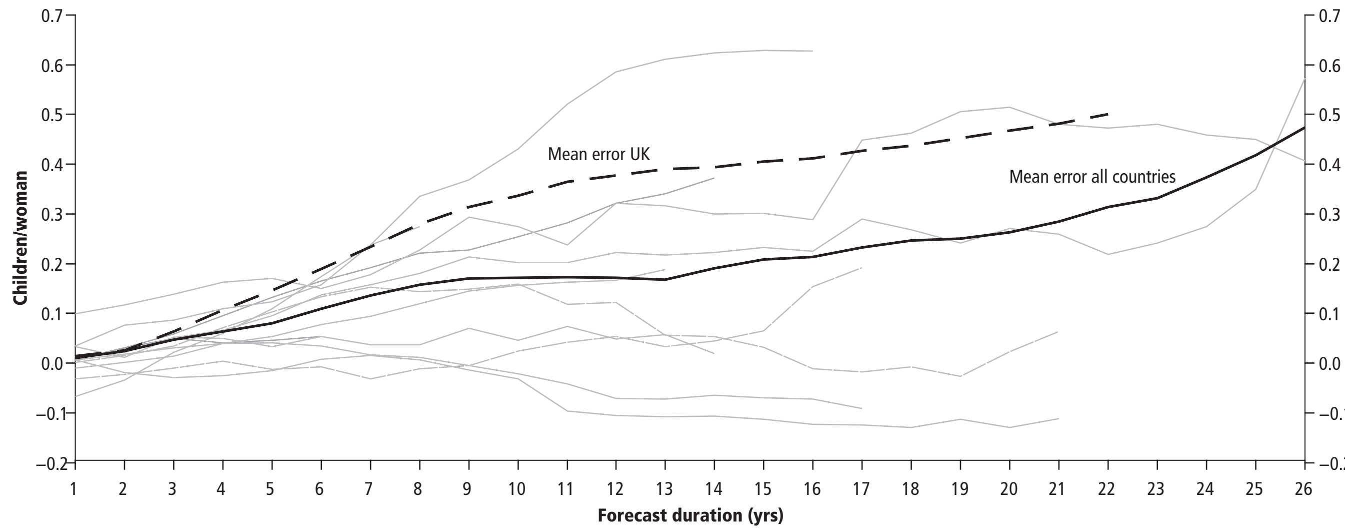
Precision and bias for assumptions on the future life expectancy at birth are illustrated in Figures 3 and 4. The first graph plots the means of the absolute errors in the life expectancy at birth of men for the 14 European countries. Some countries had fewer than ten observed errors, even at the shortest forecast duration. There is no individual line for these countries, but their errors are included in the mean of the pooled errors. By way of exception, I included the UK (dashed line), although I had only

Figure 1 Observed mean absolute error for TFR in 14 countries



Each unlabelled line represents one country. Data only shown where there are ten or more observations (unless otherwise indicated in text)

Figure 2 Observed mean error for TFR in 14 countries



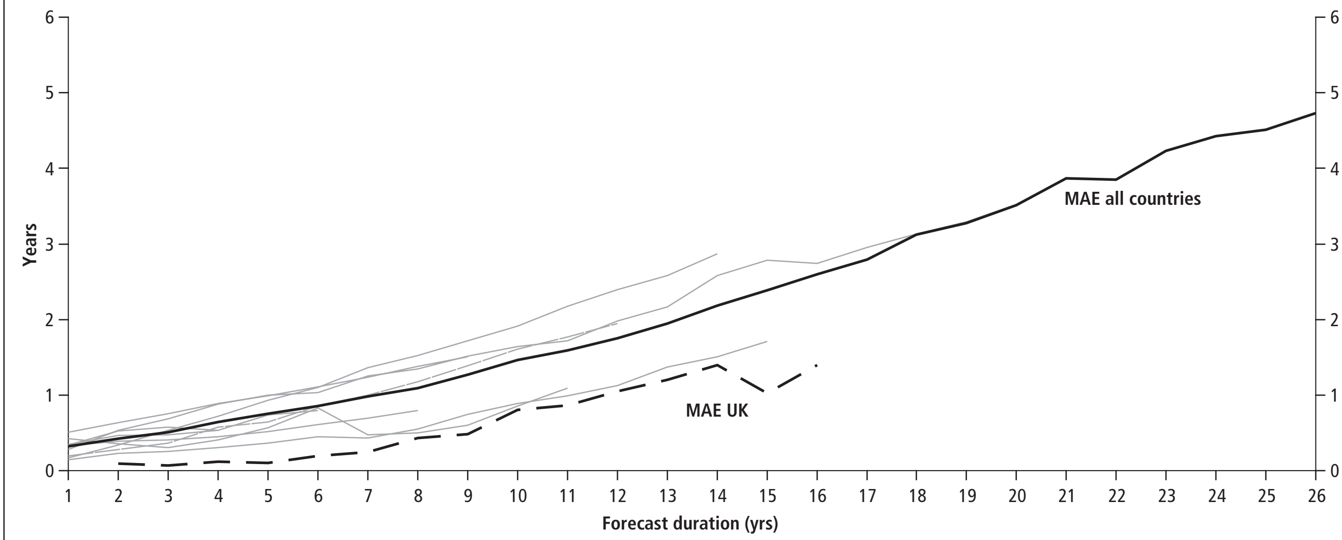
Each unlabelled line represents one country. Data only shown where there are ten or more observations (unless otherwise indicated in text)

up to eight observations at each duration. The figure shows a slightly accelerating growth in inaccuracy by forecast horizon, with errors in the life expectancy at birth increasing by 0.2 years per year for forecast horizons 10–25 years, but somewhat slower rising errors for shorter durations. UK errors are among the lowest. The line is quite stable, in spite of the few observations.

The negative mean errors in Figure 4 indicate that life expectancy forecasts have been too low on average. Across all 14 countries, the underprediction amounted to 1.4 and 3.4 years of life expectancy at forecast horizons of 10 and 20 years ahead, respectively. This confirms earlier findings for selected industrialized countries (Keilman 1997[9]). Again, the errors for the UK are relatively small.

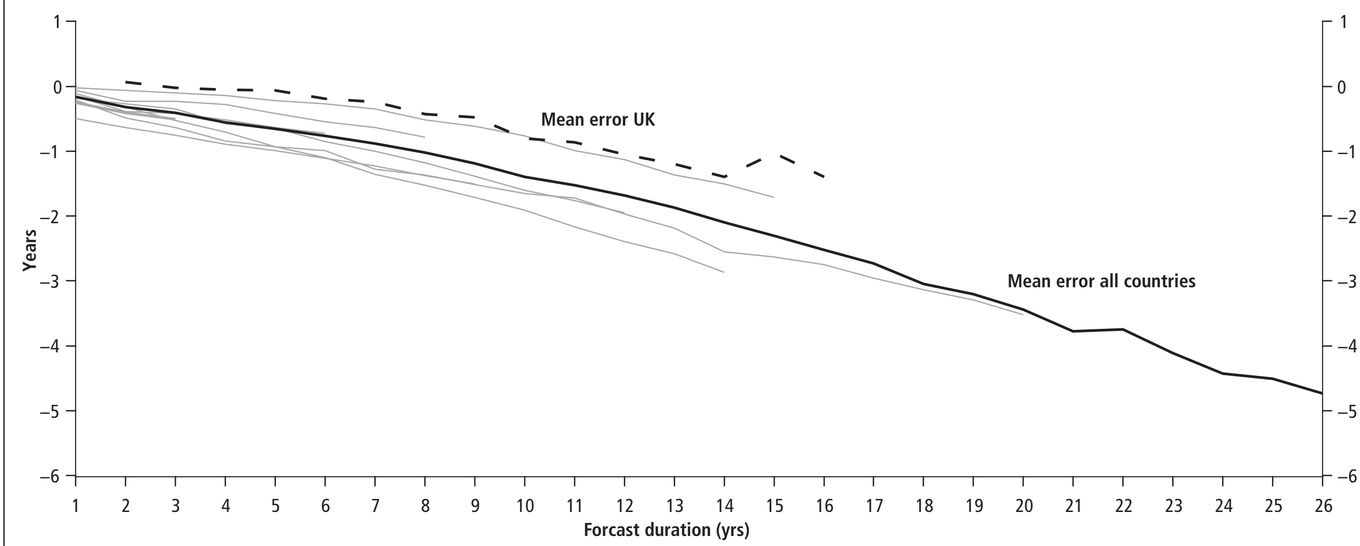
Similar graphs for women are not included here, because the general patterns for women are very similar to those for men, both for the MAE and the ME. However, for the UK, Shaw (2007[1]) noticed that life expectancy errors were greater for men than for women. This is also the

Figure 3 Observed mean absolute error for life expectancy at birth for men in 14 countries



Each unlabelled line represents one country. Data only shown where there are ten or more observations (unless otherwise indicated in text)

Figure 4 Observed mean error for life expectancy at birth for men in 14 countries



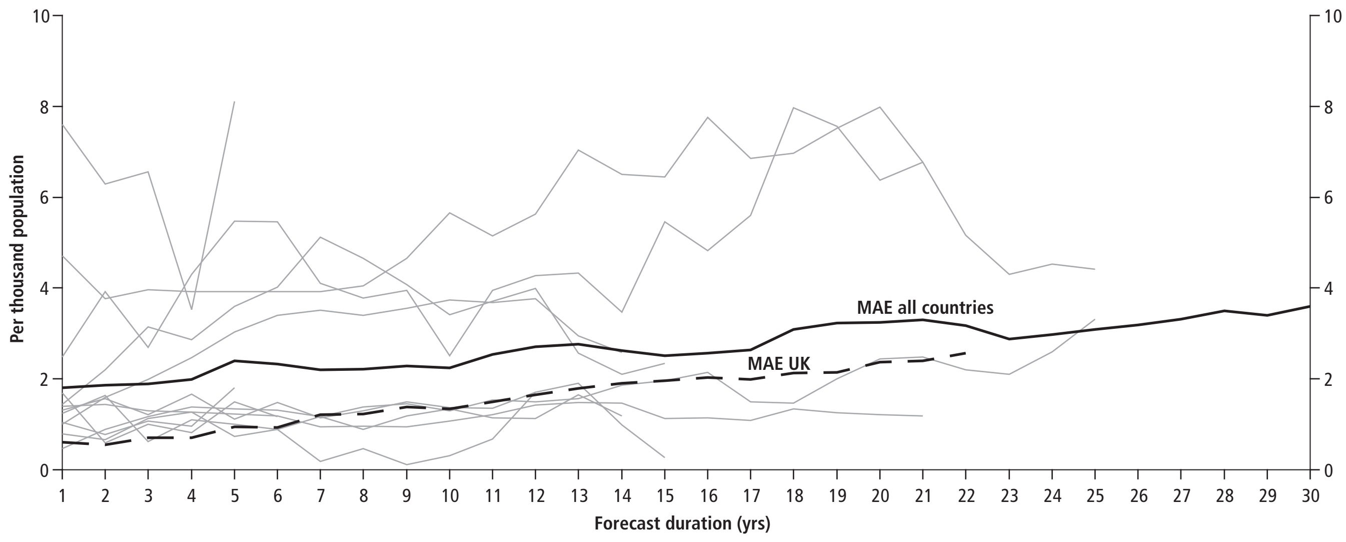
Each unlabelled line represents one country. Data only shown where there are ten or more observations (unless otherwise indicated in text)

case for a few other countries, in which actual mortality improved faster for men than for women after 1970 (Keilman and Pham 2004[13]). This effect disappears when the average error is computed for all forecasts (starting in 1950) in the 14 countries. But, in the multivariate analysis of forecast errors discussed below, after controlling for a number of disturbing factors, there was an independent effect of larger forecast errors in the life expectancy of men compared to women.

Migration

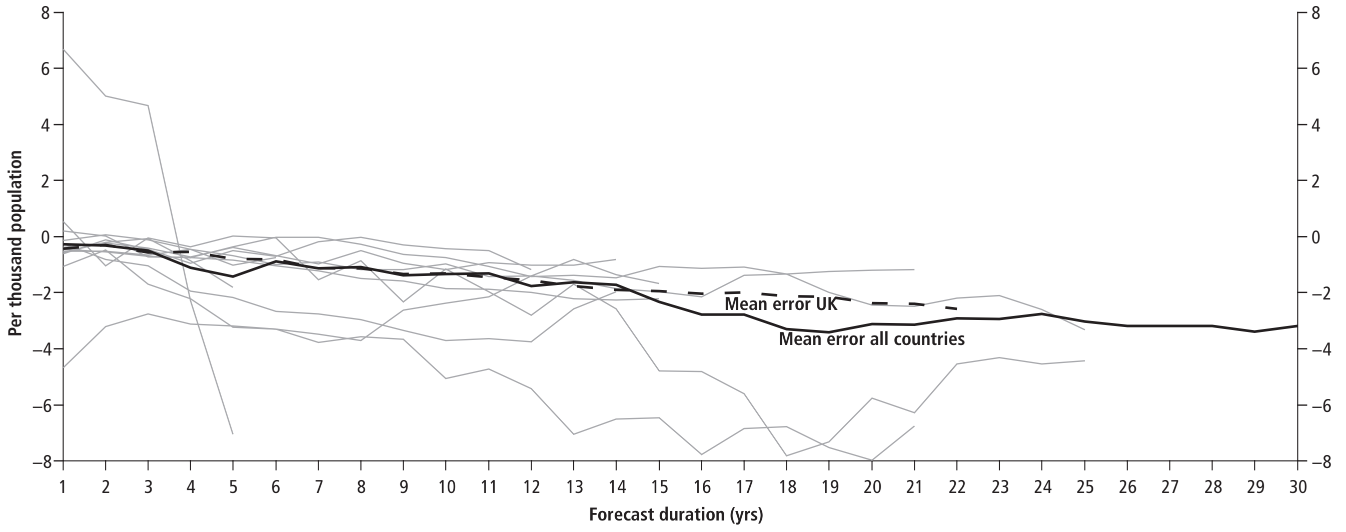
Net migration is defined, for a certain year, as the number of immigrants minus the number of emigrants. To facilitate comparison across the 14 European countries, I have scaled all national migration numbers by the national population sizes as of 1 January 2000. The latter data were taken from the Council of Europe's international data base (CoE 2002[15]).

Figure 5 Observed mean absolute error in migration forecasts in 14 countries (scaled per 1000 population as at 1 January 2000)



Each unlabelled line represents one country. Data only shown where there are ten or more observations (unless otherwise indicated in text)

Figure 6 Observed mean error in migration forecasts in 14 countries (scaled 1000 population as at 1 January 2000)



Each unlabelled line represents one country. Data only shown where there are ten or more observations (unless otherwise indicated in text)

Thus the unit of measurement is 'net migration per 1,000 population' (note that the denominator is not a population at risk). All averages were computed based on these scaled numbers. As an alternative, I could have used the population in the year in question as the scaling factor. However, I opted for a fixed factor based on the population in the year 2000 because the time series were also to be used in predictions of future migration levels (Keilman and Pham, 2004[13]). A fixed factor facilitates scaling back to the original units of migration.

Many historical projections ignored migration. I have assumed that the implicit assumption for those projections was a net migration level of zero. In these cases, the signed error is simply the negative of the observed level of net migration (that is, if the observed actual level was +10,000, the projection error was –10,000). The assumption is justified, because users are likely to have interpreted these projections as proper forecasts, reflecting plausible future demographic developments.

Figure 5 shows that precision deteriorates slowly. The mean absolute migration error for all countries taken together rises from just under two per thousand in the first forecast years, to about three to four per thousand at forecast durations of twenty years or more. For a country of the size of the UK (approximately 60 million inhabitants in 2000) the international average would imply an absolute error of 120,000 in the first forecast years rising to about 240,000 migrants per year in thirty years' time. At the other end of the spectrum, for a country the size of Luxembourg (436,000 persons in 2000), the error would be only approximately 900 to 1,700 migrants if the international average applied.

There are two distinct groups of countries. One group, consisting of the Nordic countries, Belgium, Netherlands, and the UK have relatively small errors. The fact that the UK belongs to this group (with otherwise small populations) may possibly be explained by a population size effect. This is discussed further below. The other group consists of Austria, West Germany, Luxembourg, Portugal, and Switzerland, with mean errors well above the average for the pooled data set. The forecasts of Germany could be expected to be less accurate than the international average, because of large immigration flows after the fall of the Berlin Wall in 1989. For Austria and Switzerland the large errors may be due to the fact that many forecasts for these countries ignored migration (Austria: those with base years 1962, 1978, 1980, 1981, 1982, 1983; Switzerland: those with base years 1951, 1956, 1964, 1967). The large errors for Portugal may be partly explained by major revisions to the migration estimates (Council of Europe 1998[16], 2002[15]).

Migration has been consistently underestimated in historical forecasts. In a number of cases, the error is negative simply because migration was omitted, and the actual migration level was positive, as noted above. In other cases, the assumption was just too low. Figure 6 shows that the mean error in scaled migration reached minus three per thousand after 20 years and then stabilized around that level. The UK error was close to the international average, at least for a forecast horizon up to 15 years.

Forecast accuracy for other industrialized countries

Accuracy evaluations have been reported for fertility and mortality forecasts for a few other industrialized countries, notably Australia (Adam 1992[17], Wilson 2007[18]), Canada (Preston 1974[2], George and Nault 1991[19]), Japan (Preston 1974[2], Feeney 1990[20]), New Zealand (Preston 1974[2]), and the United States (Preston 1974[2], Ahlburg 1982[21], Long 1987[22], 1995[23], Mulder 2002[24]). I refer to them only briefly, for two reasons. First, many of those studies are more than ten years old. Second, unlike the study for the 14 European countries, these other evaluations are difficult to compare, since they use different measures for forecast accuracy, they analyse the accuracy of different demographic variables, they refer to very different time periods, etc.

The analyses of these countries both for fertility and for mortality confirm the findings for the 14 European countries. Fertility assumptions in forecasts made in the 1960s for Australia, Canada, and the United States did not anticipate the steep fall in the 1970s. For instance, the Crude Birth Rate (CBR) in the United States Census Bureau (USBC) forecasts made in 1963 and 1966 were about 55 per cent too high after only ten years (Mulder 2002[24], Table 6). Crude Death Rates (CDR) show smaller errors in the assumptions than Crude Birth Rates, because CBR errors are affected both by assumptions about the **level** of fertility (how many children do women have on average?) and the **timing** of fertility (at which age do they have their children?). CDR errors are only affected by errors in assumptions about the timing of mortality: at which age do individuals die on average? Thus the CDR assumptions in the 1963 and 1966 forecasts of the USBC after ten years were too high by 3 and 8 per cent, respectively. Later USBC forecasts show larger errors in the CDR assumption, but these are 13 or 14 per cent at most (forecasts made between 1969 and 1976; see Mulder 2002[24] Table 9).

Preston (1974[2]) evaluated forecasts of age-specific mortality made around 1950 and 1960 in Australia, Canada, Japan, New Zealand, and the United States. He concluded that for most age/sex groups, mortality declined faster than was foreseen in the forecasts, in particular for women. Indeed, errors were often large since mortality was generally assumed to remain constant in those old forecasts. Forecasts made in the 1970s and later have usually assumed a fall in future death rates, but the decline has generally been underpredicted.

Are recent projections more accurate than older ones?

Given the scientific progress in population studies and in demography, one might expect that recent projections would be more accurate than older ones. Nowadays, we have better data, more sophisticated methods of analysis, and more refined theories of demographic behaviour than 30 years ago (Crimmins 1993[25]; Preston 1993[26]). The demographic and statistical literature witnesses a continuous accumulation of knowledge. This should have had positive consequences for the accuracy of demographic forecasts, other factors remaining the same.

However, in a recent analysis I found that the projections published by the statistical agencies in the 14 countries have not become more accurate over the past 25 years (Keilman 2007[27]). Scientific progress in population studies during the past two to three decades may have been too slow to keep up with less predictable demographic behaviour of populations in those countries. I shall summarize the results here, and compare my findings for the UK with those for other countries.

Recent forecasts have had a shorter life time than older ones and Figures 1 to 4 show that forecast errors increase with forecast duration. So, if one wants to compare the accuracy of recent forecasts with that of older forecasts, it is necessary to control for forecast duration. My statistical analysis took this fact into account. It also allowed for the fact that demographic behaviour may be more difficult to predict at some points in time than others. An example is the 1960s and 1970s, when the baby boom suddenly came to an end in many countries, and fertility fell unexpectedly fast. Finally, I controlled for the possibility that there are specific country circumstances which may make the population in some countries easier to project than in others.

For fertility, I estimated a multivariate model with the absolute error in the Total Fertility Rate as the dependent variable, with similar models for mortality and migration[28]. Forecast base years are included as independent variables in the form of dummy variables. I also control for the effects of forecast period, forecast duration, and country. The statistical analysis results in an indicator for each base year, which summarizes forecast errors for the forecasts produced in that year. An increase in accuracy should be reflected in lower values of the forecast error indicator for recent base years than for earlier base years.

After controlling for all of these possible effects, the analysis suggests that the accuracy of fertility forecasts improved until the end of the baby boom but, since the mid 1980s, has not improved further. Also, there has been no improvement in the accuracy of mortality forecasts after the end of the 1970s. I found a moderate decrease in the accuracy of migration forecasts after the early 1950s followed by a slight increase in the 1980s and 1990s. But the recent improvement is not statistically significant.

These results suggest that scientific progress in population studies during the previous two to three decades might have been too slow to keep up with less predictable demographic behaviour of populations in European countries. The emergence of consensual unions, the delay of childbearing, the fast increase in life expectancy, and ever-growing migration flows to European countries are some of the relevant factors here (Keyfitz 1982[29]). The lesson is that population forecasts are intrinsically uncertain, and hence should be couched in probabilistic terms. I shall come back to the issue of probabilistic forecasts at the end of this article.

Are projections more accurate in some countries than others?

The statistical model includes a country effect. This effect reflects the particular circumstances in the various countries which may relate to the production of a population forecast: for example, the quality of the available data, the relative volatility of demographic trends, government policies which may have subsequently influenced demographic trends, the number and the skills of the forecasting staff, and so on. In addition, the country effect captures the fact that large populations may be easier to forecast than smaller populations, other things being equal[30]. Such a size effect has been demonstrated in forecasts of total population size, both at the national and sub-national level (Smith *et al* 2001[31], NRC 2000[12]). Figures 7, 8, and 9 present the estimated country effects for precision in

Figure 7 Country effects for errors in TFR forecasts

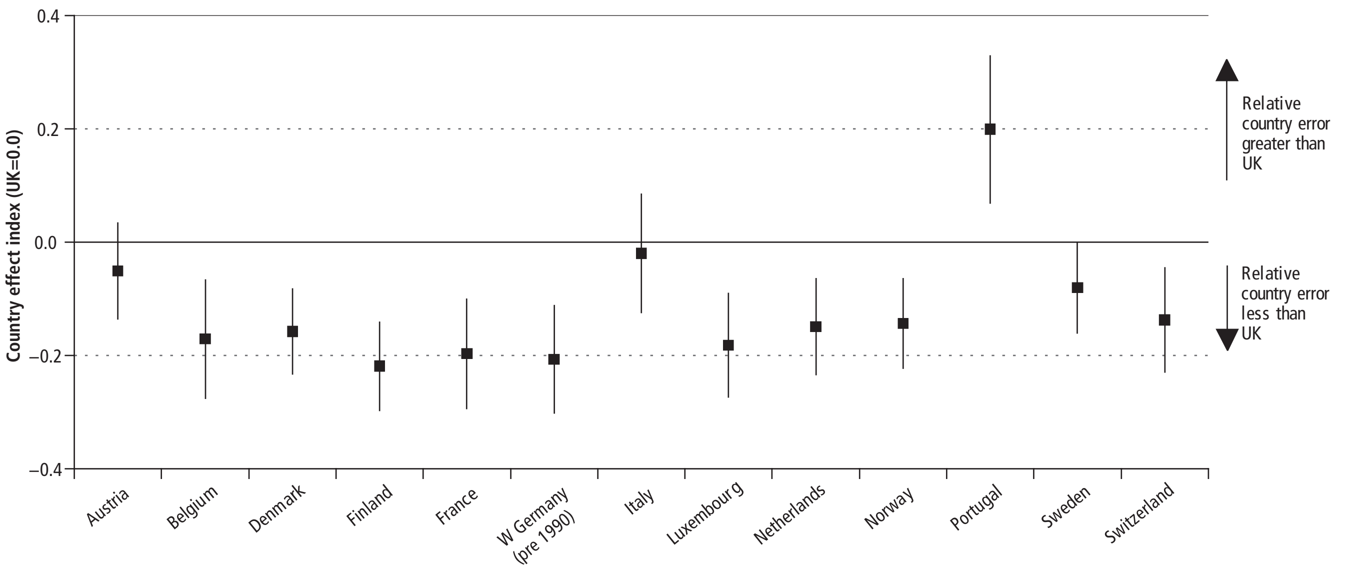


Figure 8 Country effects for errors in life expectancy forecasts

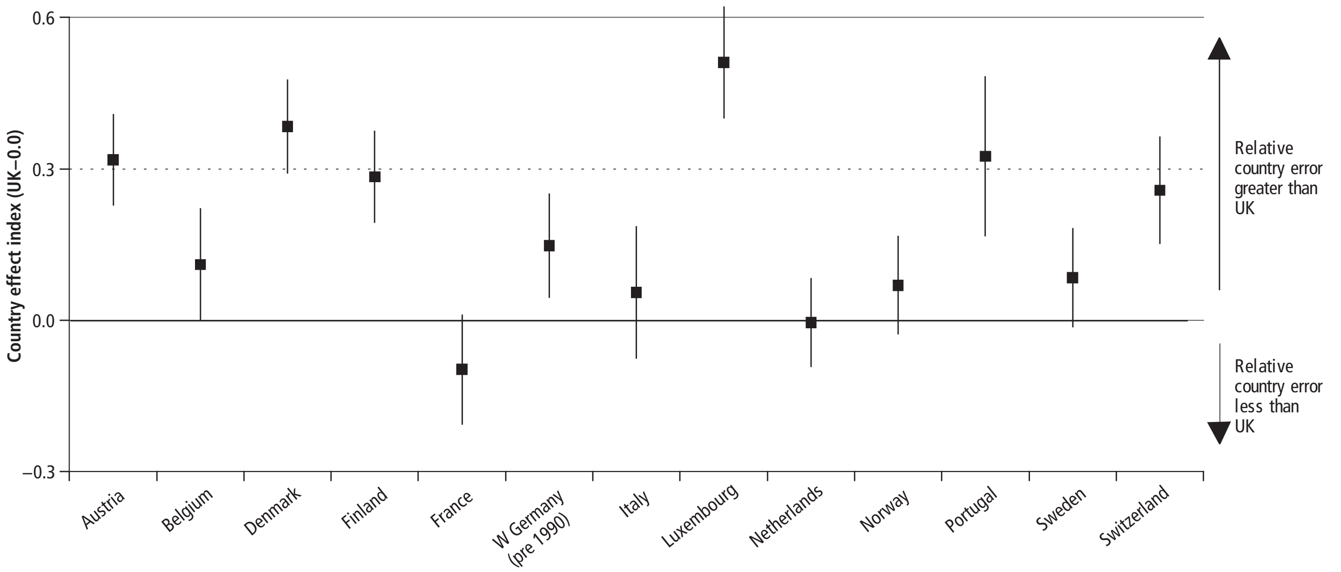
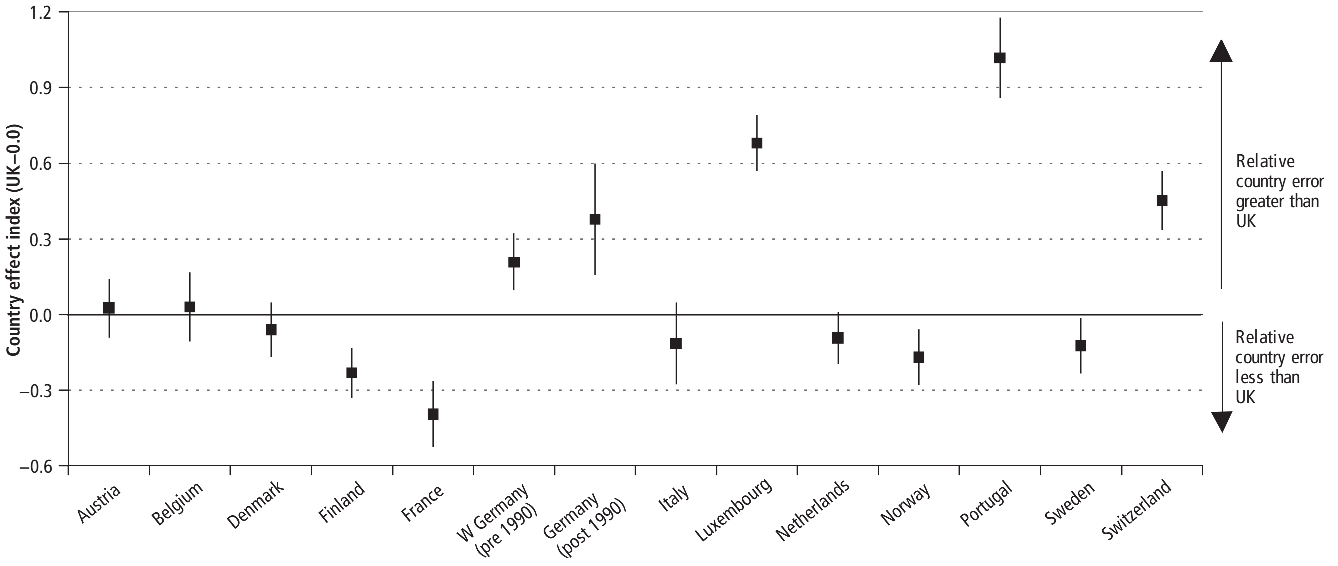


Figure 9 Country effects for errors in net migration forecasts



the assumptions for the TFR, life expectancy at birth (men and women combined), and net migration (scaled by population size). The estimates for each country are shown with the UK as the reference category. Any value larger than zero indicates that for the country concerned (having controlled for base year, forecast period and forecast duration), the country effects were on average larger than those for the UK. A value lower than zero indicates that the country effects were lower than those in the UK. The square marks are the estimates, while the vertical lines represent the corresponding 95 per cent confidence intervals for the estimates.

The multivariate analysis for fertility confirms the earlier findings of the descriptive analysis (Figure 1). In many countries it was easier to predict fertility accurately than in the UK. In Austria, Italy, and Sweden it was as difficult as in the UK to predict fertility accurately, since the estimates for these countries are not significantly different from zero. Only in Portugal was the country effect significantly greater than in the UK.

Mortality, in terms of the life expectancy at birth, was easier to predict accurately in the UK than in most other countries, see also Figure 3. The estimated country effect for France in Figure 8 is below zero, but the difference is not statistically significant. Also, the effects for Belgium, Italy, the Netherlands, Norway and Sweden are not significantly different from that of the UK.

Figure 9 shows that it was significantly easier to predict migration accurately in Finland, France, Norway, and Sweden than it was in the UK. In Germany (pre and post unification)[14], Luxembourg, Portugal and Switzerland it was more difficult.[32] A theory that there is a population size effect[16], that is, that errors are likely to be smaller in countries with large than with small populations, receives very little support from this graph. Only two of the 14 countries have effects that agree with this theory: France and Luxembourg. Large countries such as Italy and the UK have country effects that are similar to those for small countries such as Austria, Belgium, Denmark, and the Netherlands. The findings for three out of 14 countries (Finland, Norway, and Sweden) contradict the assumed relationship, as they have small populations and small migration errors. Germany is the most populous country of the 14, and has large errors. However, this is explained by its extraordinarily large immigration around 1990.[32]

Unfortunately, it is very difficult to determine what particular circumstances in the UK, or any of the countries, may have accounted for their better or poorer forecasting performance. Earlier I mentioned a number of factors which might be related to forecast accuracy. Of these, a possible population size effect was not supported by the data. I tested a possible effect of the volatility in the actual trends (as measured by the standard deviation and the absolute slope in the observed data in the years preceding the forecast's base year), but the result was negative here as well. Other factors, such as the number of staff involved in the production of the forecast, the quality of the data that are available, policy measures unexpectedly introduced by authorities, or sudden political events that have an impact on demographic developments, are difficult to quantify, or the relevant data are not available.

Probabilistic population forecasts: quantifying our ignorance

A recent review of demographic forecasting in industrialized countries concludes that significant progress has been made in the past in increasing our understanding of the causal factors and processes that determine demographic events (Booth 2006[33]). Yet this has not resulted in improved theory-informed forecasts, as shown above. The author states that demographic behaviour is too complex to be easily modelled and forecast – it may be inherently unpredictable. This provides motivation for handling forecast uncertainty in probabilistic terms.

Forecast errors are inherent to forecasting, and forecast users should be informed about the magnitude of those errors. The type of analyses reported here and in the previous article (Shaw 2007[1]) give the user useful information about the accuracy of old forecasts, but not for the current one. However, a **probabilistic** forecast computes the future population in the form of probability distributions, unlike traditional **deterministic** forecasts which just give single estimates. Thus the user of such a forecast is informed about the likely magnitude of the errors, and how these errors vary across age groups or between the sexes. When a decision maker is able to take forecast uncertainty into account, this may lead to better decision making. As soon as he or she knows the expected costs of a decision based on an erroneous forecast, an optimal strategy can be chosen. Unfortunately, nearly all official forecasts are deterministic, not probabilistic – Statistics Netherlands is the only known exception (Alders and De Beer 1998[34]). But demographers and statisticians have developed methods to calculate probabilistic forecasts. By way of illustration, I shall present a probabilistic forecast for the United Kingdom.

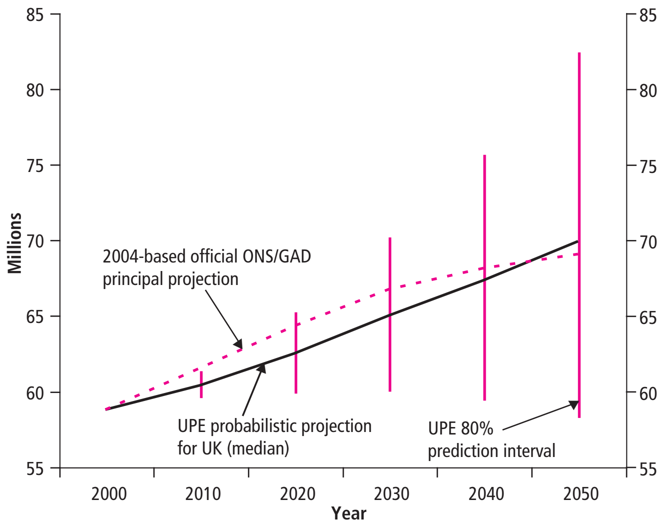
The probabilistic forecast for the United Kingdom is part of a recent large project, called 'Uncertain Population of Europe' (UPE). The aim of that project was to compute the probability distributions of future demographic variables, such as population size, age groups etc. for 18 countries in Europe (designated as the EEA+). These countries were the 15 members of the European Union prior to the accession of the new member states in 2004, plus Norway, Iceland, and Switzerland. Except for Switzerland, these countries made up the so-called European Economic Area, hence EEA+. Uncertainty was quantified by applying the traditional cohort component model for each country 3,000 times, with a specified base population, and probabilistically varying values for age-specific mortality, age-specific fertility, and net migration. The forecast horizon was 2050. Time series analysis, historical forecast errors, and expert opinion were used to formulate assumptions on central estimates (point predictions) and uncertainty parameters for fertility, mortality and international migration. These probabilistic forecasts were based on the population at 1 January 2003 and results were first published in 2004. More details can be found at www.stat.fi/tup/euupe/ and in Alho *et al* (2006)[35] and Alders *et al* (2007).[36]

The results of the 3,000 runs of the cohort component model for the period up to 2050 were assembled in a data base containing the future population for each country, and for the EEA+ as a whole, broken down by age, sex, forecast year, and forecast run. For each variable of interest, for example total population in 2030, or the old age dependency ratio in 2050, one can construct a histogram based on the 3,000 simulated values, and read off prediction intervals for any chosen probability range. I shall give some selected results for the UK.

The results show that, based on the probability assumptions made for the UPE project, there is an 80 per cent chance that the population of the UK, now 60 million, will number between 60 and 65 million in the year 2020, and between 58 and 82 million in 2050; see Figure 10. For mid-2006, the 80 per cent interval was 59.5 million to 60.2 million. However, ONS has recently published a higher estimate for mid-2006 of 60.6 million.[37] Three years ago, when the probabilistic forecast was published, such a high number was considered very unlikely, although not impossible. The probabilistic forecast expected lower net migration than has actually occurred and was also based on official UK population estimates prior to an upward revision published by ONS in September 2004. The range for 2050 illustrates that long-term uncertainty is quite large. Although continued growth to 2050 is probable, a decrease in population size cannot be excluded. The UPE results indicate that the probability of the UK population in 2050 being smaller than the current 60 million is an estimated 14 per cent.

Figure 10 shows the probability distribution for the total population size of the UK at 2050 based on the assumptions made for the UPE project. The black solid line is the 50 per cent line (or median): chances are 50 per cent that population size in 2050 will be less than 69.4 million; but a population larger than 69.4 million is equally likely. The red vertical lines represent 80 per cent intervals. There is an 80 per cent probability that population size in 2050 will be between 58.4 and 82.4 million. Thus chances are 10 per cent that it will be less than 58.4 million, but also 10 per cent that it will be more than 82.4 million. Note that uncertainty regarding predicted population size increases over time, as the 80 per cent intervals widen rapidly. The latest official UK principal projection (GAD, 2005[38]), is shown as the dotted line on the chart. This is somewhat above the UPE median forecast in the short to medium term (but well within the 80 per cent interval) although very similar for the year 2050.

Figure 10 UPE probabilistic projection* of total population size, 2000–50



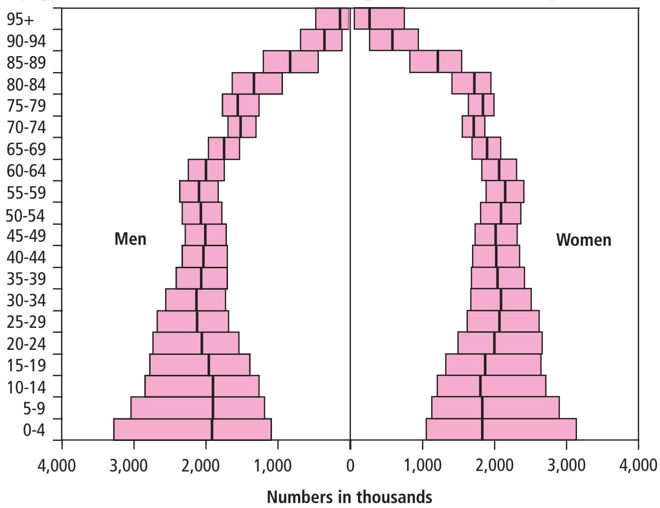
* UPE projection from 'Uncertain Population of Europe' research project. See text for details

Uncertainty is largest for the youngest and the oldest age groups, because these are largely dependant on the accuracy of fertility and mortality assumptions. So, for example, prediction intervals in 2030 for the population younger than 20 years of age are rather wide. There is rather less uncertainty about the size of the working age population as everyone who will be aged over 25 at 2030 has already been born. Of course, international migration also shows large prediction intervals around expected levels, but its impact on the age structure a few decades into the future is relatively modest. But by 2050, uncertainty has accumulated so strongly, that intervals are very large, not only for the young, but also for the old age groups, at least relative to the size of the median forecast (see Figure 11).

Figure 11 shows the age pyramid for the UK in 2050 based on the probability assumptions made for the UPE project. For example, as shown by the orange areas, the probability is 80 per cent that the number of boys aged 0–4 in the UK in 2050 will be between 1.1 million and 3.3 million. And, as indicated by the black line inside the orange areas, there is a 50 per cent chance that there will be at most 1.9 million women aged 65–69 in the UK in 2050.

Figure 11 UPE probabilistic projection* for the UK

Age pyramid at 2050, median (black line), 80% prediction interval (orange area).



* UPE projection from 'Uncertain Population of Europe' research project. See text for details

These examples illustrate how one could use a probabilistic population forecast to express forecast uncertainty. However, clearly, the results of a probabilistic forecast also depend on the assumptions that the forecaster has formulated. In this case, assumptions are needed for the form of the probability distributions of the demographic rates. Hence one can say that the prediction intervals of probabilistic forecasts are uncertain themselves. Thus the producer of a probabilistic population forecast should carefully document the choices that were made in the assumption making process – similarly to the producer of a traditional forecast. But there are also important differences between a probabilistic population forecast and one that results from the traditional deterministic approach.

First, uncertainty is not quantified in a deterministic forecast. Second, high and low variants in the traditional approach do *not* show a range of likely future developments, as users may interpret. In the high variant, migration is assumed high in *every* year of the forecast period (and *vice versa* for the low variant). In other words, one assumes perfect serial correlation. However, it is extremely unlikely that migration will follow the path as defined in the high (or low) variant for an extended period, say thirty or more years. In the official UK projections (GAD, 2005[38]) it is stressed that variants do *not* represent limits for future demographic behaviour, and that they should be used for sensitivity analyses: '…these migration variants should be regarded as giving an indication of the implications for the future, if average migration levels were to differ significantly from those assumed in the principal projection.' But one may question how useful such sensitivity analyses are for the user of the forecast, when we know that it is extremely unlikely that the high and the low variants will ever materialize. In contrast, a probabilistic forecast does not assume perfect serial correlation (in any one run the level of migration will be very volatile as opposed to a smooth high or low time path in traditional variants). A user could consider one run, or just a few runs, of the probabilistic forecast if the interest is in the possible deviations from the central path.

A third drawback of the traditional deterministic approach is that the use of high and low variants for more than one component in a particular projection is inconsistent from a statistical point of view (Lee 1999[39], Alho 1998[40]). So when two or more mortality variants are formulated, in addition to two or more fertility variants, high (or low) population growth variants are often produced by combining high fertility with high life expectancy (or low fertility with low life expectancy). In that case, *any* year in which fertility is high, life expectancy is high as well. In other words, one assumes perfect correlation between fertility and mortality, in addition to perfect serial correlation for each of the two components. Assumptions of this kind are unrealistic, and they cause inconsistencies: a variant that is extreme for one variable need not be extreme for another variable.

In his British Academy Annual Lecture on 1 December 2004[41], the Bank of England's Governor Mervyn King stressed that in a wide range of collective decisions it is vital to think in terms of probabilities. We must accept the need to analyse the uncertainty that inevitably surrounds these decisions. In order that public discussion can be framed in terms of risks, the public needs to receive accurate and objective information about the risks. Transparency and honesty about risks should be an essential part of both the decision-making process and the explanation of decisions. If population projections are to inform policy decisions, then uncertainty of these projections must be assessed. In some areas, greater uncertainty might lead to postponement of action. In other policy arenas such as education planning, greater uncertainty might indicate that the best polices would be those most easily changed as the future unfolds. For example, a school planner facing uncertain projections of enrolment growth might decide to rent additional space for schools rather than building or buying space. Explicitly estimating the degree of uncertainty in population projections encourages consideration of alternative population futures and the full range of implications suggested by these alternatives (Lee and Tuljapurkar 2007[42]).

But the public has great difficulty in understanding probabilities, and handling them. Forecasters, whether occupied with weather, or inflation, or population trends in the future, should develop appropriate techniques for communicating uncertainty to the users of their services. The type of charts as presented in Figure 10 for future population size of the UK, are commonly used by the Bank of England for their inflation reports, and the Met Office for their weather reports. Population forecasters should also consider using such charts.

Acknowledgement

Parts of this article were written during my stay at the Centre for Advanced Studies in Oslo, the hospitality of which I gratefully acknowledge.

References

1. Shaw C (2007) Fifty years of United Kingdom national population projections: How accurate have they been? *Population Trends* **128**, 8–23.
2. Preston S H (1974) An evaluation of postwar mortality projections in Australia, Canada, Japan, New Zealand, and the United States. WHO Statistical Report **27**, 719–745.
3. Calot G. and Chesnais J C (1978) 'La prevision démographique'. Dans A.-C. Decouflé *Traité élémentaire de prevision et de prospective*. PUF: Paris.
4. Inoue S. and Yu Y C (1979) 'United Nations new population projections and analysis of ex post facto errors'. Paper Annual Meeting Population Association of America, Philadelphia, April 1979.
5. Keyfitz N (1981) The limits of population forecasting. *Population and Development Review* **7**, 579–93.
6. Stoto M (1983) The accuracy of population projections. *Journal of the American Statistical Association* **78**, 13–20.
7. Pflaumer P (1988) 'Confidence intervals for population projections based on Monte Carlo methods'. *International Journal of Forecasting* **4**, 135–42.
8. Shaw C (1994) Accuracy and uncertainty of the national population projections for the United Kingdom. *Population Trends* **77**, 24–32.
9. Keilman N (1997) 'Ex-post errors in official population forecasts in industrialized countries'. *Journal of Official Statistics* **13**, 245–77.
10. Keilman N (2000) 'Précision et incertitudes des prévisions nationales de population'. Pp. 33–67 in C. Wattelar et J. Duchêne (red.): *Le défi de l'incertitude: Nouvelles approches en perspectives et prospective démographiques*. Actes de la Chaire Quetelet 1995. Louvain-la-Neuve: Academia-Bruylant/L'Harmattan.
11. Keilman N (2001) 'Data quality and accuracy of United Nations population projections, 1950–95. *Population Studies* **55(2)**, 149–64.
12. National Research Council (2000) *Beyond six billion: Forecasting the world's population*. Panel on Population Projections. Bongaarts J and Bulatao R (eds.). Committee on Population, Commission on Behavioral and Social Sciences and Education. Washington DC: National Academy Press. Internet http://books.nap.edu/books/0309069904/html/index.html.
13. Keilman N and Pham DQ (2004) Empirical errors and predicted errors in fertility, mortality and migration forecasts in the European Economic Area. Discussion Paper no. 386. Statistics Norway: Oslo. Available at www.ssb.no/publikasjoner/DP/pdf/dp386.pdf.
14. For Germany, I have fertility and mortality errors for the (former) Federal Republic of Germany (FRG) for the period 1952–2002. For migration, I have errors for the FRG in forecasts made between 1952 and 1989, and for the reunified Germany in forecasts made since 1990.
15. Council of Europe (2002) *Recent demographic developments in Europe 2002*. Council of Europe Publishing: Strasbourg.
16. Council of Europe (1998) *Recent demographic developments in Europe 1998*. Council of Europe Publishing: Strasbourg.

17. Adam A Y (1992) The ABS population projections: Overview and evaluation. *Journal of the Australian Population Association* **9**, 109–30.
18. Wilson T (2007). The forecast accuracy of Australian Bureau of Statistics national population projections. *Journal of Population Research* **24:1** 1–27
19. George M V and Nault F (1991) The accuracy of Statistics Canada's demographic projections. Paper Annual Meeting of the PAA, Washington DC, March 1987.
20. Feeney G (1990) The demography of aging in Japan: 1950–2025. NUPRI Research Paper Series no. 55. NUPRI: Tokyo.
21. Ahlburg D (1982) How accurate are the U.S. Bureau of the Census projections of total live births? *Journal of Forecasting* **1**, 365–74.
22. Long J F (1987) The accuracy of population projection methods of the U.S. Census Bureau. Paper Annual Meeting of the PAA, Chicago, April/May 1987.
23. Long J F (1995) Complexity, accuracy, and utility of official population projections. *Mathematical Population Studies* **5**, 203–16.
24. Mulder T J (2002) Accuracy of the U.S. Census Bureau national population projections and their respective components of change. Population Division Working Paper Series no. 50. U.S. Census Bureau: Washington.
25. Crimmins E M (1993) 'Demography: The past 30 years, the present, and the future'. *Demography* **30(4)** 579–91.
26. Preston S H (1993) 'The contours of demography: Estimates and projections'. *Demography* **30(4)** 593–606.
27. Keilman N (2007) European demographic forecasts have not become more accurate during the past 25 years. Submitted. Available at http://folk.uio.no/keilman/publications.html
28. All dependent variables are in logarithmic form.
29. Keyfitz N (1982) Can knowledge improve forecasts? *Population and Development Review* **8**, 729–51.
30. Two factors explain this finding. First, small populations are often more strongly influenced by unforeseeable migration flows than large populations. Second, forecasters often give less attention to small populations than to large populations in a multi-population forecast. So far, a size effect has not been tested for variables other than total population. Because of the second explanation, one could expect to find it for errors in migration forecasts. See comments in connection with Figure 9.
31. Smith S K, Tayman J, and Swanson D (2001) *State and Local Population Projections: Methodology and Analysis*. Kluwer Academic/Plenum Publishers: New York. (The Plenum series on demographic methods and population analysis).
32. In addition to main effects, the model for migration errors includes a few interaction effects, in order to take into account that the effects for some countries were very different for certain periods. (The models for fertility and mortality errors include main effects only.) The migration model includes interaction effects for Portugal with the period 1970–74 (high emigration of labourers to other European countries), and for the FRG, Germany, and Austria with the period 1990–94 (large immigration flows from Central and Eastern Europe after the fall of the Berlin Wall). The estimates for these interactions are 0.69, 0.74, 0.61, and 0.26, respectively. They have to be added to the estimates in Figure 9, which shows main effects only. As a consequence, the country effects for the FRG, Germany, and Portugal become much larger, while Austria moves to the group of countries for which migration was significantly more difficult to predict than for the UK.
33. Booth H (2006) Demographic forecasting: 1980 to 2005 in review. *International Journal of Forecasting* **22**, 547–81.
34. Alders, Maarten and Joop de Beer (1998) Kansverdeling van de bevolkingsprognose ('Probability distribution of population forecasts'), *Maandstatistiek van deBevolking* **46**, 8–11.
35. Alho J, Alders M, Cruijsen H, Keilman N, Nikander T, Pham D Q (2006) New forecast: Population decline postponed in Europe. *Statistical Journal of the United Nations ECE* **23**, 1–10.
36. Alders M, Keilman N and Cruijsen H (2007) Assumptions for long-term stochastic population forecasts in 18 European countries. *European Journal of Population* **23**, 33–69.
37. UK population grows to 60,587,000 in mid-2006. ONS News Release (22 August 2007).
38. Government Actuary's Department (2005) UK 2004-based national population projections are available at: www.gad.gov.uk/Population/index.asp
39. Lee R (1999) 'Probabilistic approaches to population forecasting', in Lutz W, Vaupel J, and Ahlburg D (eds.). *Frontiers of Population Forecasting*. Supplement to Vol. 24 of *Population and Development Review*, 156–90.
40. Alho J M (1998) A stochastic forecast of the population of Finland. Reviews 1998/4. Statistics Finland: Helsinki.
41. King M (2004) What fates impose: facing up to uncertainty. British Academy lecture www.britac.ac.uk/news/release.asp?NewsID=154
42. Lee M. and Tuljapurkar S (2007) The Degree of Certainty in Population Projections. Population Reference Bureau. Available at www.prb.org/Articles/2007/PopulationProjections.aspx