How Accurate Are the United Nations World Population Projections?
Author(s): Nico Keilman
Source: *Population and Development Review*, Vol. 24, Supplement: Frontiers of Population Forecasting (1998), pp. 15-41
Published by: Population Council
Stable URL: http://www.jstor.org/stable/2808049
Accessed: 07/05/2011 09:52

How Accurate Are the United Nations World Population Projections?

Nico Keilman

Producers of population forecasts rarely report about the expected accuracy of their forecast results. The fact is surprising, because accuracy is an important aspect of the quality of the forecast. Other aspects, such as the information content (e.g., only total population, or also age groups? which regional level?) and the usefulness for policy purposes (e.g., does the projected trend imply immediate policy measures?), are relevant as well, but the degree to which the forecast may be expected to reflect real developments in the future is a key factor in assessing its quality.

Most forecast reports do not give the expected accuracy of the population forecast, simply because the forecast producer does not know this himself. But for a number of cases, in particular, forecasts for industrialized countries made since the 1960s, we know a great deal from independent assessments (see Keilman 1997 for an overview). Ex post comparisons between projected and observed trends in population variables have revealed that the forecast accuracy of fertility is better than that of mortality—behaviorally determined variables are difficult to forecast. Large errors in forecasts of age structures have been found for both the young and the old after a forecast period of 15 years: results that were up to 30 percent too high for the age group 0–4, and 15 or more percent too low for women aged 85+ are not uncommon. This suggests that those earlier forecasts supplied useful information perhaps up to 10–15 years ahead, but certainly not longer. Finally, detailed studies for a few countries have found only a weak association between improvements in forecast accuracy and the introduction of more sophisticated forecast methods.

These findings relate to the accuracy of forecasts produced for industrialized countries. Much less is known about the reliability of population forecasts for developing countries. Inoue and Yu (1979) investigated the errors in total population size of six rounds of United Nations projections, with base years from 1950 to 1970 and observed data for the period 1950–

75. They found a consistent overestimation of the projected growth rate in developing countries after 1960, which was largely explained by the rapid slowdown of population growth in China. They also concluded that errors in the base population and in the growth rate of population immediately preceding the starting year were important determinants of errors in the projected population size of developing countries. Keyfitz (1981) and Stoto (1983) analyzed, for various countries of the world, errors in projected population growth rates in projections made by the United Nations during the 1950s and 1960s. Important findings were that errors varied strongly by region and by base year: regions in which population growth was high had large errors, as did forecasts made in the early 1950s. Moreover, Keyfitz concluded that the error in the growth rate was more or less independent of forecast duration. These conclusions were confirmed by Pflaumer (1988), who analyzed the predicted growth rates in 101 countries with at least a million inhabitants (excluding China). Forecasts were those made by the United Nations between 1963 and 1978, and actual growth rates applied to the period 1960–80. Pflaumer also found evidence for an improvement over time in the accuracy of the projected growth rates. Furthermore, errors were relatively small in countries with large population sizes.

The purpose of this chapter is to extend the analyses mentioned above, which were focused on growth rates and total population sizes. I investigate the accuracy of UN forecasts of the age structure and crude birth and death rates in seven major regions of the world: Africa, Asia, Europe, the former Soviet Union, Latin America, Northern America, and Oceania. I also include findings for a few large countries that may dominate their region: China and India (Asia), and the United States (Northern America). UN forecasts made between 1951 and 1988 are evaluated. Projected numbers for total population size, crude birth rate, crude death rate, and age structure in five-year age groups for the period 1950–90 are compared with corresponding ex post observed numbers. Two broad questions are addressed in this chapter. First, does accuracy differ strongly among regions? In other words, are population trends in some regions easier to project than those in other regions? And second, did the UN forecasts improve over time? The results indicate that the latter was indeed the case, not only because data quality improved, but also because unforeseen declines in birth rates became less important for projection errors. Concerning the first question I find that population trends are much more difficult to project in some regions than in others. Age structure projections for the former Soviet Union, Africa, and Asia show larger errors than on average. For Asia this is explained by errors in base populations. Errors caused by wrong assumptions regarding fertility and mortality explain why the projected age structure of the former Soviet Union is inaccurate, and also that of Oceania, Northern America, and Europe.

Data and methods

Measuring accuracy

Many errors analyzed in this chapter depend on the size of population subgroups. To facilitate comparison between regions and over time, most of the measures presented here are relative errors, such as the percentage error (PE) or the absolute percentage error (APE). PE is defined as

PE = (forecast – observation) x 100 /observation.

Thus, a positive PE indicates that the forecast was too high, and a negative value reflects an underestimation. When we know the errors in a series of forecasts, we can compute mean errors. The mean percentage error (MPE) provides an average measure of bias: a positive MPE indicates that forecasts tended to be too high on average, and a negative MPE reflects forecasts that were too low. When forecast results are not size dependent (e.g., crude birth rate) the term (mean) error is used, and this error is defined as the forecast minus the observed value. Absolute errors ignore the direction of the error. They indicate forecast accuracy by telling how much the forecast went wrong, irrespective of whether it was too high or too low.

UN projections between 1950 and 1985

The 12 UN forecasts that were analyzed are listed in Table 1. They are labeled in this chapter by base year. The last base year included is 1985. For some base years, the forecast was revised a few years later: 1950I, II, and III; 1975I and II; 1980I and II; and 1985I and II.

UN projections are based on the cohort-component approach for all countries of the world, except for those with a population size of under 150,000. A base population by sex and five-year age group is exposed to an assumed set of mortality and fertility rates, and to net immigration numbers by age and sex. This leads to numbers of deaths, births, and net migrants for the first five projection years, and these numbers are used to update each sex-age group in the base population to find the next age group five years later. Repeated application results in projections of population size and age groups every fifth year and projections for demographic indicators for continuous five-year periods. Four variant assumptions are formulated for fertility in each country (high, medium, low, and constant). Mortality has only one variant, and migration usually also one variant.[1]

Compared to the situation in the 1950s, the projections have expanded in regional and age detail, in time coverage, and in methodological sophistication (El-Badry and Kono 1986; Frejka 1994: 7). For instance, Africa was absent from the 1950I series, because of the unreliability or even lack of data. Country detail was available only for Latin America and the Far East. The 1950II series attempted to derive country projections from the

TABLE 1 United Nations population forecasts analyzed in this chapter

Number	Label	Source (year of publication)	Base year	Remark
1.	1950I	*Population Bulletin of the United Nations* 1 (1951) Sales no. E52.XIII.2	1950	Only total population by region for 1980
2.	1950I	*Future Population Estimates by Sex and Age. Report I: The Population of Central-America (including Mexico), 1950–1980* (1954) Sales no. 1954.XIII.3. *Report II: The Population of South-America 1950–1980* (1955) Sales no. 1955.XIII.4. *Report III: The Population of South-East Asia 1950–1980* (1959) Sales no. 1959.XIII.2. *Report IV: The Population of Asia and the Far East 1950–1980* (1959) Sales no. 1959.XIII.3	1950	Same as number 1, but the reports give five-year age groups for the years 1955, 1960, ..., 1980
3.	1950II	*Proceedings of the World Population Conference 1954* Vol. III (1955) Sales no. E.55.XIII.8, pp. 265–328	1950	Update of 1950I; only total population by region for 1955, 1960, ..., 1980
4.	1950III	*The Future Growth of the World Population* (1958) Population Studies 28, Sales no. E.58.XIII.2	1950	Update of 1950II; 5-year age groups for 1960; broad age groups for 1960 and 1975; total population for 1960, 1965, ..., 1990
5.	1960	*World Population Prospects As Assessed in 1963* (1966) Sales no. E.72.XIII.2	1960	Broad age groups for 1965, 1970, ..., 1980
6.	1965	*World Population Prospects As Assessed in 1968* (1973), Population Studies 53, Sales no. E.72.XIII.4	1965	Five-year age groups up to 70+ for 1970, 1975, ..., 1990
7.	1970	*World Population Prospects As Assessed in 1973* (1977) Sales no. E.76.XIII.4 and corrigenda	1970	Broad age groups for 1985 and 2000 only; total population for 1975, 1980, ..., 1990
8.	1975I	*Selected Demographic Indicators by Country 1950–2000: Demographic Estimates and Projections as Assessed in 1978* (1980) ST/ESA/SER/.R/38	1975	Five-year age groups for 1975, 1980, and 1990; total population and broad age groups for 1975, 1980, ..., 1990
9.	1975II	*Demographic Indicators by Country: Estimates and Projections As Assessed in 1980* (1982) Sales no. E.82.XIII.5 and corrigendum	1975	Five-year age groups for 1975, 1980, ..., 1990
10.	1980I	*World Population Prospects: Estimates and Projections As Assessed in 1982* (1985) Sales no. E.83.XIII.5	1980	Five-year age groups for 1980, 1985, 1990
11.	1980II	*Global Estimates and Projections of Population by Sex and Age: The 1984 Assessment* (1987) ST/ESA/SER/.R/70	1980	Update of number 10
12.	1985I	*Global Estimates and Projections of Population by Sex and Age: The 1988 Revision* (1989) ST/ESA/SER/.R/93	1985	Five-year age groups for 1985 and 1990
13.	1985II	*The Sex and Age Distribution of Population: The 1990 Revision of the United Nations Global Population Estimates and Projections* (1991) Sales no. E.90.XIII.33	1985	Update of number 12

projected totals in each of the 25 regions. A more innovative series was produced in the 1950III projections, based on the theory of demographic transition and on stable population theory.

The 1960 and 1965 series used stable and quasi-stable population theory and indirect estimation methods to estimate basic indicators from incomplete data. Available computer facilities made it possible to prepare the 1965 projections by age and sex for each country, and to compute a large number of other indicators. Various sets of model schedules for fertility, mortality, and migration were applied. Backward projections, starting from the base year 1965 and going back to 1950, were also prepared for each country.

These developments continued into the 1970s. More detailed indicators were computed; the complex link between socioeconomic, political, and cultural factors in fertility and mortality change was taken into account; and baseline data were improved. Finally, in the 1980s, the cycle of revisions was shortened from every five years to every two years. In the 1990s better software for projections permitted staff in the Population Division of the United Nations to try many fertility, mortality, and migration assumptions. The findings in the following sections illustrate to what extent these gradual improvements in projection approach have resulted in increased accuracy.

For most forecasts, the United Nations has computed more than one variant, typically a high, a medium, a low, and a constant variant. In such cases I limited the analysis to the medium variant, as this is the one users most often select as the best guess. Since the focus in this chapter is on comparative accuracy across regions and over time, the choice of the medium variant is not likely to have had a major impact upon the conclusions.

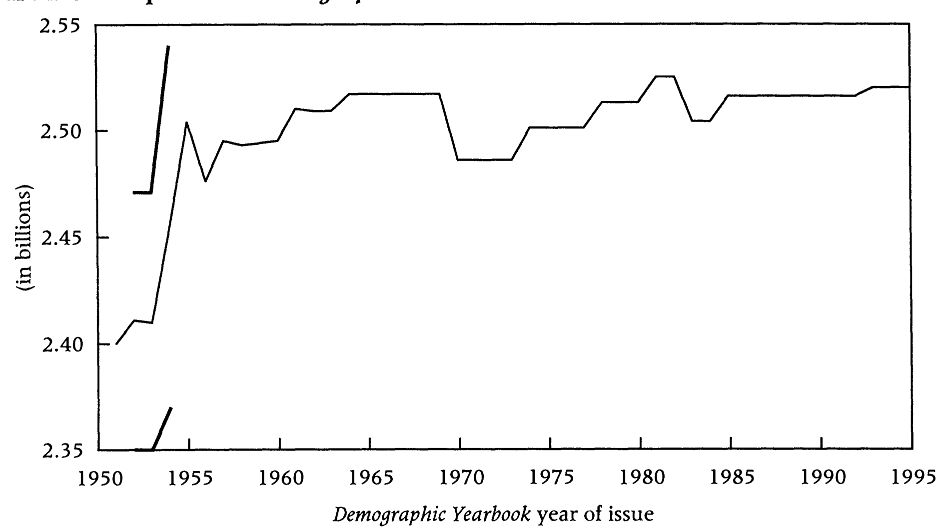
Observed population numbers

Observed population numbers have been taken from UN (1994).[2] The United Nations prefers to speak of "estimates" instead of "observations." This is understandable, because the UN continuously revises previously published demographic numbers for many countries, as new data become available and better techniques are developed. However, to avoid confusion with the notions of "projection" and "forecast," the word "observation" will be used here instead of "estimate."[3]

The frequent revision of observed data poses a problem. It has direct implications not only for the accuracy of the base population, but indirectly also for population growth in the period prior to the base year, and probably in the first few projection years as well (Inoue and Yu 1979). El-Badry and Kono (1986: 37) note that the 1950III projections included an upward revision of the world population compared to the previous round, mainly because the base population of China had to be increased by 100 million after the results of the 1953 census became available. This led in turn to substantial reductions in assumed mortality levels. Such surprises have also occurred in recent years, and not only in developing countries. Poursin (1994: 20) reminds us of the November 1991 census in Nigeria that resulted in a population size of 88 million, 35 million lower than the 123 million that had been expected. In the 1990 round of censuses in Europe, Italy counted almost 1 million fewer persons than expected on the basis of vital statistics, and Portugal missed half a million persons (Cruijsen and Eding 1995: 10).

Against which yardstick should one evaluate projected numbers? For instance, the world's population size as of 1950 has been revised on many occasions. The 1951 to 1994 issues of the UN *Demographic Yearbook* contain mid-year estimates for the world's total population of 1950, as shown in Figure 1. Frequent revisions, carried out in the light of new data and better methods, have led to increasingly higher estimates of the 1950 population size. The upward adjustment after the Chinese census in 1953 is clearly visible. The *Yearbooks* of 1952, 1953, and 1954 give an interval for the world population, ranging from a low of 2.35–2.37 billion to a high of 2.47–2.54 billion. Both the intervals and the arithmetic averages for these years are included in Figure 1. Since the 1960s there seems to be agreement on a number of roughly 2.5 billion, but even in 1992 small adjustments were carried out. Then the question arises with which number should we compare the base-year population of the 1950I forecast (and which sources should we use for evaluating other forecast results)? If we make use of the 1951 figure for the population size in 1950, we do justice to the circum-

FIGURE 1 Mid-year estimates of world population size for 1950 from subsequent UN *Demographic Yearbooks*



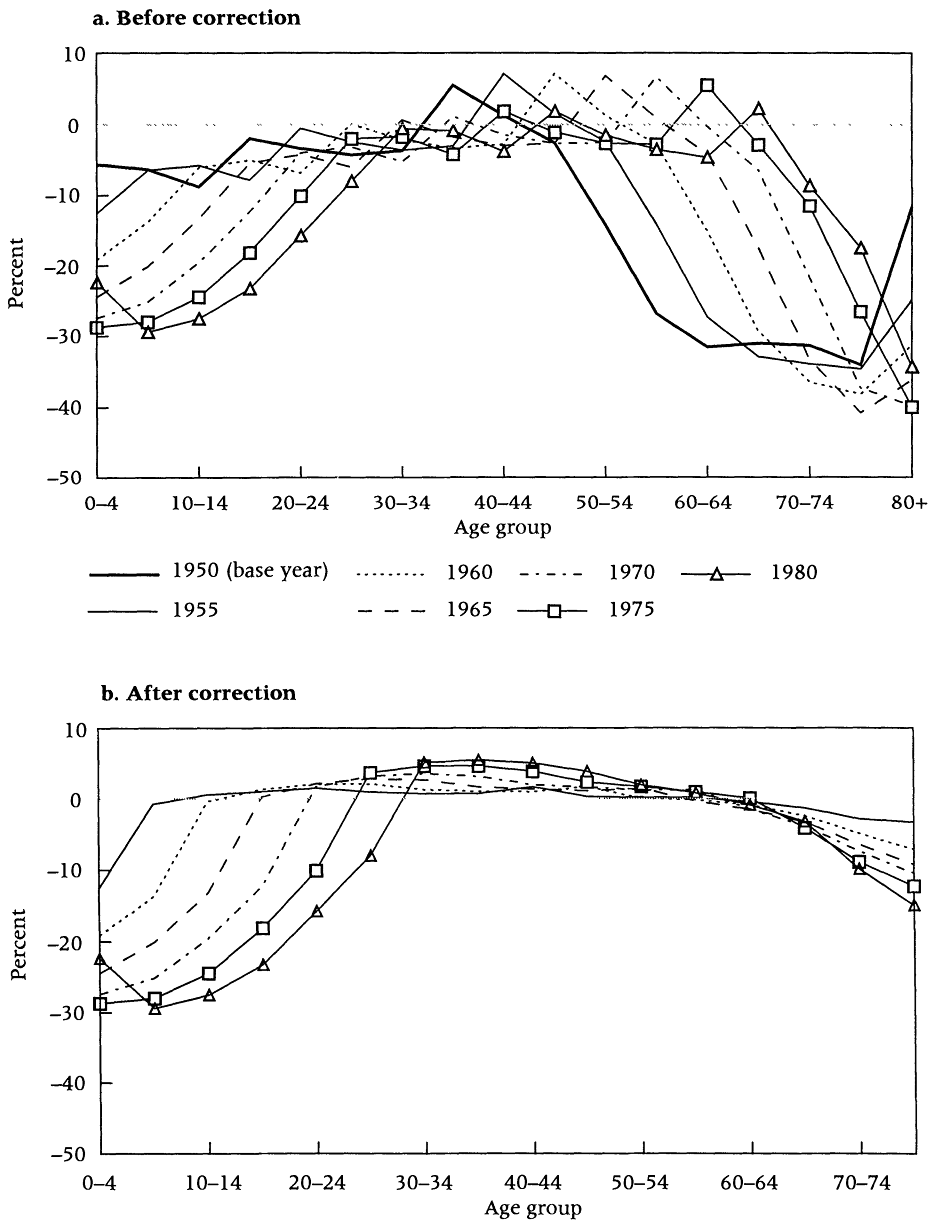
stances under which the forecasters had to work. Although better data about the situation in 1950 became available in later years, the forecasters could not have known the revised figures. But there are two objections to selecting the early number. First, the earliest data for 1950 were probably outdated when observed population trends for later years were assessed. This may cause inconsistencies in the time series. Second, from the point of view of the user, the most recent data for the year 1950 have to be preferred, because these may be assumed to be closest to the real (but for many countries unknown) numbers. Therefore I chose to base the observations for the period 1950–90 on the data published in United Nations (1994).

A correction method

The disadvantage is, of course, that this creates an error in the base population. In some cases the revision has been so large that the base-year error dominates errors in subsequent years. Moreover, the base-year errors are strongly age dependent, and the errors propagate through the age structure in later years. As a consequence, an observed forecast error in the age structure after, say, five years into the forecast period is the result of two error sources: 1) the initial error in the base-year population, caused by bad data quality, and 2) an error caused by wrong assumptions on fertility, mortality, and migration during the five-year period since the base year. I simply assumed that the observed error is the sum of these two. In this way I could isolate the error due to wrong assumptions, by subtracting the error in the base population from the observed forecast error.

As an example, consider Figure 2a, which shows percentage errors by age for the 1950I projections for Central America, as compared with observations for every fifth year between 1950 and 1980. An error of +5 percent in the age group 35–39 in 1950 comes back every five years as an error of the same magnitude in the next five-year age group. But much more striking is the pattern of rapidly falling errors for age groups between 35–39 and 75–79 in 1950 (from +5 percent down to –35 percent), which repeats itself in later years for higher age groups. Thus Figure 2a mixes errors due to wrong assumptions (too high death rates and too low birth rates for the five-year periods between 1950 and 1980) with errors that were already present in the base population of 1950, caused by revisions of "observed" data. Errors in the age structure caused by wrong assumptions alone can be isolated from the overall error by subtracting the error in the base population by age from the errors for later years, thereby following five-year birth cohorts. For example, the percentage error in age group 35–39 in 1950 is subtracted from that of age group 40–44 in 1955, and from that of age group 45–49 in 1960, and so on.[4] This correction was applied to the error in each five-year age group. Corrected percentage errors (CPE) by age for the 1950I projections for Central America are displayed in Figure 2b. The underestimation of the elderly is much smaller here than in the case of uncorrected errors in Figure 2a. The correction removed the effect of errors in the base population caused by bad data quality: corrected errors for the elderly are caused by wrong death rates—

FIGURE 2 Percentage errors in projected age structures, 1950I projections, Central America



in this case, too pessimistic assumptions for the age group 60+.[5] The positive errors for adults between ages 20 and 50 are a combination of two factors: first, and probably more important, too low outmigration, and second, too low mortality. Corrected errors in the base year are all zero by definition, and therefore omitted from Figure 2b. Cohorts born after 1950 were not included in the base population. For these cohorts, no correction is necessary and thus the uncorrected values are given in Figure 2b. Five-year cohorts cannot be followed into the highest open-ended age group (80+) and therefore corrected errors have been computed up to age 75–79.

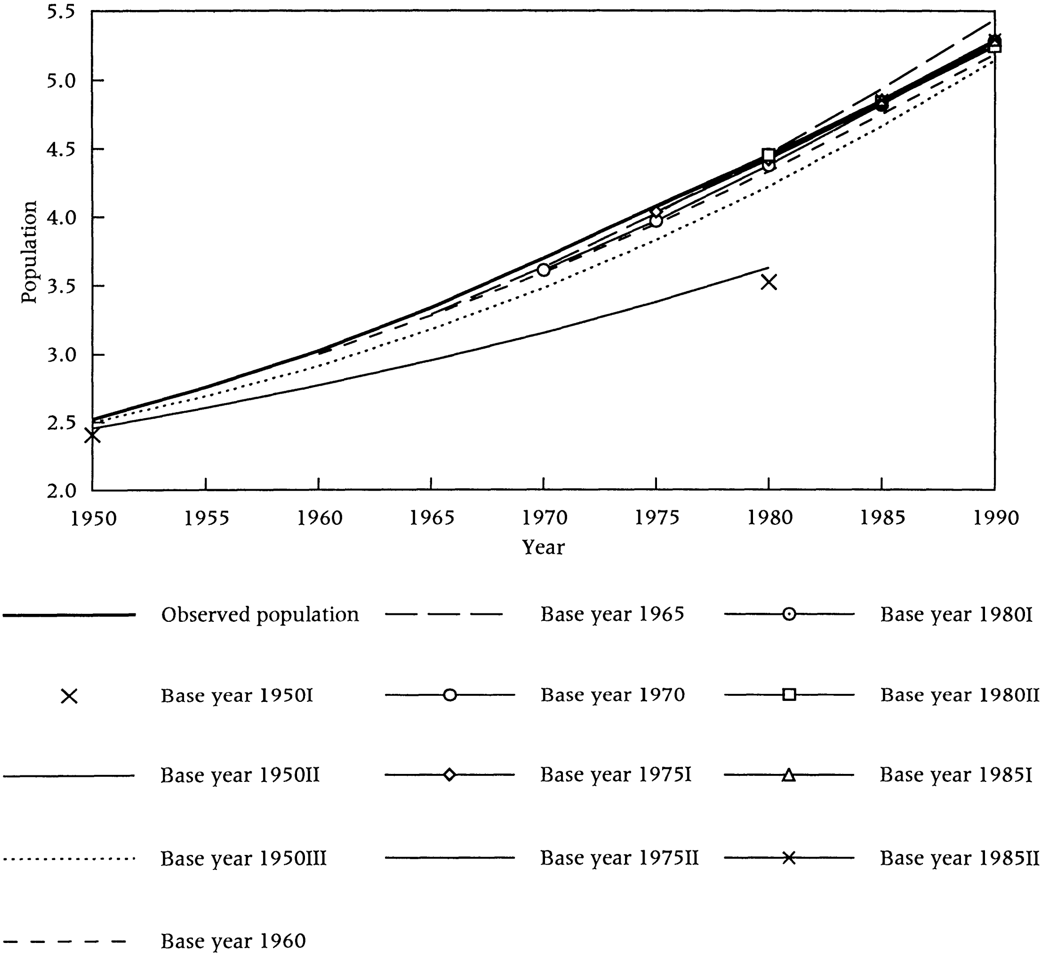
In the following sections, both corrected and uncorrected percentage errors in the age structure will be analyzed. Mean errors in the age structure for the world as a whole and for the seven regions have been calculated. Forecasts with base years 1950 and 1960 had very little age detail and are therefore omitted. For each age group, the mean error was taken over the subsequent forecasts, controlling for forecast duration. Thus mean errors are presented for the base population and for the age structure at forecast durations of five years and ten years ahead. Results for a duration of 15 years have also been computed, but these are not analyzed here, as they are based on two or three observations only.

The world

Figure 3 shows that the first two forecasts for the world as a whole with base year 1950 (i.e., 1950I and 1950II) both had too low population growth and a base population that was too small. The above-mentioned underestimation of China's population by 100 million is part of the explanation. The 1950III forecast and those made later were much more accurate. Noteworthy is the 1965 forecast, for which the growth rate was too low until 1975, and too high for the years thereafter. The result was that the world's projected population size crossed the observed size in 1980 and was higher than the observations in the decade thereafter. This is mainly caused by the 1965 projection for India, in which annual growth rates were too high by 0.35–0.45 percentage points in the period 1965–80. But unforeseen drops in birth rates in Europe and Northern and Latin America have also contributed to this trend, as seen below.

Concerning forecasts of the world's age structure since 1965, the mean percentage error is very modest for most age groups (see Figure 4). Except for the population aged 80+, the mean errors are limited to between –2 and +2 percent. The underestimation of the age group 40–70 lasted until the 1980II forecast. It can be explained by the failure to appreciate the pace of mortality declines that already characterized the global population forecasts in the 1940s and 1950s made by the United Nations and others (Coale 1983; Lee 1991). Beginning with the 1980II forecast the errors for

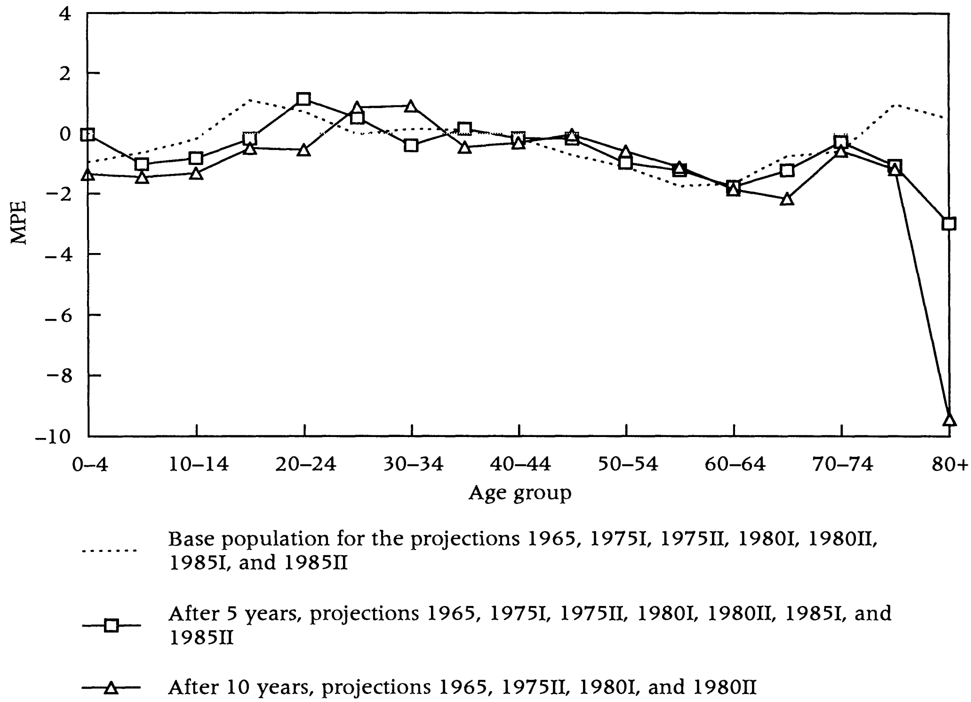
FIGURE 3 Observed and projected world population, 1950–90 (billions)



ages over 40 are much smaller. After correction for bias in the base population resulting from bad data quality, the age structure errors caused by wrong fertility and mortality assumptions become less than one percent in absolute value. Note that age groups 0–4, 5–9, and 80+ in Figure 4 have larger errors at a duration of ten years ahead than five years ahead. In other words, fertility and mortality assumptions become more uncertain when forecast duration increases. The effect is even stronger for developed countries as seen below.[6]

Errors for the elderly are caused by wrong mortality assumptions; those for children are caused by wrong assumptions for fertility and infant mortality. Therefore I analyzed errors in assumed crude birth rates (CBR) and assumed crude death rates (CDR). Data on the CBR and CDR forecasts have been published, or could be constructed, for ten forecasts and all seven

FIGURE 4 Mean percentage errors (MPE) in projections of age structures, world total



major regions. Other indicators, such as the total fertility rate (TFR) and the life expectancy at birth, are to be preferred for the analyses (at least in principle), because they are independent of the age structure. Hence errors in the forecasts of the TFR or of life expectancy are not influenced by errors in the forecasts of age structures. But the necessary data have only been found for eight forecasts (1965, 1970, 1975I, 1975II, 1980I, 1980II, 1985I, 1985II). Errors in the life expectancy at birth will be presented in the concluding section. Those in the TFR will not be discussed here, as preliminary analyses showed that the overall behavior in these errors was not very different from that in the CBR. Apparently, the influence of errors in the age structures on those in the TFR has been small.

Table 2 presents the errors in crude death rates in ten forecasts for the world as a whole. The 1950I forecast is not included. It built on vital rates continuing at the 1946–48 level, which led to CDR forecasts of between 2.2 and 2.5 percent for the whole forecast period. The 1950III forecast was the first one, as far as could be traced, that had an explicit extrapolation for mortality. The rates apply to the eight five-year periods from 1950–55 to 1985–90. There are three (dependent) dimensions in the errors for a given region: the calendar year or period for which the errors have been computed, the base year, and the forecast duration (calendar year of observation minus base year). Therefore three types of averages are presented:

TABLE 2 Errors in projected crude death rates (percentage points) by base year, forecast period, and forecast duration, world total

Forecast period (1 January)	Forecast base year 1950III	1960	1965	1970	1975I	1975II	1980I	1980II	1985I	1985II	Mean error	Duration in years	Mean error
1950–55	.52										(.52)	0–5	.10
1955–60	.77										(.77)	5–10	.16
1960–65	.54	.03									(.29)	10–15	.17
1965–70	.76	.10	.06								(.31)	15–20	.26
1970–75	.93	.19	.11	.11							.34	20–25	(.36)
1975–80	.60	.18	.06	.09	.05	.04					.17	25–30	(.37)
1980–85	.67	.12	.02	.07	.04	.03	.03	.02			.13	30–35	(.67)
1985–90	.77	.14	.03	.09	.08	.08	.07	.06	.06	.05	.14	35–40	(.77)
Mean error	.70	.13	.06	.09	(.06)	(.05)	(.05)	(.04)	(.06)	(.05)	.22		.22

NOTE: Numbers in parentheses are averages based on fewer than four observations.

TABLE 3 Errors in projected crude birth rates (percentage points) by base year, forecast period, and forecast duration, world total

Forecast period (1 January)	Forecast base year 1950III	1960	1965	1970	1975I	1975II	1980I	1980II	1985I	1985II	Mean error	Mean absolute error	Duration in years	Mean error	Mean absolute error
1950–55	.16										(.16)	(.16)	0–5	.02	.06
1955–60	.34										(.34)	(.34)	5–10	.09	.15
1960–65	.17	–.14									(.02)	(.16)	10–15	.18	.18
1965–70	.32	–.09	.00								(.08)	(.14)	15–20	.28	.28
1970–75	.61	.15	.23	.06							.26	.26	20–25	(.35)	(.35)
1975–80	.87	.33	.38	.28	.11	.02					.33	.33	25–30	(.52)	(.52)
1980–85	.94	.23	.33	.25	.10	–.01	–.03	–.05			.22	.24	30–35	(.94)	(.94)
1985–90	1.01	.16	.22	.15	.07	.07	–.08	–.09	.02	.02	.16	.19	35–40	(1.01)	(1.01)
Mean error	.55	.11	.23	.19	(.09)	(.03)	(–.05)	(–.07)	(.02)	(.02)	.20			.20	
Mean absolute error	.55	.18	.23	.19	(.09)	(.03)	(.05)	(.07)	(.02)	(.02)		.23			.23

NOTE: Numbers in parentheses are averages based on fewer than four observations.

by period (average taken over base years), by base year (over periods), and by duration (over base years).[7] The last average has been computed alongside the main diagonal (for duration 0–5 years) and subdiagonals (for longer durations) of the table. For instance, the mean error of 0.26 percentage points at a duration of 15–20 years was computed as the average of 0.76 (forecast 1950III, period 1965–70), 0.18 (forecast 1960, period 1975–80), 0.02 (forecast 1965, period 1980–85), and 0.09 (forecast 1970, period 1985–90).

First we note that all errors in Table 2 are positive.[8] Hence real mortality levels were always lower than projected ones. In the 1950III round of forecasts, the crude death rate was assumed to fall slowly, from 2.5 percent throughout the 1950s, to 2.1 percent in the 1960s and the beginning of the 1970s, reaching an ultimate level of 1.7 percent from 1975 onward. The data we now have, 40 years after the forecast was made, indicate that these levels were too pessimistic by 0.7 percentage points on average. The situation became considerably more favorable beginning with the 1965 forecast: since then, mean errors have been around 0.06 percentage points only. Ignoring averages computed on the basis of fewer than four observations, there is a clear tendency of improvement over time: during the period 1970–90, mean errors were roughly halved. During the first 20 years of forecast duration, the mean error increased more than twofold, from 0.10 to 0.26 percentage points.

Errors in crude birth rates are shown in Table 3. Here we note that the forecasts with base years 1980I and 1980II had too low birth rates, as indicated by the minus signs. Hence the mean error in those two forecasts was –0.05 and –0.07 percentage points for the whole of the 1980s. The 1950III forecast overestimated the falling crude birth rates for the world by 0.55 percentage points. The forecasts of the 1960s and 1970s had too high birth rates as well, but much less so. The period averages indicate that it was relatively difficult to give accurate forecasts of the CBR during the years 1970–85, as absolute errors are between 0.24 and 0.33 percentage points on average. Increasing uncertainty in the CBR as forecast duration grows is reflected by the mean absolute error in the last column, which grows from 0.06 percentage points in the first five years of the average forecast to 0.28 percentage points for periods of 15–20 years ahead.[9]

In principle, one would expect it to be easier to extrapolate mortality than fertility—after all, everyone dies exactly once, and the only uncertainty connected to mortality is around the timing of death, whereas for fertility we have to guess not only the timing (mean age at childbearing), but also the number of children a woman has. Yet, the overall mean absolute error (MAE) in the CBR is 0.23 percentage points, which is almost equal to that for the CDR (0.22 percentage points, see Table 2). This suggests that it has been equally difficult to give accurate extrapolations for mortality and fertility for the period and forecasts studied in this chapter.

TABLE 4 Mean errors and mean absolute errors in projected crude death rates and crude birth rates (percentage points) by major region, projections with base years as in Tables 2 and 3, period 1950–90

	Crude death rates		Crude birth rates	
Region	Mean error	Mean absolute error	Mean error	Mean absolute error
Africa	.23	.28	–.01	.09
Asia	.34	.34	.26	.32
Europe	–.02	.08	.15	.17
Soviet Union	–.15	.15	.09	.12
Northern America	.03	.06	.25	.30
Latin America	.08	.14	.33	.36
Oceania	.02	.06	.19	.24
World total	.22	.22	.20	.23

The poor quality of vital data, in particular for Africa and Asia in the 1950s, explains this counterintuitive finding. Table 4 shows that the overall MAE for mortality in Africa and Asia was higher than the corresponding overall MAE for the world as a whole. For the other regions the error was much lower, ranging from 0.06 percentage points for Northern America and Oceania to 0.15 percentage points for the former Soviet Union. A comparison with the last column in Table 4 demonstrates that fertility was indeed more difficult to extrapolate than mortality in Europe, Northern America, and Oceania, that is, in regions with relatively good data.

Tables and figures similar to Tables 2 and 3 and Figures 2 and 4 have been produced for all seven major regions, and for China, India, and the United States. A few of them will be presented below, and the other major findings will be summarized. Detailed results are available from the author upon request.

Africa

Young adults (20–40) and older age groups (55–80) have been underestimated in African forecasts since 1965. Figure 5 shows the mean percentage error (MPE) for forecasts of age structures five and ten years ahead, together with the MPE for the base populations. The pattern is mildly sloping downward with advancing age (except for age group 80+). When we correct for errors that are already present in the base population, the resulting error pattern becomes almost flat at a level close to zero (figures not shown here). Hence the pattern in the MPE in Figure 5 has been caused mainly by errors in the base populations, and much less so by wrong fertility and mortality assumptions in the forecasts.

FIGURE 5 Mean percentage errors (MPE) in projections of age structures, Africa

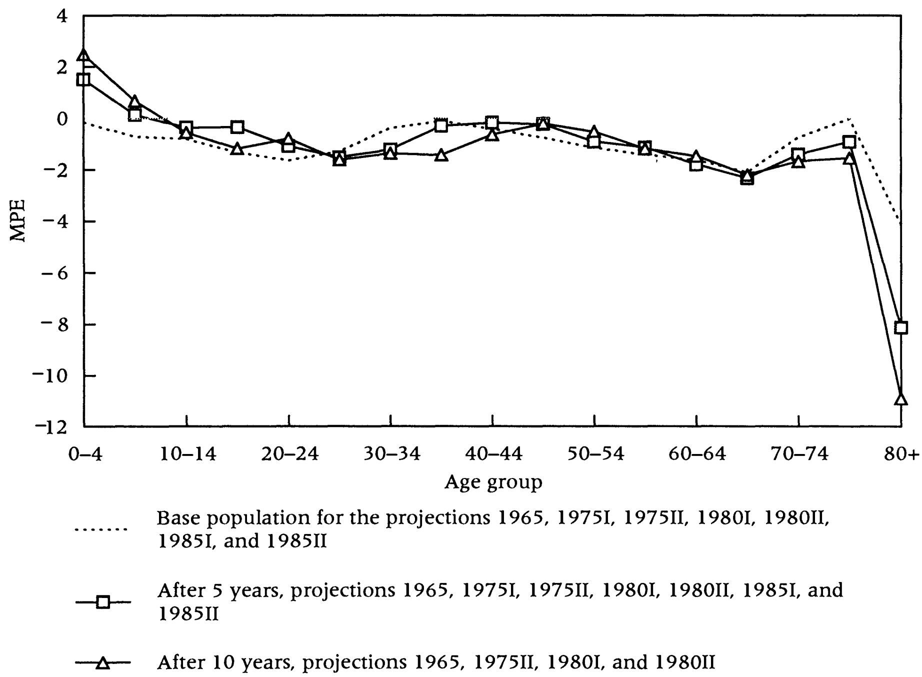
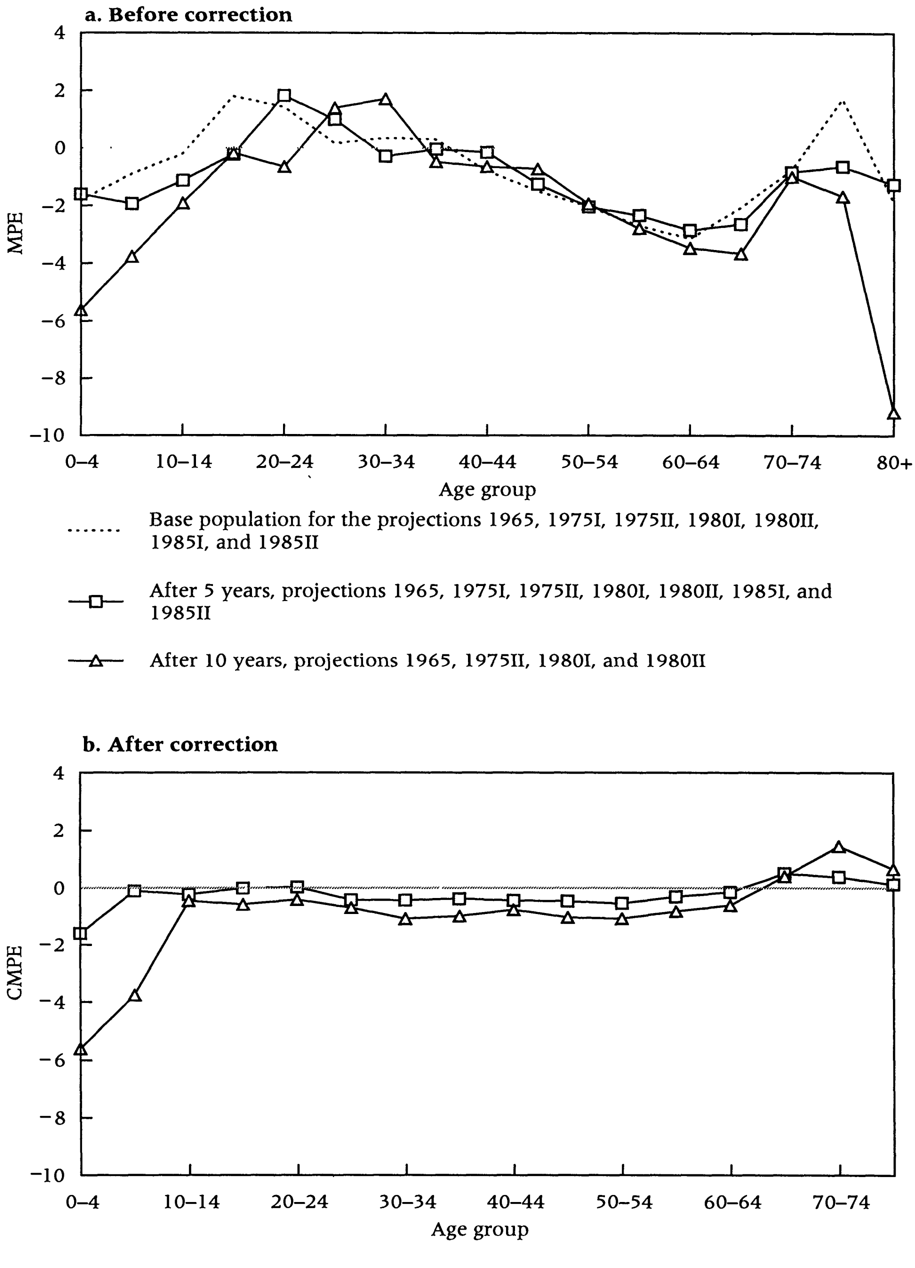


Table 4 shows that Africa's mean error in crude death rates is much higher than that in crude birth rates. This is entirely due to the 1950III forecast, for which the death rate shows an excessively large error. On the other hand, fertility assumptions in the 1950III forecast were much more accurate. Apparently, the paucity of data had a larger impact on mortality than on fertility assumptions.

Asia

Asian forecasts since 1965 have a pronounced error pattern for the age groups over 40, as seen in Figure 6a. After correction for errors in the base population (Figure 6b), the errors for this age group are more than halved, indicating that wrong mortality assumptions have had less impact than wrong base population data. Numbers of children below age 15 have been underestimated. One explanation is too low fertility rates in the five forecasts beginning with 1975II, for which the mean error in crude birth rates has been −0.08 percentage points.[10] The birth rates in the five forecasts from 1950III to 1975I were far too high (so that the overall error in crude birth rates became +0.26 percentage points; see Table 4), but two of the forecasts (base years 1950III and 1960) lacked age detail and could therefore not be included in the age structure errors. Much of the error for children

FIGURE 6 Mean percentage errors (MPE and CMPE) in projections of age structures, Asia



under 15 is caused by the forecasts for India and China. Chinese actual birth rates were higher (by 0.19 percentage points on average) than those foreseen in the six forecasts beginning with 1975; the same is true for the two forecasts for India with base year 1980 (also 0.19 points on average).

Forecasts for the age structure of China are characterized by large baseline errors, underestimations of the age groups 0–10 (between –6 and –15 percent after a forecast duration of ten years), small errors up to age 70 (plus or minus one percent), and rapidly increasing errors for the elderly (up to 10 percent after ten years for the age group 80+). The overestimation of the elderly diminishes somewhat after correction for base population errors, but mortality assumptions have obviously been too optimistic.[11]

Europe and Northern America

Quite striking are the relatively large errors for Europe and Northern America since 1965: strong positive errors (overestimations) at young ages, and equally strong negative errors (underestimations) at more advanced ages; see Figures 7 and 8. The overestimations among the young in the two regions started around 1965, and were caused by unforeseen sharp declines in birth rates. The elderly were underestimated because forecasters have been too pessimistic regarding mortality, in particular for women. This error pattern has

FIGURE 7 Mean percentage errors (MPE) in projections of age structures, Europe

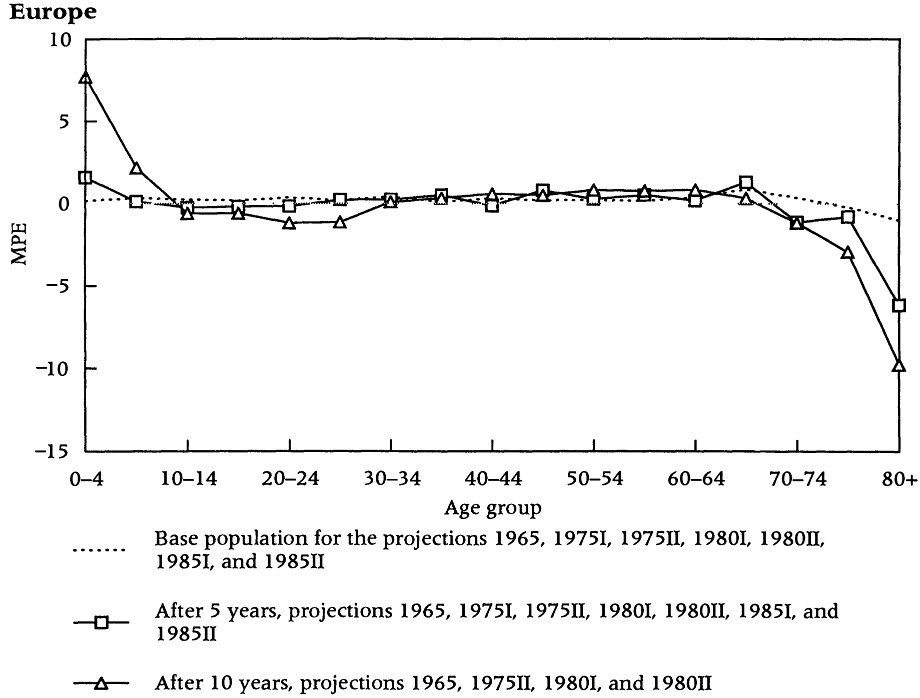
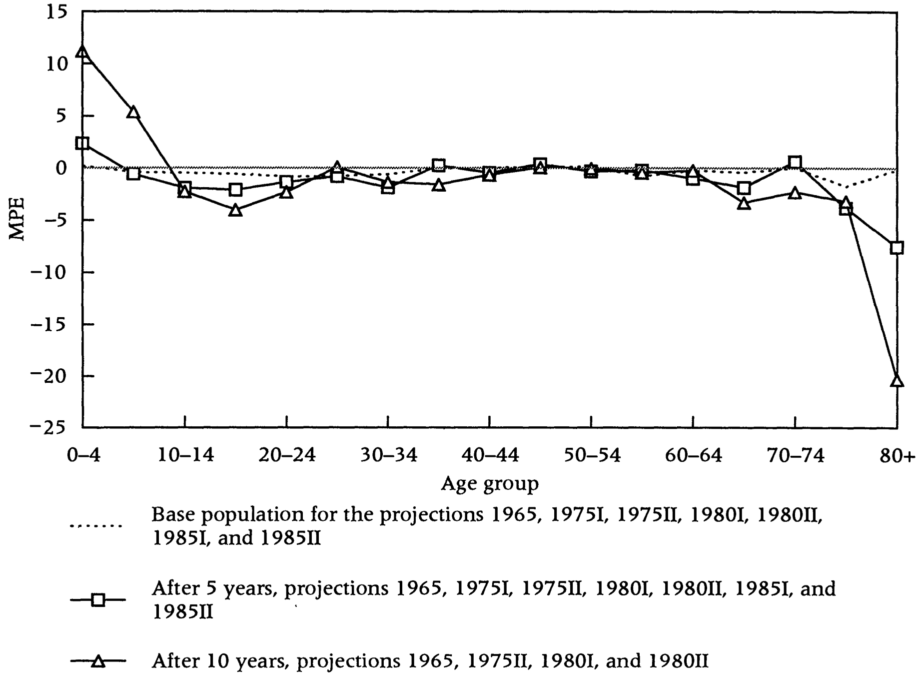


FIGURE 8 Mean percentage errors (MPE) in projections of age structures, Northern America



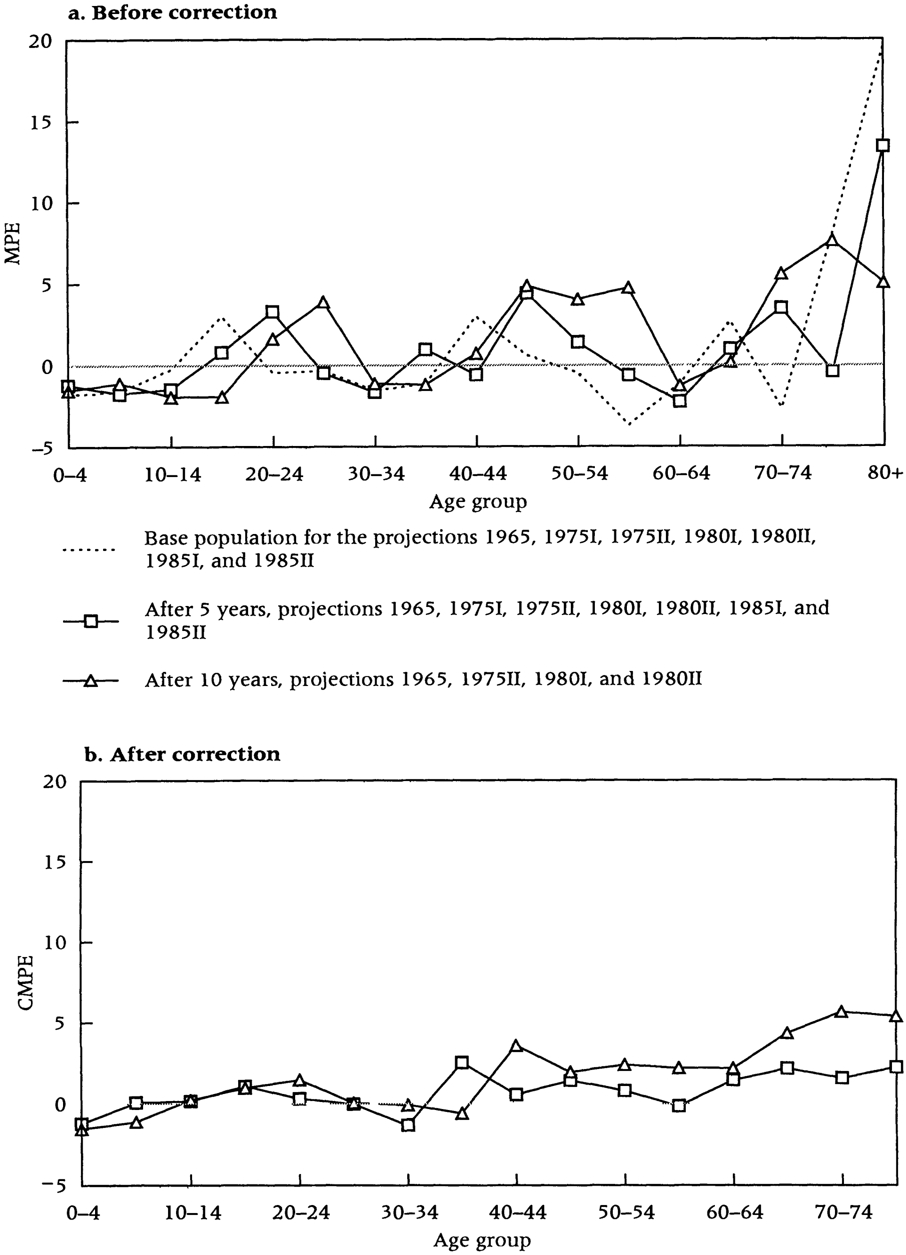
also been established for various individual countries: Canada, Denmark, the Netherlands, Norway, and the United Kingdom (see Keilman 1997). European birth forecasts of 1965 and 1970 were too high by 0.30 and 0.20 percentage points, whereas the average error (see Table 4) was 0.17 percentage points. The overestimation of birth rates was even stronger in Northern America.

Unforeseen levels of immigration have caused moderate errors at ages between 20 and 30 for the two regions.[12] The negative values indicate that immigration has been underestimated. Because base-year populations show only minor errors in Europe and Northern America, corrected errors are not very different from uncorrected errors.

The former Soviet Union

There are two striking features in the error patterns of forecasts for the former Soviet Union in Figures 9a and 9b. First, the errors are much more irregular than those for other regions, both before and after correction for base population errors. A possible explanation is a sudden recent improvement in data quality. The United Nations notes the problems with mortality data in particular (United Nations 1995: 25). Second, there is a sub-

FIGURE 9 Mean percentage errors (MPE and CMPE) in projections of age structures, Soviet Union



stantial overestimation of the age groups 65 and older. The death rates are too low in nearly all forecasts for which we have data. Mortality trends in the former Soviet Union indicated significant health problems before its dissolution. Between 1970 and 1985, life expectancy stagnated and, for males, even declined. These trends came unexpectedly, and the earlier levels were not well enough reflected in the available data. The result was that, on average, forecasts for crude death rates were 0.15 percentage points too low (see Table 4). This was caused by the forecasts of 1950III (–0.23) and 1960 (–0.17) in particular. More recent forecasts have smaller errors, and the two forecasts with base year 1985 have been on target during the years 1985–90.

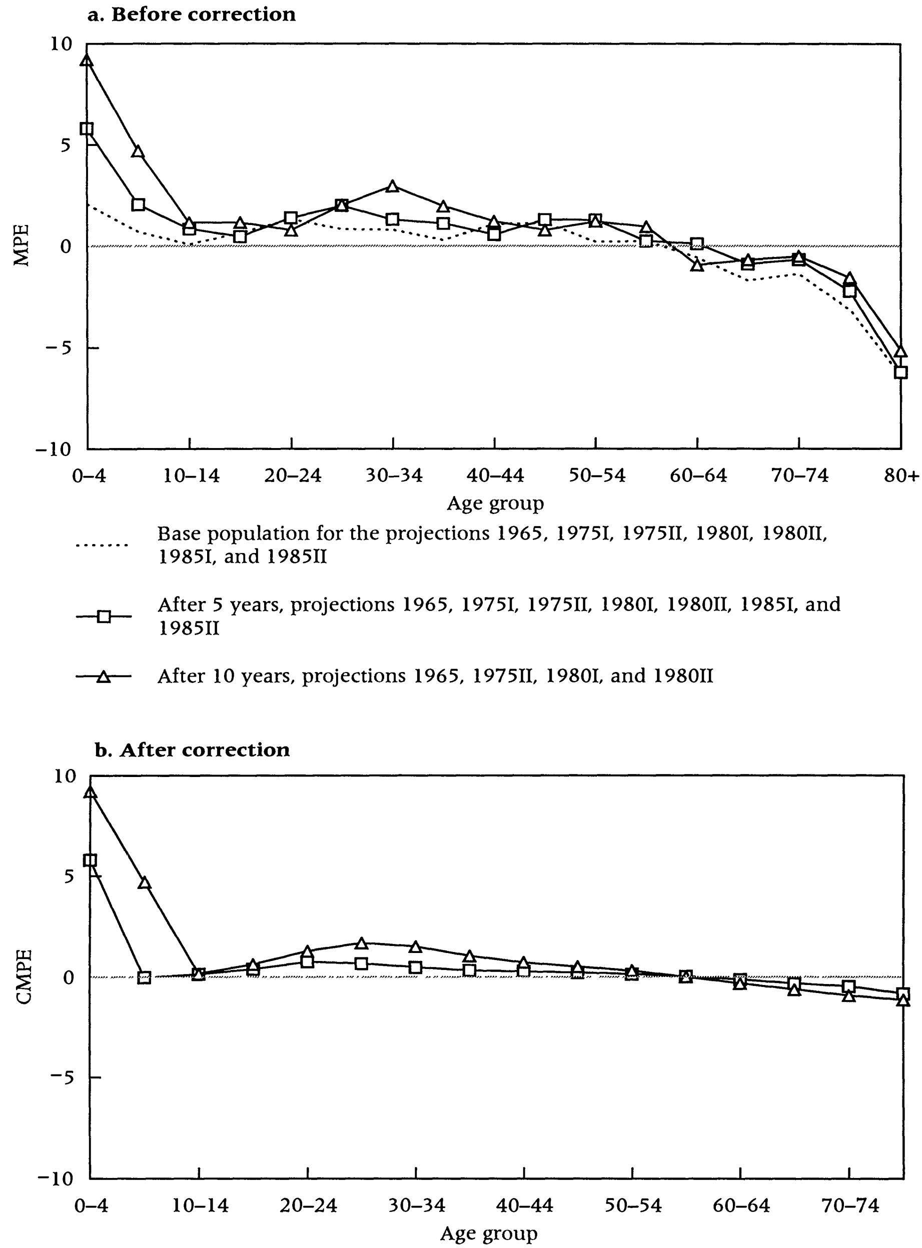
Latin America

For Latin America (i.e., South and Central America, and the Caribbean) the error pattern for the youngest age groups in Figures 10a and 10b is similar to that of Europe and Northern America, the result of an unforeseen fall in birth rates starting in the mid-1960s. Errors caused by wrong mortality assumptions are much more moderate than those due to high birth rates, as witnessed by the age groups above 60 (Figure 10b). Note that the forecasts for adults between ages 20 and 35 have been too high on average, which indicates that outmigration from Latin America was higher than expected. At the same time, forecasts for that age group for Northern America (i.e., United States and Canada, plus three small countries with fewer than 150,000 inhabitants) were too low (see Figure 8), suggesting that at least part of the unforeseen outmigration from Latin America was directed toward Northern America. The UN forecasts do not include assumptions on migration between countries, only net immigration or net emigration for a single country. Therefore this issue could not be analyzed further.[13]

Oceania

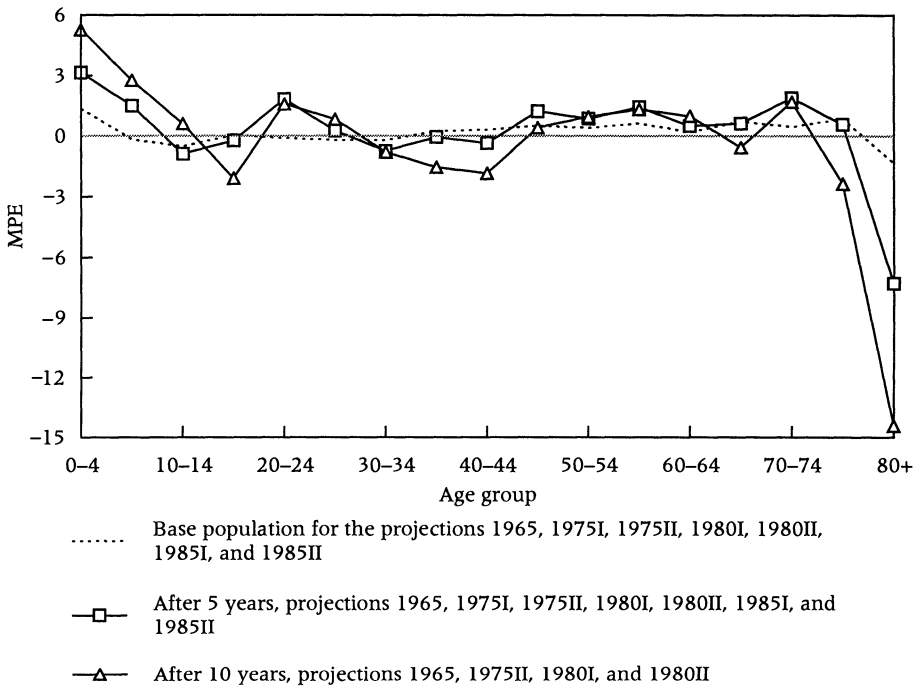
The general pattern of errors in age structure forecasts since 1965 for Oceania is similar to that for Europe, Northern America, and Latin America: too many children (ages 0–10), too few elderly (75+), and quite accurate forecasts for the intermediate age groups (see Figure 11). The overestimation of children at a forecast duration of up to ten years in Oceania is smaller than that in the other three regions. The crude birth rates for forecasts since 1965 up to ten years ahead were too high by 0.10 percentage points. For Europe and Northern America, the corresponding errors were 0.06 and 0.09 points, respectively. Thus the overestimation of fertility in Oceania was larger than that in the other two regions—yet the error in the number of children was smaller. I assume that unexpected immigration to Oceania also contributed to the forecast errors for children under ten years of age. Between 1965 and 1990, the two typical immigration countries, Australia

FIGURE 10 Mean percentage errors (MPE and CMPE) in projections of age structures, Latin America



and New Zealand, represented 77–80 percent of Oceania's population. Immigration is not included in the forecasts for these countries (nor for the other countries of Oceania). This results, all other things equal, in an un-

FIGURE 11 Mean percentage errors (MPE) in projections of age structures, Oceania



derestimation of migration-sensitive age groups, namely, 0–10 and 20–45. For young children, the net result of too high fertility and no immigration was a moderate overestimation. My assumption receives some support from the negative errors for age groups 15–19 and 30–44 in Figure 11, although wrong mortality assumptions for these and neighboring ages may have blurred the picture. The lack of reliable migration data meant this issue could not be analyzed further.

Conclusions and recommendations

An overall assessment of the quality of the United Nations population projections should include more aspects than just their accuracy. For instance, the strength of these projections is that they contain detailed results for each country of the world computed on the basis of one consistent methodology, and that they are updated frequently in the light of the most recent data and methods. Is the accuracy of the UN projections a weak point? I do not think so. When compared with ex post observed real trends, the 12 sets of population projections that the UN prepared between the 1950s and the end of the 1980s show a clear tendency over time toward greater accuracy, and there is no reason to believe that projections made by others would do substantially better. Part of the accuracy improvement is attributable to better data for base populations. The number of elderly persons for the world as a whole has been systematically underestimated by up to 8 percentage points in the base populations for the forecasts made in the 1960s and 1970s. This feature disappeared in the 1980s, when errors in the world's base population became as small as between –1 and +1 percent in almost every five-year age group. Not only the base population, but also implied growth rates have become more accurate. Forecasting was difficult in the 1950s and 1960s, when growth rates were too low by 0.3–0.5 percentage points. In recent years, the growth has been overestimated, but only by 0.2 percentage points or less.

There is considerable regional heterogeneity in the accuracy of the UN projections. Even on the aggregate level of only seven major regions (Africa, Asia, Europe, the former Soviet Union, Northern America, Latin America, and Oceania) and a few large countries (India, China, and the United States) the differences are clear. Mortality has been relatively difficult to project in Africa and Asia. For fertility, errors for Asia, Northern and Latin America, and Oceania are above the world average.

For an appropriate appreciation of these findings, one has to remember that not only the forecasts, but also data on actual population trends are uncertain. Future revisions for countries with poor data for the period 1950–90 may lead to different conclusions.

How can knowledge of these errors observed for historical forecasts be applied when one tries to assess the uncertainty connected to the current or future forecasts made by the United Nations? Can we simply assume that forecasting today is as difficult as it was in the past, and use the historical errors? At first sight, this would be too conservative. After all, there is a clear improvement in accuracy over time. Baseline errors have become less important for developing countries, and for the developed countries there is no sign that the sharp fall in birth rates that occurred in the 1960s and 1970s will repeat itself, or that it will be reversed. Yet we have to be prepared for surprises. The UN projections, like most projections produced by official agencies, are surprise-free. They are based on an assumption of steady social and economic development. In reality we have seen that unanticipated trends can suddenly appear. The decline in life expectancy in Central and Eastern Europe is an example. Migration from countries affected by war, famine, or simply unfavorable economic development is another case in point. Therefore it would be good to keep the historical errors in mind when we make a best guess of the predictability of the world's population. Simply considering the UN's low, medium, and high variants is not enough, as these do not take account of errors in mortality, migration, or base population data—only uncertainty around fertility levels is included. This leads to the following recommendation.

The United Nations ought to include more than one variant for mortality in the projections. A comparison between observed and projected life expectancies at birth in Table 5 shows that mortality assumptions, with the exception of the 1985II series, have been too pessimistic. But there was no systematic lag between observed and projected values. For instance, the life expectancies contained in the 1965 projections come closest to observed ones, whereas those of the projections with base years between 1970 and 1985I are much lower. As late as 1985I, the underestimation was 1.5 years for the period 1985–90. Two years later when the 1985II projections were prepared, the assumed level was suddenly too high by 0.9 years. In summary, Table 5 clearly shows the difficulty of extrapolating mortality. At the same time, this component can cause important errors in forecast results. The findings in this chapter have shown that errors are relatively large for the elderly in Europe, Northern America, and the former Soviet Union. For the near future, the AIDS epidemic only adds to the uncertainty, as no national life tables that include AIDS-induced mortality are yet available for high-prevalence countries.

Thus difficulties with extrapolation together with considerable impact on the results necessitate the use of several sets of mortality assumptions, for instance a low, medium, and high set of life expectancies. This can be implemented quite easily, as shown by the practice of national agencies in several countries.[14] High, medium, and low sets of projection results can be prepared for each country by combining high fertility with high life expectancy, medium fertility with medium life expectancy, and low fertility with low life expectancy.[15] Then the high and low projection variants can still be "thought to bracket the probable range of future population change

TABLE 5 Life expectancy at birth (both sexes), observed since 1965 and projected, world total

	1965–70	1970–75	1975–80	1980–85	1985–90
Observed	55.9	57.9	59.7	61.3	63.0
Projected, base year					
1965	53.1	55.5	58.1	60.4	62.5
1970		55.2	57.2	59.0	60.7
1975I			57.4	59.2	61.0
1975II			57.5	59.2	60.8
1980I				58.9	60.4
1980II				59.5	61.1
1985I					61.5
1985II					63.9

for each country" (United Nations 1993: 84), as is the case with the variants in current UN forecasts.

Notes

Research assistance by Svenn-Erik Mamelund and comments by Helge Brunborg and by participants at the IIASA seminar "Rethinking International Population Projections" in June 1996 are gratefully acknowledged. This project has been supported by grant no. 111939/730 of the Norwegian Research Council.

1 For most countries migration is relatively low, and the United Nations did not assume any future migration until the first round of projections starting from 1990. The revised 1990 round, however, accommodated the continuing extensive migratory movements within Europe and the rapid growth in the number of refugees in Africa and elsewhere, and assumptions on migration were heavily revised.

2 Because the forecasts evaluated here apply to the period 1950–90, before the dissolution of the Soviet Union, I chose to maintain the Soviet Union as one of the regional entries. Observed numbers published in United Nations (1994) for the 15 states of the former Soviet Union have been aggregated.

3 No distinction is made in this chapter between a forecast and a projection.

4 In general, a corrected percentage error for age group x in year t, written as $CPE(x,t)$, is the difference between two percentage errors $PE(x,t)-PE(x-n,t_0)$, where $t_0=t-n$ represents the base year and n is the forecast duration.

5 The correction procedure assumes that the error in the base population is independent of that caused by wrong assumptions. Although this may seem a reasonable assumption, it is not always valid. An example of dependence between errors in the base population and those in extrapolated mortality levels was given earlier for the case of China.

6 Inoue and Yu (1979) have indicated that until the 1965 rounds, the dominant source of error was in the baseline data. Starting with the 1965 round, the errors in the base population were relatively small, so that the major source of error was in the forecast assumptions. On the other hand, El-Badry and Kono (1986: 39) locate the diminishing importance of base population errors in the 1970s—a decade later than Inoue and Yu did. These apparently conflicting conclusions must be seen in the light of the revision of data for real trends. Inoue and Yu evaluated their forecasts against data as assessed in 1978, whereas El-Badry and Kono had access to the 1984 assessment. The fact that I used data as assessed in 1994 explains why, according to my findings, it took even longer than El-Badry and Kono assumed before the errors in the base populations became relatively small. I found that the percentage errors in the base population of world forecasts with base years between 1965 and 1980I are between –4 and +3 percent for ages up to 35 (with clear cohort effects when subsequent forecasts are compared). For higher ages the pattern is sloping downward, to between –4 and –8 percent for the elderly. The underestimation of the elderly is a consistent feature of the base populations for the forecasts made in the 1960s and 1970s, but it disappears with the 1980II forecast. From then on, errors are between –1 and +1 percent at almost all ages, and clear age patterns are no longer present. If future revisions lead to even higher estimates for the number of elderly in the world during the 1980s, base population errors of forecasts 1980II and later will of course become larger than shown here.

7 Because the three dimensions are correlated, simple marginal averages do not give an entirely correct impression of the errors by period, duration, or base year. For instance, the mean errors for the 1950s contain errors for short durations only, whereas mean errors for the 1980s contain errors for long durations as well. Since errors tend to grow when duration increases, the means for

subsequent periods are not entirely comparable. A multivariate (or age-period-cohort-type) model may be used to disentangle the effects of period, duration, and base year and to obtain better estimates for the effects of these three dimensions (see Keilman 1991 and 1997 for applications). There are so few entries in Table 2, however, that precise estimates are hard to obtain.

8 Because all errors have the same sign, each mean error is identical with the corresponding mean absolute error (MAE), and the MAEs are not presented here.

9 Uncertainty concerning mortality and fertility rapidly increases when we look further into the future. Also when analyzing the net effect of fertility and mortality, that is, world population growth rates, an increase in forecast uncertainty was noted. Mean absolute errors in projected annual growth rates were 0.13, 0.20, 0.24, 0.31, and 0.24 percentage points for durations 0–5, 5–10, 10–15, 15–20, and 20–25 years, respectively. On the other hand, Keyfitz (1981) and Pflaumer (1988) found that the error in the growth rate was more or less independent of forecast duration. These apparently contradictory findings are explained by different definitions of growth rates. Keyfitz and Pflaumer analyzed overlapping durations (0–5 years, 0–10 years, 0–15 years, etc.), whereas I looked at subsequent durations (0–5, 5–10, 10–15, etc.). Errors computed for overlapping durations show less variability than those defined for subsequent durations.

10 Too low infant and child mortality may also have contributed to the error.

11 Too high assumptions for Chinese death rates since 1965 and too low birth rates since 1975 have resulted in an underestimation of annual population growth. This is in contrast with the finding of Inoue and Yu (1979), who reported an overestimation of population growth in China during the 1960s and 1970s. The opposite conclusions are explained by the revision of the estimates of China's actual population size. For instance, in 1970 the estimated mid-year population as of 1970 was 760 million; five years later the estimate for 1970 was 772 million. Similarly, the estimate of the 1975 population size increased from 839 to 895 million between 1975 and 1980. Thus actual population growth in China in the 1960s and 1970s was stronger than was thought 20 years ago.

12 The underestimation in the age group 15–19 in Northern America is probably attributable to too pessimistic mortality assumptions, in particular regarding the so-called accident hump in the age pattern of mortality.

13 For every five-year period since 1955, natural growth in Latin America was higher than total population growth (see United Nations 1995: 502). If there had been no errors in the measurement of births, deaths, and total population, this would indeed indicate net emigration. But census under-enumeration combined with underregistration of births and deaths makes it problematic to assess the direction of migration streams, let alone the levels.

14 Several national agencies that prepare official population projections work with more than one mortality variant, and the number of such agencies is increasing. For instance, a survey of 30 statistical agencies in industrialized countries carried out in the mid-1980s showed that eight of these prepared several mortality variants (Hämäläinen 1992: 82). Four of the 15 countries that currently make up the European Union (EU) did so at that time. By 1994, nearly half of the EU countries worked with more than one mortality variant (Cruijsen and Keilman 1994).

15 This would give the greatest range, but it can be criticized from a statistical viewpoint (e.g., see the chapter by Lee in this volume). However, it is rather common practice among statistical agencies.

References

Coale, Ansley J. 1983. "A reassessment of world population trends," *Population Bulletin of the United Nations* 14-1982: 1–16.

Cruijsen, Harri and Harold Eding. 1995. "The 1990-based long-term population scenarios for the European Economic Area and Switzerland: How good were they in the short run?" paper presented at the Eurostat/Statec/CEPS/INSTED Seminar "New Long-Term Population Scenarios for the European Economic Area," Luxembourg, November.

Cruijsen, Harri and Nico Keilman. 1994. "Population projections by statistical agencies in countries of the European Economic Area," paper presented at the Annual Meeting of the Population Association of America, Miami, May.

El-Badry, M. A. and Shigemi Kono. 1986. "Demographic estimates and projections," *Population Bulletin of the United Nations* 19/20-1986: 35–44.

Frejka, Tomas. 1994. "Long-range global population projections: Lessons learned," pp. 3–15 in Wolfgang Lutz (ed.), *The Future Population of the World: What Can We Assume Today?* London: Earthscan Publications.

Hämäläinen, Hannu. 1992. "International comparison of mortality assumptions," pp. 75–86 in Nico Keilman and Harri Cruijsen (eds.), *National Population Forecasting in Industrialized Countries.* Amsterdam/Lisse and Berwyn, PA: Swets and Zeitlinger Academic Publishing Division.

Inoue, Shunichi and Y. C. Yu. 1979. "United Nations new population projections and analysis of ex post facto errors," paper presented at the Annual Meeting of the Population Association of America, Philadelphia, April.

Keilman, Nico. 1991. "Analysing ex-post observed errors in a series of population forecasts," *Zeitschrift für Bevölkerungswissenschaft* 17(4): 411–432.

———. 1997. "Ex-post errors in official population forecasts in industrialized countries," *Journal of Official Statistics* 13(3): 245–277.

Keyfitz, Nathan. 1981. "The limits of population forecasting," *Population and Development Review* 7(4): 579–593.

Lee, Ronald D. 1991. "Long-run global population forecasts: A critical appraisal," *Population and Development Review* 16 (Supp.): 44–71.

Pflaumer, Peter. 1988. "The accuracy of U.N. population projections," pp. 299–304 in *Proceedings, Annual Meeting, American Statistical Association, New Orleans, August 1988, Social Statistics Section.*

Poursin, Jean-Marie. 1994. "Les fausses surprises de la démographie mondiale," *Futuribles* 183 (janvier): 19–31.

Stoto, Michael A. 1983. "The accuracy of population projections," *Journal of the American Statistical Association* 78: 13–20.

United Nations. 1993. *World Population Prospects: The 1992 Revision.* New York.

———. 1994. *The Sex and Age Distributions of the World Populations: The 1994 Revision.* New York.

———. 1995. *World Population Prospects: The 1994 Revision.* New York.