Visible and Invisible Sources of Error

in World Population Projections

Rodolfo A. Bulatao

Consultant
8201 16th St.
Silver Spring, MD 20910
Bulatao@erols.com

Visible and Invisible Sources of Error in World Population Projections

Summary

World population projections have been quite accurate for the world as a whole but less accurate for individual countries. Where they have gone wrong has been explored by the National Research Council (NRC, 2000), which apportions the blame among inaccurate base populations and mistaken fertility, mortality, and migration trends. I reexamine the validity of their results using an alternative method for partitioning error in projections. This decomposition method also allows further analysis of the sources of projection error.

The decomposition leads to slightly different estimates of the contribution of different factors to projection error. Among sources of error in shorter projections, base error is less dominant than in the NRC results. Fertility and mortality error are both somewhat more prominent, and migration error slightly less. The NRC approach probably represents a fair representation of the factors in error overall, emphasizing, perhaps naturally, those country projections where more error exists. The current approach, in contrast, appears to be a somewhat more accurate representation of the situation for the average country, where error is more limited.

I also consider an alternative decomposition of error that serves to emphasize, perhaps even more than the NRC analysis, the importance of the demographic data from which one starts. If one adds to base population error the error due to the starting levels of fertility, mortality, and net migration, the sum is more than half of all error in projections as long as 25 years. While it is important to accurately predict trends in growth components from whatever starting point is chosen, the choice of the starting point is at least as important.

This analysis permits an examination of both visible and invisible error. The error due to trend reversals in component rates tends to be invisible. Projected component trends in a single country projection can be for one period higher and for the next period lower than the actual trend, therefore being offset. Such trend reversal error is much more likely for migration than for fertility or mortality and increases the overall error.

I also investigate the much larger error that becomes invisible when components offset. Offsets are most likely between fertility and mortality, perhaps because socioeconomic development, accelerated or decelerated more than expected, may influence fertility and mortality decline in parallel. I find in fact that too many projected births goes more often with too many than with too few projected deaths. Forecasts have tended to simultaneously underestimate national progress in reducing fertility and reducing mortality, or, less often, to overestimate both together, in either case offsetting error in one with error in the other.

I also examine the way these errors in different groups of countries contribute to error in world projections. World error tends to be much smaller than the error for the average country because country errors have tended to offset. World error also tends to be explained by errors for a few large countries. For the typical forecast, about half of world absolute error is due to only six, usually large, countries: the three with the largest positive and the three with the largest negative errors. Though these countries are not identical across forecasts, four countries have been particularly important contributors to world projection error: China, India, Nigeria, and the U.S.A.

The decomposition of projection error used here is simple and fairly direct, allowing more extensive analysis than the regression-based decomposition used by the NRC. Since it gives fundamentally similar results, the implications for forecasts are similar. Improvement of current population estimates continues to be important in making projections more accurate; fertility trends need to be carefully considered, and approaches need to be developed to predict turning points in fertility; international migration trends deserve more attention than they usually receive, though projecting them (and even obtaining current data) is obviously difficult; and mortality trends deserve some attention too, particularly for longer-term projections. I also note the considerable importance of choosing the proper starting point from which to project component rates.

World projections would benefit from continuing improvements in base data. But improvement in some countries, coupled with lack of improvement in others, could have unexpected effects on projections. Forecasters remain at the mercy of unpredictable tradeoffs between errors in some countries and errors in others.

Visible and Invisible Sources of Error in World Population Projections

1. World population projections have been quite accurate for the world as a whole but less accurate for individual countries (National Research Council 2000). Where they have gone wrong has been explored by the National Research Council, which apportions the blame among inaccurate base populations and mistaken fertility, mortality, and migration trends. Their results are not entirely as expected. For instance, they show migration error to contribute as much to error in projections as fertility error. We reexamine the validity of their results using an alternative method for partitioning error in projections. This method also allows further analysis of the sources of projection error.

Background

2. In a comprehensive assessment of world population projections, the National Research Council (2000) has reviewed worldwide trends in fertility, mortality, and international migration and attempted to determine whether these support the future trends assumed in current world forecasts. It has also looked at past forecasts, to see how well these have fared.

3. The verdict on past forecasts is generally positive. The council shows that 11 world forecasts made since 1957 missed the year 2000 population by an average of only 2.6 percent. All but one forecast came within 4 percent of the actual number, a display of substantial accuracy given that some forecasts were longer than 40 years.

4. Yet, as the council recognized, these modest errors conceal much larger ones. World projections are a composite of projections for individual countries. Council analysis of country projections indicates that, in projections of 5 to 15 years, country projections have on average seven or eight times the error of world projections. In addition, the error in any projection combines errors in the components of population change: fertility, mortality, and migration. These projected components, as well as base population estimates, may have offsetting errors that do not appear as mistaken population forecasts. In order to understand where projections go wrong, therefore, it is useful to partition the error in them, paying attention to both visible error and errors that offset, therefore becoming invisible.

5. The council reported a partitioning of the variance in projected population error from multivariate analysis. Assessing selected forecasts of the 1970s to the 1990s from the U.N., the World Bank, and the U.S. Census Bureau, the council first estimated the error in population projected to 1975, 1980, 1985, and so on up to 2000. Then it attempted to account for this error in a multivariate context using parallel estimates of error in the factors that enter a cohort-component projection: base population, total fertility, life expectancy, and net migration. This approach contrasts usefully with approaches in which base population error is ignored in order to concentrate on the growth rate (e.g., Stoto 1983), providing instead a more complete partitioning of the variance in projected population error (see Figure 4 below).

6. The council interpreted its results to indicate that base error contributes heavily to projected population error in the short run. As the importance of base error declines with projection length, the importance of misspecified component trends increases, accounting for five times as much variance as base error in 30-year projections. Of the components of population change, fertility trend misspecification accounts for three to four times as much error as mortality misspecification, though both increase in

importance with projection length. Migration error is at least as important as fertility error, though its effect is essentially constant by projection length (National Research Council 2000:48-49). These results can be reconsidered, and further insight into visible and invisible error can be provided, if we partition error in a different way.

Method

7. The analysis relies on direct partitioning of error using crude rates. Unlike the error in total fertility and life expectancy, error in crude rates can be expressed in the same metric across components.

8. **Decomposition.** Let P_t* stand for the projected population after t years. One measure of error is the ratio of P_t* to the actual population at that date, P_t, which can be expressed as a product of error in the base population (P_0) and in the growth rate r.

$$P_t^*/P_t = (P_0^*/P_0)\ (e^{r^*t}/e^{rt}) \qquad (1).$$

Growth rate error can be further decomposed to show how much is due to the birth (b), death (d), and net migration[1] (m) rates:

$$P_t^*/P_t = (P_0^*/P_0)\ e^{(b^*-b)t}\ e^{(-d^*+d)t}\ e^{(m^*-m)t} \qquad (2).$$

To obtain additive components of error, we take natural logs:

$$\ln(P_t^*/P_t) = (\ln P_0^*-\ln P_0) + (b^*-b)t + (-d^*+d)t + (m^*-m)t \qquad (3),$$

or, if the projection covers i periods of length y,

$$\ln(P_t^*/P_t) = (\ln P_0^*-\ln P_0) + y3(b_i^*-b_i) + y3(-d_i^*+d_i) + y3(m_i^*-m_i) \qquad (4).$$

9. This decomposition of error is difficult to compare with the National Research Council results from multivariate analysis. Take a hypothetical example. Assume that, in a set of countries, only one component contains any errors. The error in projected population, therefore, would be entirely due to this one component, which a multivariate analysis should confirm. Yet, it could be the case with the current decomposition that the error in this component, averaged across countries, would be zero, if positive and negative errors cancel.

10. We avoid this problem by taking absolute values for each component in the last equation; we designate these figures as U or unsigned estimates of error in different components. This is of course no longer a decomposition of $\ln(P_t^*/P_t)$ but of a new quantity that is the sum of the absolute values of the terms on the right hand side. We apply the term U or unsigned error also to this sum. We do not label the sum absolute error because it differs from what usually goes by that name. For instance, the quantity referred to

[1] The net migration rate, actual and projected, is easily obtained given the growth, birth, and death rates. The National Research Council (2000:247) argues that these estimates, though usually obtained as residuals, are still worth analysis.

as mean absolute proportional error in the literature (e.g., Pflaumer 1988) involves taking absolute values of error by country or other unit rather than by component per country.

11. Some error is still concealed: that due to trend reversals, when, for example, the birth rate is overprojected for one period but underprojected for the next. If one is concerned only with the final projected population, this invisible error is inconsequential. However, if one is also concerned with population projected for intermediate periods, or with the age structure of the final population, the error due to trend reversals is worth taking into account. To include such error in the estimates, we take absolute values of the period-by-period component error in the last equation. That is, we take absolute values before rather than after summing errors across periods. We designate the resulting estimates as T or total trend estimates of error.

12. Additional error may still be concealed. In forecasting a trend in one growth component, the forecaster has to choose a starting level for the first period and a slope (not necessarily linear) to define levels for subsequent periods (e.g., Keilman 2000). Either or both may be wrong. If they are wrong in opposite directions, some of the offsetting error will not be visible. Take for instance, the term for error in births from the last equation, which can be reexpressed, assuming the projection covers k periods, as

$$y3(b_i^*-b_i) = yk[b_1^*-b_1] + y[b_2^*-b_2-(b_1^*-b_1)] + y[b_3^*-b_3-(b_1^*-b_1)] + \ldots + y[b_k^*-b_k-(b_1^*-b_1)] \quad (5).$$

This amounts to adjusting the birth rate error in each period based on the error in the starting period of the projection. The error in the starting period can be considered the starting-level error, while the remaining error, if any, for subsequent periods can be considered the slope error. Adding up these errors, after taking absolute values of each term in square brackets, gives an estimate of what we label S for summed (starting level plus slope) error. S error should be greater than T error. It incorporates error from trend reversals and, in addition, error when the starting level for a component is off in one direction and the projected slope is off in the other direction: for instance, a too high starting birth rate being matched with projected decline in the birth rate when it actually rises.

13. The U or unsigned estimates make visible some error that does not appear in projected country populations: when component errors have opposite signs, some of this error eventually cancels. U estimates indicate which components are the least reliable. They do not help explain, however, why a projection is too high or too low. For instance, the largest raw error could be in births and could be positive, but if the errors in the other components were all negative and entirely offset birth error, the final population could have negative error that would not be explained by the positive error in births.

14. To estimate sources of error while eliminating offsets across components, we take this approach. Of unsigned estimates with opposite sign, the smaller, in the absolute, is set to zero and the larger is reduced by the size of the smaller. If, on the larger side, more than one component is involved, each is reduced in proportion to its original size. The absolute values of the resulting figures are designated as V or visible error.

15. To illustrate these alternative error measures, consider the birth rates in a two-period country projection (Table 1). In this example (an actual case), the starting birth rate is too high, the raw error being 1.4 points per thousand, the second-period rate too low, the error being -1.9 points. U error for births, therefore, is (.5 x 5 years). T error, based on taking absolute values before summing across periods, is (3.3

x 5 years). S error requires summing the absolute values of starting-level and slope error per period, and is still higher. V error depends on the error in other components. If the algebraic sum of the errors in other components is greater than (.5 x 5 years), then V error for births would be zero.

16. **Aggregates.** These measures apply to country projections. In the global forecasts we examine, countries or territories are the main units projected, and aggregates are obtained by summation. While the same measures can be applied to aggregates, we can go further and decompose aggregate error into the error due to specific countries. Two steps are involved. The amount of error in an aggregate projection due to the projection for country j is defined as

$$E_j = (P_{tj}^* - P_{tj}) / \Sigma P_{tj} \qquad (5).$$

Country error may then be broken down into parts due to base population, births, deaths, and net migrants using proportions from equation (3) above. For instance, the error due to base population would be

$$E_j \, (\ln P_0^* - \ln P_0) / \ln(P_t^* / P_t) \qquad (6).$$

Some of these terms will be positive and some negative. In the case of world projections (we do not look at other aggregates), we generally focus on signed error, keeping all error visible except that from trend reversals and from splitting starting point and slope error. However, we reverse the sign for error in deaths (and call it error in survivors), so that, like all the other measures, a positive error indicates an overprojection of the population.

17. **Data.** The data for this analysis, drawn from the same forecasts evaluated by the National Research Council (2000:245), are seven world forecasts from the U.N. and the World Bank dating from 1973-1994. (Two forecasts they used are left out because of unavailable crude rates.[2]) We assess their accuracy, as did the National Research Council, against estimates of past demographic trends from the 1998 U.N. (1999) estimates and projections.

18. Projections up to 30 years long can be evaluated in these data, with the 1973 projections essentially running from 1970 to 2000. The base year for each projection is assumed to be the year ending in 5 or 0 that precedes the forecast date.

Results

19. Table 2 shows the various estimates of error across all countries and forecasts. For all components combined, U or unsigned error is .119, which translates (taking exponentials) into the typical projection being 13 percent off. (Since we are dealing here with absolute error, it could be either too high or too low.) If the invisible error due to trend reversals in included, the error estimate (T) rises to .134, or 14 percent off. If starting-level and slope error are distinguished and summed, the estimate (S) rises further to .185, or 20 percent off. In the other direction, if offsetting error across components is excluded from U error, the error estimate (V) falls to .072, or 7 percent off. Before looking at the relative importance of components, we

---

[2] Including one from the U.S. Census Bureau. The work of the International Institute for Applied Systems Analysis (e.g., Lutz 1996), where aggregates are directly projected, is too recent to examine.

examine first the reasons why the S, T, U, and V estimates differ.

20. **Visible and invisible error.** In the U (unsigned) error estimates, error appears to be 7 percent smaller on average than in the T (total trend) error estimates. Base population error is not affected, but the error in the other components is, for all projections of longer than one period. The effect should be greater in longer projections, which necessarily involve more intermediate estimates of rates. Figure 1 confirms this. The proportion of T error not visible in U error estimates increases substantially for net migrants in longer projections, reaching 32 percent in 30-year projections. Net migration can vary and even change sign from period to period, and projected numbers may therefore be alternately too high and too low. Error due to trend reversals in deaths and births also increases with projection length, though only at something like half the rate for net migration.

21. Figure 1 also shows the reduction in error as one goes from S to U error. This reduction is two or three times as large as the reduction in going from T to U error. However, the pattern of these reductions by projection length is similar, with greater reductions in error in longer projections, particularly error in migrants.

22. The amount of error from trend reversals varies across countries. For six world regions, the average reduction in going from T to U error is 7-11 percent for births, 6-9 percent for deaths, and 9-20 percent for net migrants. Only for net migrants does the variation across regions appear to be large. Similarly, the average reduction in going from S to U error is 14-22 percent across regions for births, 15-20 percent for deaths, and 24-38 percent for net migrants.

23. To illustrate why S error is so large, focus on slope error between the first and the last projection period. Draw a line between the starting level for a component and its level in the last period of a projection. If this line has a steeper upward angle than the actual trend, regardless of the level at which it starts and of any intermediate values, we can say that the slope is too positive. If it has a steeper downward trend than the actual trend, we say the slope is too negative. Table 3 shows how often these mistakes in slope are combined with particular mistakes in the starting level. We see that errors in slope tend to counter errors in starting level at least as often as they reinforce each other. For migration, in fact, opposing errors are considerably more common than reinforcing errors. If the starting level of net migration is too high, the level is projected to decline too fast or grow too slowly. If the starting level is too low (perhaps too negative), the level is projected to rise too fast (perhaps toward zero) or decline too slowly. Taking these offsetting errors into account makes S error for migration much larger than U error.

24. Going from U to V error, we find that the error concealed when components offset--one leading to too high a projection and another leading to too low a projection--is also quite large: about 40 percent. Offsets are particularly likely between births and deaths. Too-high births go with too-low survival in 41 percent of all cases, and too-low births with too-high survival in 20 percent of all cases (Table 4). As Figure 2 illustrates, the association between errors in these two components, as represented by gamma coefficients, is the most negative association among components and varies only slightly by projection length.

25. In contrast, errors in base population and in births reinforce each other in 51 percent of cases and at least partially offset each other in only 30 percent of cases. (For the remaining cases, base population error is zero.) Some tendency also appears for error in survivorship and error in net migrants to be positively associated.

26. When the offsets are taken into account in calculating V (visible) error, error in deaths is reduced by the greatest proportion--58 percent on average relative to unsigned error. Error in births is reduced least (46 percent). For the other two components, the reduction depends on projection length (Figure 3), being 10 or even 20 percentage points greater in longer than in shorter projections.

27. **Decomposition of error.** How the offsets alter the relative contribution of different components to error is shown in the last panel of Table 2. Avoiding the split between starting-level and slope error and removing trend reversal error increase the relative importance of base error and decrease the importance of migration error. In contrast, removing offsetting component error mainly reduces the relative contribution of mortality error.

28. These effects are largely confirmed when error is examined by projection length (Figure 4). It might be noted in addition, however, that the reduction in the relative importance of mortality error takes place mainly in shorter projections, while in longer projections the relative importance of fertility error appears to increase.

29. Comparisons with the National Research Council results can only be made with projection length held constant, as in Figure 4. These earlier results, which can be considered a partitioning of the R^2 for raw error, are designated R in the figure, with the partitioning represented leaving out joint effects in the council's multivariate analysis as well as unexplained variance.

30. These results differ more from the partitionings of S, T, U, and V error than the latter differ among themselves. In general, the new results reduce the relative importance of base error, especially in shorter projections; somewhat reduce the importance of migration error, particularly in longer projections; and somewhat increase the importance of fertility and mortality error. Nevertheless, the earlier conclusions drawn by the National Research Council remain valid. Base error is still the most important source of error in short projections, declining sharply in importance as projections lengthen. Fertility and migration error are of roughly equal importance, but fertility error increases in importance with projection length, while migration error does not. And mortality error contributes less than these other components to overall projection error.

31. The sharpest contrast between the current decompositions and the council results is in estimating the contribution of base population error in five-year projections. The council analysis was essentially based on the relationship between error in each component and total error in projected population. We can approximate this analysis more closely if we group country projections by level of error. In Figure 5, we distinguish four levels of error: very low U error, under .05; low error, from .05 to .09; medium error, from .10 to .14; and high error, from .15 up. Using these distinctions, we see that the contribution of base population error to total error does increase sharply as total error increases, at higher levels of error becoming much more similar to the council results. (Results are similar for S, T, and V error but are not shown.) For longer projection lengths, the initial contrast with the council results is smaller, and the distinction between these groups has more limited effect.

32. How the importance of sources of error varies across countries can be examined. (The National Research Council does not present such an analysis, but instead does multivariate analysis to examine variation in error in each component separately.) In Figure 6, countries are dichotomized in three ways: by population size in the base year of each forecast; by total fertility in the initial projection period; and by demographic history, in particular whether each country experienced a demographic quake any time since

the 1970s. A demographic quake, in the National Research Council (2000:254-255) usage, is an abrupt change in the population growth rate of at least 2.5 percentage points between successive five-year periods, and generally occurs in fewer than 5 percent of possible cases.

33. The sum of unsigned error clearly differs between these groups of countries, with more error in these forecasts evident where populations are under 1 million, total fertility is 4.5 births or higher, and a demographic quake has occurred. The differences can generally be assigned to specific components. For countries under 1 million in comparison to larger countries, error in fertility and mortality is roughly the same, but error in migration and base population is substantially larger. For high fertility countries in comparison with low fertility countries, error is greater for each of the four components. The relative difference is particularly large for base population error. Among countries that have experienced a demographic quake, we see no greater error in fertility than among other countries. But on the other three components, and especially on migration, countries that have experienced a quake show much greater error.

34. A similar partitioning of sources of unsigned error in Figure 7 is shown for countries classified into six regions: Europe and the former Soviet Union; the Pacific Rim industrial countries (the U.S.A. and Canada, Japan, and Australia and New Zealand); Latin America and the Caribbean; Asia; the Middle East and North Africa; and Sub-Saharan Africa. The regions can be considered in pairs. The first two, of industrial countries, have the least amount of overall unsigned error; the second two, have intermediate levels; and the last two have the highest levels of error.

35. For countries in the industrial regions, migration and fertility are the main sources of projection error. Base populations on average have little error (though somewhat more in the Pacific Rim industrial countries than in Europe), and mortality rates show a fairly steady trend over time with few upsets, therefore being relatively easy to project.

36. For countries in developing regions, error in each component tends to be larger. Between projections in Latin America and Asia, errors tend to be larger in base population and migration for Latin America, in mortality for Asia. Countries in the remaining two regions have the most unsigned error. Migration is an especially important source of error for the Middle East and North Africa, base population and mortality especially important for Sub-Saharan Africa.

37. An alternative decomposition can be provided of S error, separating starting-level error from slope error (Figure 8). Like base population error, starting-level error in births, deaths, and net migrants becomes proportionally less important in longer projections, as slope error becomes proportionally more important. Even in 30-year projections, however, base population error and starting-level error combined still account for half of all error.

38. **Country contributions to world projection error.** These comparisons of regions have involved not regional projections but average country projections within each region. Where projections of aggregates are concerned, one can assess the contribution of projections for individual countries to aggregate error. We look not at regional aggregates but only at the world projection as a single aggregate.

39. We have 27 world projections to examine, with the earlier forecasts contributing more cases. Average signed or visible error across these projections is only 4.2 per thousand. As Figure 9 shows, positive and negative errors often offset each other across countries. In each forecast, positive error increases with projection length. Negative error increases with length in some cases but not others, and

generally more modestly, so that positive error is greater in longer projections.

40. We can consider the specific country projections that contribute most to projection error in each forecast. We need to examine only a few countries, because a small number accounts for a substantial amount of the error in most forecasts. Taking the three countries that contribute the most positive and the three that contribute the most negative error usually accounts for half the total absolute error, and taking five countries in each case accounts for almost 60 percent of the total error.

41. Table 5 lists five countries responsible for the most positive and the most negative error in each forecast. We show only the results for 5-year and 15-year projections. The list sometimes changes between 5-year (sometimes 10-year) projections and longer projections, but does not change much among longer projections. The percentage of error that each country accounts for is calculated on the base of the total absolute error in each forecast. Typically, countries on the list are large countries, with China and India appearing often. They are also mainly developing countries, though the U.S.A. and the former Soviet Union are prominent sources of error in some cases.

42. Four countries are worth particular attention, appearing most frequently in the lists of top five positive or top five negative sources of error: China, India, Nigeria, and the U.S.A. Each of these countries appears 24 out of 27 possible times, except for the U.S.A., which appears 26 times. Combined, these countries account on average for almost 40 percent of absolute error in world projections.

43. China is a major source of error in every forecast except the 1994 forecast. In fact, in the 1973 five-year projection, more than half of the world total absolute error is attributable to errors in projecting China. Errors for China have almost without exception been negative, involving underprojection. The 1973 forecast involved a substantial underestimate of the base population, which was aggravated by a substantially higher projection of mortality than actually occurred (see the Appendix Table). This mortality error was not repeated in later forecasts. But the majority of later forecasts did contain negative base error, aggravated in the 1983 forecast by underprojection of net migration. The 1980 and 1994 forecasts were the exceptions, where base error was not important, and for the 1984 forecast it was relatively small. In two of these forecasts where base error was less prominent, however, fertility was underprojected, providing a similar though smaller contribution to world error.[3]

44. India has produced less error than China, and the direction of error has been less consistent. In the 1973 forecast, India was the leading source of positive error for longer projections. In several subsequent forecasts it was either the leading source of negative error or second in this to China. The base population was badly underestimated in the 1973 and 1983 forecasts. In the 1973 forecast, however, this error was submerged under a much larger overprojection of fertility. For the 1980 forecast, the problem was both underprojection of fertility and overprojection of mortality. In the 1983 forecast, underprojection of fertility was even worse, but mortality was less of a problem.

---

[3] The case of base error for China illustrates why we do not comment on contrasts between the U.N. and World Bank forecasts. There appears to be less base error for China in the more recent U.N. forecasts in comparison to the World Bank forecasts. However, the two agencies disagree about the actual current population. Therefore, using the 1998 U.N. estimates as the standard against which forecasts are evaluated, which we do of necessity, does not provide an impartial standard.

45. Unlike India, the error in Nigeria has a single source: overestimation of the base population. Because of such overestimation, Nigeria is the main source of positive error in most forecasts, accounting for up to a fifth of world absolute error. For some of the longer projections, fertility is also somewhat overprojected.

46. The U.S.A. is a somewhat surprising source of important error in world projections. This error is always negative, typically accounting for 2-5 percent of world absolute error--enough, in some cases, to make it the leading source of negative error. The base population has generally been slightly underestimated, and parallel underprojection of net migration has added to error. Mortality was also overprojected in earlier forecasts though not in later ones, and fertility has been underprojected in some later forecasts. A combination of factors, generally operating in the same direction, has therefore been responsible for the error in U.S.A. projections.

47. With regard to the underprojection of net migration for the U.S.A., it is worth noting that Mexico, the main recent source of migrants to the U.S.A., is the next most frequently listed source of error after these four countries. It has been overprojected, with consistently too high base population and consistently too high net migration. Migration error is of course in a special class where world projections are concerned. If the forecaster properly arranges it, the total of world net migrants will be zero, so that all of the error in net migration will be invisible in a world projection.

48. Since we cannot similarly examine the projection for every other country, we summarize for groups of countries what they contribute to world projection error, averaging across forecasts. Figure 10 shows world projection error broken down between the "top five" countries--the five with the most positive error combined with the five with the most negative error--and all other countries. The top five contribute more negative error and somewhat less--though not a lot less--positive error than all other countries. For the top five, base error, particularly negative base error, is substantially larger than for all other countries. Migration error, however, is substantially smaller. Error in births and survivors is more closely balanced between groups, with the top five having less error in births and slightly more error in survivors.

49. By population size (Figure 11), we note first how small is the contribution to world error of countries under 10 million, even though they include over 60 percent of countries. Countries of 10-99 million contribute large positive error, especially in births, combined with some positive base error. Opposed to this is the large negative error, especially in base population, for countries of 100 million or more, mitigated somewhat by positive fertility error. Projected aggregates of these two groups of countries would show substantially more proportional error than the world aggregate, which combines both groups and allows their opposite errors to offset.

50. By region (Figure 12), we note the important contribution of negative error in Asian countries. Much of this is due to negative base error, but underestimates of survivors (or overestimates of mortality) also contribute. Both types of errors are only partly balanced by large positive errors in births. The second largest regional contribution to total error, from Sub-Saharan Africa, is shown not to be due to any single source but to come almost equally from base error and error in births, with substantial contributions also from error in survivors and migrants.

Discussion

51. Some previous examinations of error in world population projections (e.g., Stoto 1983, Keilman 1998, but see Keilman 2000) have attempted to examine "true" projection error by eliminating the error due to mistaken base estimates. Instead we look at base error as an important part of projection error and compare it directly with the error in component trends. The error in these trends, in addition, is broken down by component.

52. **Alternative partitionings of error.** Reexamining the error in forecasts that date from the 1970s to the 1990s, we have shown that the recent accounting for this error in a National Research Council (2000) report can be duplicated with a simpler approach. We confirm that, as the council concluded, error in the base population used in a projection is a major factor but declines in importance as projections lengthen; fertility error increases in importance with longer projections; migration error is similarly large but does not increase in importance; and mortality error provides the smallest contribution.

53. Although these findings are parallel to the earlier findings, our simple decomposition leads to slightly different figures. Among sources of error in shorter projections, base error is less dominant in our results. Fertility and mortality error are both somewhat more prominent, and migration error slightly less.

54. The reasons for these differences are difficult to explicate. Although we use the same data, the analytical approach is quite different. The council analysis used multivariate analysis to partition proportional projection error, transformed to a more normal distribution (National Research Council 2000:271). The transformation should have affected the relative impact of small and large errors. They attempted to relate this error measure to error in total fertility and life expectancy rather than in crude birth and death rates. The council approach probably represents a fair representation of the factors in error overall, emphasizing, perhaps naturally, those country projections where more error exists. The current approach, in contrast, appears to be a somewhat more accurate representation of the situation for the average country, where error is more limited. This contrast may account for what differences exist.

55. We also show an alternative decomposition of error that serves to emphasize, perhaps even more than the council analysis, the importance of the demographic data from which one starts. If one adds to base population error the error due to the starting levels of fertility, mortality, and net migration, the sum is more than half of all error in projections as long as 25 years. While it is important to accurately plot trends in growth components from whatever starting point is chosen, the choice of the starting point is at least as important.

56. **Visible and invisible error.** Our analysis allow us to examine both visible and invisible error. The error due to trend reversals in component rates, for instance, is not investigated in the council analysis. Projected component trends in a single country projection can be for one period higher and for the next period lower than the actual trend. Adding such trend reversal error to visible unsigned error would raise the total about 8 percent. We show that such trend reversal error is much more likely for migration than for fertility or mortality.

57. We also investigate the much larger error that becomes invisible when components offset. Offsets are most likely between fertility and mortality. The National Research Council (2000:198-199) gives one reason why this should happen: because socioeconomic development, accelerated or decelerated more than expected, may influence accelerate or decelerate both fertility and mortality decline. However, it also offers

a reason for the opposite result: fertility may fall while mortality rises in an unpredicted crisis, an epidemic or famine. What we actually find is that too many projected births goes more often with too many projected deaths than with too few deaths. These forecasts have tended to simultaneously underestimate national progress in reducing fertility and reducing mortality, or, less often, to overestimate both together, in either case offsetting error in one with error in the other.

58. Other pairs of components are less likely to offset. In particular, an overestimate of the base population tends to be paired with an overprojection of births. If base population is not known precisely, too high an estimate of recent fertility could lead to both an overestimate of current population and an overprojection of future fertility. A similar though weaker relationship exists between error in deaths and in net migrants. Too optimistic a view of the death rate goes with overprojection of net migrants--either too many immigrants or too few emigrants or both.

59. In contrast to other components, a greater proportion of error in deaths--more than half--is offset by error in other components. This is partly because error in deaths is smaller in the absolute. But much error in other components is also offset. This underlines the importance, in assessing projection uncertainty, of attending to cross-correlations of component error, as well as to correlations of error over time.

60. **Contrasting countries.** Our analytical approach allows us to easily examine the contribution of each component to error in different groups of countries. The contrasts are largely interpretable. For instance, small countries of under 1 million have much greater migration error than larger countries. Migration can affect a small country dramatically, and absolute net rates, as opposed to actual numbers, can be many times those for even such a traditional receiving country as the U.S.A. Error in predicting such flows can have a substantial effect on projected population.

61. For high fertility countries in contrast to low fertility countries, all components are likely to have greater error, but the contrast is greatest with regard to base population. This suggests that high fertility countries have poorer demographic data. When population estimates that are not current have to be updated, higher fertility levels mean greater scope for error.

62. For countries that have experienced a demographic quake as opposed to countries that have not, error is greater on every component except fertility. Quakes are often related to sudden inflows of migrants (such as oil workers into Persian Gulf states) or sudden outflows (as during military or civil crises). They could also be associated with sudden increases in mortality. Why base populations in such cases also have greater error may be explained in various ways. If the quake occurred before the forecast, the changes in vital rates may make the base population difficult to determine. The underlying events may also have obstructed the collection of adequate demographic data. Finally, countries prone to such quakes may have weaker statistical systems to begin with.

63. Regional contrasts are similarly interpretable, often with reference to the factors just considered. Industrial countries have better base data than developing countries, and with lower fertility and fewer demographic quakes should be easier to project. Fairly sudden flows of migrants, sometimes amounting to demographic quakes, have been an important part of the recent experience of the Middle East. And uncertainty about actual population size and about death rates produces substantial error for Sub-Saharan Africa.

64. **World error.** We examine finally the way these errors in different groups of countries contribute to

error in world projections. World error tends to be much smaller than the error for the average country because country errors have tended to offset. World error also tends to be explained by errors for a few large countries. For the typical forecast, about half of world absolute error is due to only six, usually large, countries: the three with the largest positive and the three with the largest negative errors. Though these countries are not identical across forecasts, four countries have been particularly important contributors to world projection error: China, India, Nigeria, and the U.S.A.

65. The specific factors in erroneous world projections have varied. In general, Asian projections have contributed much error from underestimated base populations, coupled with overprojected death rates. By contrast, Sub-Saharan projections have contributed error from overestimated base populations. Overprojections of births have been a major contributor for large countries of intermediate size (10-99 million), especially in Asia, though the largest countries have also contributed error from this source.

66. **Implications.** The decomposition of projection error we have used is simple and fairly direct, allowing more extensive analysis than the regression-based decomposition used by the National Research Council. It does have one disadvantage: crude rates are not the parameters that forecasters typically use as inputs. They are not refined rates and are affected by age structure.

67. Nevertheless, this simple decomposition gives fundamentally the same results as the National Research Council partitioning. The implications for forecasts should be essentially similar. Improvement of current population estimates continues to be important in making projections more accurate; fertility trends need to be carefully considered, and approaches need to be developed to predict turning points in fertility; international migration trends deserve more attention than they have usually receive, though projecting them (and even obtaining current data) is obviously difficult; and mortality trends deserves some attention too, particularly for longer-term projections. We also note the considerable importance of choosing the proper starting point from which to project component rates.

68. Our investigation of visible and invisible error also confirms the National Research Council's (2000:196-197) attention to intercorrelations of error. Correlated errors in country projections, across time or across components, can reinforce each other, and even when they offset still affect the statistical uncertainty in forecasts. We have gone further, however, in quantifying these effects, verifying which components typically correlate and when correlations are more likely to exist across periods.

69. Taking both visible and invisible error into account in estimating uncertainty is essential. What can or should be done to reduce such error is less obvious. For instance, it is arguable that the larger trend reversal error for migration than for fertility or mortality is not much of a problem. It indicates that migration rates may be more volatile, but the long-term trend may still be fairly well projected. In contrast, a consistent bias--though in which direction for a particular country we cannot say--is more likely in fertility and mortality.

70. Across countries, the positive association between base population error and fertility error deserves some attention. We hypothesize that this is due to the need to update base estimates from earlier censuses. An overestimate, due to higher assumed recent fertility, also affects the projected fertility trend. In such cases, better base data would obviously help, as well as some attempt to estimate base population and the fertility trend independently, to the extent possible, from separate data.

71. We also find a positive association between fertility error and mortality error (i.e., a negative

association between fertility and survivorship). Since fertility and mortality are often assumed to move in tandem, this may be difficult to avoid. Notably, this association could become worse, and appear even more systematic, if simultaneous equation models (Sanderson 1998) were used extensively across countries, since such models usually assume some consistent dependence of fertility on mortality. In a sense, the association may be considered a good thing for a forecast, because fertility and mortality errors cancel. Yet even if projected total population is not greatly affected, such other projected parameters as the age structure, as well the fertility and mortality trends themselves, may be distorted. Remedying this probably requires improved analysis of fertility and mortality, to understand not only their interrelationships but also the way these relationships may be stronger or weaker for particular countries.

72. The implications of this analysis for world projections are somewhat different from those for country projections. For world projections, the importance of focusing on the largest countries to improve accuracy is evident. This may require little additional emphasis, since the proportion of error that is due to a handful of countries, and sometimes to only one or two countries, is quite striking. In addition, however, errors in world projections appear more haphazard and subject to specific historical factors than the error in country projections. Historical factors are important in country projections too, such as the impact of demographic quakes in producing error particularly in migration. With a larger number of cases from which to generalize, error in the country projections can be interpreted in more systematic terms. With fewer cases, the error in world projections is more readily tied to specific historical circumstances.

73. For instance, several forecasts show simultaneous anomalies in too low a base population for China and too high a base for Nigeria. These two errors have offset each other more than once. However, unlike the offset between too low net migration for the U.S.A. and too high net migration for Mexico, the base population errors for China and Nigeria are unrelated and unlikely to be duplicated in current and future forecasts. A large amount of error is offset when countries are combined into aggregates; no fundamental reason exists (except in the case of net migration) why these country errors should offset rather than complementing each other. Current and future world projections could therefore, largely by chance, become less rather than more accurate than in the past if errors for whatever reason offset each other less often.

74. This is not to deny that world projections would benefit from continuing improvements in base data. But improvement in some areas, coupled with lack of improvement in others, could have unexpected effects on projections. Forecasters remain at the mercy of such unpredictable tradeoffs, as well as being at the mercy of unexpected demographic developments.

References

Keilman, Nico
1998 How accurate are the United Nations World Population Projections? *Population and Development Review* 24(Suppl.):15-41.

Keilman, Nico
2000 United Nations population projections 1950-1995: Data quality and accuracy. Paper presented at the 34th meeting of the German Association for Population Science, Bochum, 5-7 April.

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

Lutz, Wolfgang (ed.)
1996 *The Future Population of the World: What Can We Assume Today?* Revised and updated edition. Laxemburg, Austria: International Institute for Applied Systems Analysis.

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

Pflaumer, P.
1988 Confidence intervals for population projections based on Monte Carlo methods. *International Journal of Forecasting* 4:135-142.

Sanderson, Warren C
1998 Knowledge can improve forecasts: A review of selected socioeconomic population projection models. *Population and Development Review* 24 (Suppl.):88-117.

Stoto, Michael A
1983 The accuracy of population projections. *Journal of the American Statistical Association* 78(381):13-20.

United Nations
1999 *World Population Prospects: The 1998 Revision.* Vol. 1, *Comprehensive Tables.* New York: United Nations.

Table 1. Error in births in a two-period country projection: An illustration

Indicator	*Period* *1st*	*2nd*	*Combined*	*Explanation*
Actual birth rate	41.6	42.3		
Projected birth rate	43.0	40.4		
Raw error	1.4	-1.9	-.5 x 5	Projected less actual, summed over periods, times period length
U (unsigned) error			.5 x 5	Absolute value of sum
T (total trend) error	1.4	1.9	3.3 x 5	Sum of period absolute values
Starting-level error	1.4	1.4	2.8 x 5	Assumed constant across periods
Slope error	0.0	-3.3	-3.3 x 5	Raw error minus starting-level error
S (summed) error			6.1 x 5	Sum of absolute values by period for starting-level and slope error

Note: V (visible) error depends on the error in other components. For instance, if the error is positive for base population, survivors, and net migrants and totals more than (.5 x 5), then V error for births would be zero.

Table 2. Error estimates when certain errors are included or excluded: Means across countries and forecasts[a]

Error indicator	*Combined*	*Base population*	*Births*	*Deaths*	*Net migrants*
Means					
S: Summed error = unsigned error + starting level error + slope error	0.185	0.029	0.046	0.025	0.085
T: Total trend error = unsigned error + trend reversal error	0.134	0.029	0.036	0.019	0.050
U: Unsigned error	0.119	0.029	0.034	0.018	0.039
V: Visible error = unsigned error - offsetting component error	0.072	0.018	0.021	0.008	0.024
Coefficients					
e^S	1.203	1.029	1.047	1.025	1.089
e^T	1.143	1.029	1.037	1.019	1.052
e^U	1.126	1.029	1.034	1.018	1.040
e^V	1.075	1.019	1.022	1.008	1.024
Mean ratios[b]					
U/S	0.79	1.00	0.80	0.82	0.68
U/T	0.93	1.00	0.90	0.93	0.85
V/U	0.61	0.50	0.54	0.42	0.50
Mean percentage distributions[b]					
S: Summed error	100	18	31	16	35
T: Total trend error	100	21	31	16	32
U: Unsigned error	100	22	31	17	30
V: Visible error	100	23	33	14	31

[a] Each case is a country projection of a specific length (up to 30 years) from a particular forecast, such as the 1973 U.N. forecast. Total cases are 4,062.

[b] These are means across individual country projections, not ratios or percentages of the preceding means.

Table 3. Starting-level and slope error, by component: Percentage of cases with each type of error

Starting level	*Slope*[a]	*Births*	*Deaths*	*Net migrants*
Too low	Too negative	15	16	12
	Correct	1	2	1
	Too positive	25	20	26
Correct	Too negative	2	3	3
	Correct	0	1	3
	Too positive	3	5	4
Too high	Too negative	22	27	37
	Correct	1	3	2
	Too positive	32	23	14
(Number of cases[b])		(2,987)	(2,987)	(2,987)
Reinforcing errors[c]		47	40	25
Opposing errors[d]		46	47	63

[a] For this table, the slope is determined from the starting level to the level for the last projection period, ignoring intermediate fluctuations.
[b] Differs from the number in Table 2 because one-period projections, with no slope error, are excluded.
[c] Starting level too low and slope too negative, or starting level too high and slope too positive.
[d] Starting level too low and slope too positive, or starting level too high and slope too negative.

Table 4. Agreement between components in the direction of error: Percentage of cases in each cell

	Negative	*Zero*	*Positive*	*Gamma*
Birth rate error	Base population error			
Negative	23	6	10	
Positive	20	13	28	0.42
Survivorship error	Base population error			
Negative	23	12	24	
Positive	20	7	14	-0.13
Net migration rate error	Base population error			
Negative	20	10	15	
Positive	23	9	23	0.09
Survivorship error	Birth rate error			
Negative	18	0	41	
Positive	20	0	20	-0.39
Net migration rate error	Birth rate error			
Negative	18	0	26	
Positive	20	0	35	0.10
Net migration rate error	Survivorship error			
Negative	30	0	15	
Positive	29	0	26	0.28

Note: Zero error for births, deaths, and net migrants involves only 4, 7, and 0 cases respectively. Survivorship error is the inverse of death rate error.

Table 5. Correlations between U error and percentage of U error due to each component, by projection length

Projection length	*Base population*	*Births*	*Deaths*	*Net migration*
5 years	0.37	-0.30	-0.25	-0.02
10 years	0.30	-0.31	-0.20	0.13
15 years	0.25	-0.27	-0.18	0.16
20 years	0.18	-0.23	-0.17	0.21
25 years	0.08	-0.16	-0.10	0.19
30 years	-0.01	-0.11	-0.09	0.19

Table 6. Percentage of absolute error in world projections accounted for by countries with the most positive and negative error, by forecast and projection length

5-year projection				*15-year projection*			
Positive error	*%*	*Negative error*	*%*	*Positive error*	*%*	*Negative error*	*%*
1973 U.N. forecast							
Nigeria	3.5	China	52.7	India	5.3	China	34.0
Afghanistan	2.3	India	4.4	Afghanistan	3.7	Ethiopia/Eritrea	2.8
Sudan	1.3	Ethiopia/Eritrea	3.8	Brazil	3.5	U.S.A.	2.2
Saudi Arabia	1.0	U.S.A.	3.7	Nigeria	3.0	Viet Nam	1.9
Brazil	0.9	Viet Nam	2.7	Indonesia	2.9	Pakistan	1.3
Combined	9.1	*Combined*	67.3	*Combined*	18.3	*Combined*	42.2
1980 U.N. forecast							
Nigeria	14.7	Ethiopia/Eritrea	9.3	Nigeria	8.8	India	12.6
Mexico	2.8	U.S.A.	9.1	Mexico	3.7	China	11.7
South Africa	2.2	India	5.6	Bangladesh	2.8	Ethiopia/Eritrea	4.1
Pakistan	2.0	China	5.0	Afghanistan	2.5	Indonesia	3.9
Burma	1.9	Indonesia	3.7	Brazil	2.2	U.S.A.	3.6
Combined	23.5	*Combined*	32.8	*Combined*	20.0	*Combined*	35.9
1983 World Bank forecast							
Nigeria	15.9	China	16.4	Nigeria	12.3	China	21.5
Mexico	3.1	India	10.3	Bangladesh	5.0	India	6.4
Afghanistan	2.1	Ethiopia/Eritrea	5.2	Mexico	3.3	U.S.A.	3.0
Bangladesh	2.0	Pakistan	4.1	U.S.S.R.	2.9	Ethiopia/Eritrea	2.6
South Africa	1.8	Indonesia/E.Timor	3.6	Morocco	2.7	Pakistan	2.3
Combined	24.9	*Combined*	39.6	*Combined*	26.1	*Combined*	35.9
1984 U.N. forecast							
Nigeria	20.1	China	11.0	Nigeria	12.8	India	13.0
Mexico	3.7	India	9.2	Bangladesh	4.1	China	10.9
Afghanistan	2.1	U.S.A.	4.0	U.S.S.R.	4.1	Pakistan	3.3
Bangladesh	1.8	Iran	3.1	Mexico	2.8	U.S.A.	2.9
South Africa	1.7	Colombia	3.1	Brazil	2.0	Congo, Dem.R.	1.7
Combined	29.4	*Combined*	30.4	*Combined*	25.9	*Combined*	31.8
1988 World Bank forecast							
Nigeria	19.2	China	23.3	Nigeria	19.0	U.S.A.	5.5
Afghanistan	3.4	Pakistan	3.8	U.S.S.R.	7.6	India	4.3
Bangladesh	2.8	U.S.A.	3.2	Bangladesh	5.7	Germany	2.5
Mexico	2.5	Iran	2.9	Brazil	3.5	Pakistan	2.4
Viet Nam	2.2	India	2.9	Viet Nam	3.0	Colombia	1.9
Combined	30.1	*Combined*	36.2	*Combined*	38.9	*Combined*	16.6

Table 6. Percentage of absolute error in world projections accounted for by countries with the most positive and negative error, by forecast and projection length

5-year projection				*15-year projection*			
Positive error	*%*	*Negative error*	*%*	*Positive error*	*%*	*Negative error*	*%*
1990 World Bank forecast							
Nigeria	21.1	China	23.2	Nigeria	20.5	U.S.A.	3.8
Afghanistan	3.9	Pakistan	3.8	U.S.S.R.	6.5	India	2.8
Mexico	2.8	Iran	3.0	Bangladesh	4.1	Germany	2.2
Bangladesh	2.6	U.S.A.	2.7	Brazil	3.4	Colombia	1.8
Brazil	1.6	Egypt	2.6	Mexico	2.7	Congo, Dem.R.	1.8
Combined	31.9	*Combined*	35.4	*Combined*	37.3	*Combined*	12.5
1994 U.N. forecast							
Nigeria	14.4	U.S.A.	4.2				
Iran	5.6	Colombia	3.9				
Pakistan	4.8	Congo, Dem.R.	1.7				
South Africa	4.5	Mozambique	1.6				
Burma	4.1	Russia	1.2				
Combined	33.3	*Combined*	12.6				

Note: Percentages are based on the total absolute error in each world projection.

Appendix Table. Error per thousand contributed to world error by four countries, and amount of error due to different components

Fore-cast	*Projec-tion length*	*Total*	*Base popu-lation*	*Births*	*Deaths*	*Net mig-rants*	*Total*	*Base popu-lation*	*Births*	*Deaths*	*Net mig-rants*
				China					**India**		
1973	5	-21.9	-15.9	-1.5	-4.3	-0.1	-1.8	-3.3	1.3	0.1	0.0
1973	10	-20.6	-15.8	2.5	-7.3	0.0	1.2	-3.3	4.4	0.1	0.1
1973	15	-20.1	-15.5	5.1	-9.5	-0.2	3.1	-3.4	6.5	0.6	-0.5
1973	20	-23.6	-15.3	2.9	-11.0	-0.3	4.8	-3.5	8.2	1.0	-0.9
1973	25	-23.2	-15.0	3.6	-11.6	-0.1	6.4	-3.6	9.4	1.3	-0.7
1973	30	-21.5	-14.7	5.6	-12.3	-0.1	7.6	-3.7	10.4	1.5	-0.6
1980	5	-0.9	0.0	-0.2	-0.8	0.2	-1.0	-0.5	0.5	-0.9	-0.1
1980	10	-2.1	0.0	-1.2	-0.8	0.0	-3.1	-0.5	-0.5	-1.6	-0.6
1980	15	-5.3	0.0	-4.3	-0.8	-0.2	-5.7	-0.5	-2.0	-2.4	-0.8
1980	20	-4.9	0.0	-4.6	-0.3	0.1	-7.4	-0.5	-2.7	-3.4	-0.8
1980	25	-3.4	0.0	-3.4	-0.1	0.2	-8.8	-0.5	-3.1	-4.6	-0.7
1983	5	-5.2	-4.3	-0.2	-0.5	-0.2	-3.3	-3.2	1.1	-0.2	-1.0
1983	10	-10.4	-4.2	-2.6	-0.8	-2.8	-3.8	-3.3	1.8	-0.6	-1.7
1983	15	-13.6	-4.1	-1.8	-0.6	-7.1	-4.0	-3.3	2.5	-1.3	-1.9
1983	20	-13.3	-4.0	0.1	-0.7	-8.6	-3.9	-3.4	3.6	-2.2	-1.9
1984	5	-2.1	-0.6	-1.4	-0.1	-0.1	-1.8	0.0	-1.6	0.4	-0.6
1984	10	-5.8	-0.6	-5.0	0.1	-0.2	-4.3	0.0	-4.1	0.7	-0.9
1984	15	-5.3	-0.6	-5.3	0.5	0.0	-6.3	0.0	-6.4	0.8	-0.7
1984	20	-3.5	-0.6	-3.8	0.8	0.1	-7.9	0.0	-8.1	0.8	-0.6
1988	5	-7.4	-6.3	-0.9	0.0	-0.2	-0.9	-0.6	0.2	-0.1	-0.4
1988	10	-3.5	-6.3	2.4	0.4	-0.1	-1.5	-0.6	-0.3	-0.2	-0.3
1988	15	0.2	-6.2	5.7	0.7	0.0	-2.0	-0.6	-0.5	-0.7	-0.3
1990	5	-6.3	-5.2	-1.0	0.1	-0.2	-0.2	-0.6	0.8	0.0	-0.4
1990	10	-3.3	-5.2	1.3	0.6	-0.1	-0.5	-0.6	0.6	-0.3	-0.3
1990	15	-0.3	-5.1	3.8	1.0	-0.1	-1.1	-0.6	0.4	-0.6	-0.3
1994	5	0.2	0.0	0.2	0.0	0.0	0.4	0.0	0.3	0.0	0.1
1994	10	1.2	0.0	1.4	-0.2	0.0	1.4	0.0	1.3	-0.1	0.2

Appendix Table. Error per thousand contributed to world error by four countries, and amount of error due to different components

Fore-cast	*Projec-tion length*	*Total*	*Base popu-lation*	*Births*	*Deaths*	*Net mig-rants*	*Total*	*Base popu-lation*	*Births*	*Deaths*	*Net mig-rants*
				Nigeria					**U.S.A.**		
1973	5	1.5	1.5	0.0	-0.1	0.0	-1.5	-1.3	0.1	0.0	-0.3
1973	10	1.6	1.6	0.2	-0.2	0.0	-1.4	-1.3	0.7	-0.3	-0.5
1973	15	1.8	1.7	0.3	-0.3	0.0	-1.3	-1.2	1.2	-0.6	-0.7
1973	20	2.2	1.8	0.7	-0.3	0.0	-1.4	-1.2	1.5	-0.8	-0.9
1973	25	2.9	2.0	1.2	-0.3	0.0	-1.9	-1.2	1.5	-1.0	-1.2
1973	30	3.9	2.1	1.9	-0.2	0.1	-2.3	-1.1	1.7	-1.3	-1.5
1980	5	2.6	2.3	0.2	0.1	0.0	-1.6	-1.6	0.3	-0.2	-0.2
1980	10	3.2	2.4	0.4	0.3	0.0	-1.5	-1.5	0.7	-0.3	-0.4
1980	15	4.0	2.6	0.8	0.6	0.0	-1.6	-1.5	0.9	-0.4	-0.6
1980	20	5.1	2.8	1.3	0.9	0.0	-2.1	-1.4	0.9	-0.6	-1.0
1980	25	6.4	3.0	1.8	1.4	0.1	-2.4	-1.4	1.0	-0.8	-1.2
1983	5	5.1	4.7	0.2	0.2	0.0	-0.4	0.2	0.3	-0.1	-0.7
1983	10	6.1	5.0	0.7	0.4	0.0	-1.1	0.1	0.3	-0.1	-1.4
1983	15	7.8	5.5	1.6	0.7	0.0	-1.9	0.1	0.1	-0.1	-2.1
1983	20	9.6	5.9	2.5	1.1	0.1	-2.5	0.1	0.2	-0.3	-2.6
1984	5	3.9	3.5	0.3	0.1	0.0	-0.8	-0.6	0.1	-0.1	-0.2
1984	10	4.8	3.8	0.9	0.2	0.0	-1.0	-0.5	0.0	-0.2	-0.3
1984	15	6.2	4.1	1.8	0.3	0.0	-1.4	-0.5	-0.1	-0.2	-0.6
1984	20	8.1	4.5	3.0	0.5	0.1	-1.6	-0.5	0.0	-0.3	-0.8
1988	5	6.1	5.4	0.5	0.1	0.0	-1.0	-0.5	-0.2	-0.1	-0.3
1988	10	7.4	5.8	1.1	0.4	0.0	-1.8	-0.5	-0.5	-0.1	-0.7
1988	15	8.8	6.3	1.7	0.8	0.0	-2.6	-0.5	-0.6	-0.2	-1.2
1990	5	5.7	5.3	0.4	0.1	0.0	-0.7	-0.5	-0.1	0.0	-0.1
1990	10	6.7	5.6	0.7	0.3	0.0	-1.2	-0.5	-0.4	0.1	-0.3
1990	15	7.8	6.1	1.1	0.6	0.0	-1.5	-0.5	-0.6	0.1	-0.5
1994	5	2.3	1.9	0.4	0.0	0.0	-0.7	-0.8	0.1	0.0	0.0
1994	10	2.9	2.0	0.8	0.1	0.0	-0.5	-0.8	0.2	0.0	0.0

Figure 1. Ratio of U (unsigned) error to T (total trend) error and S (summed) error, by component and projection length

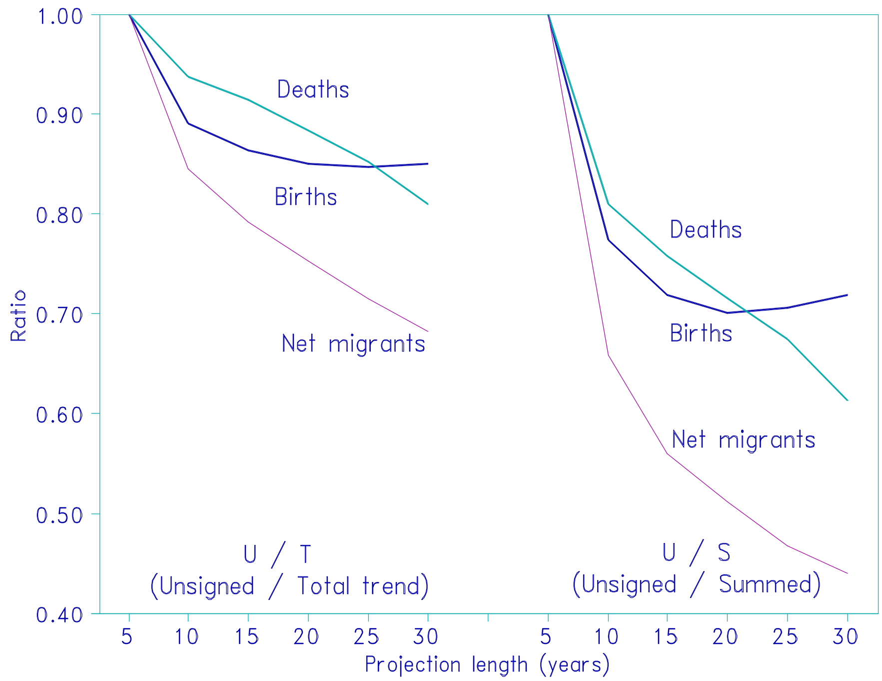
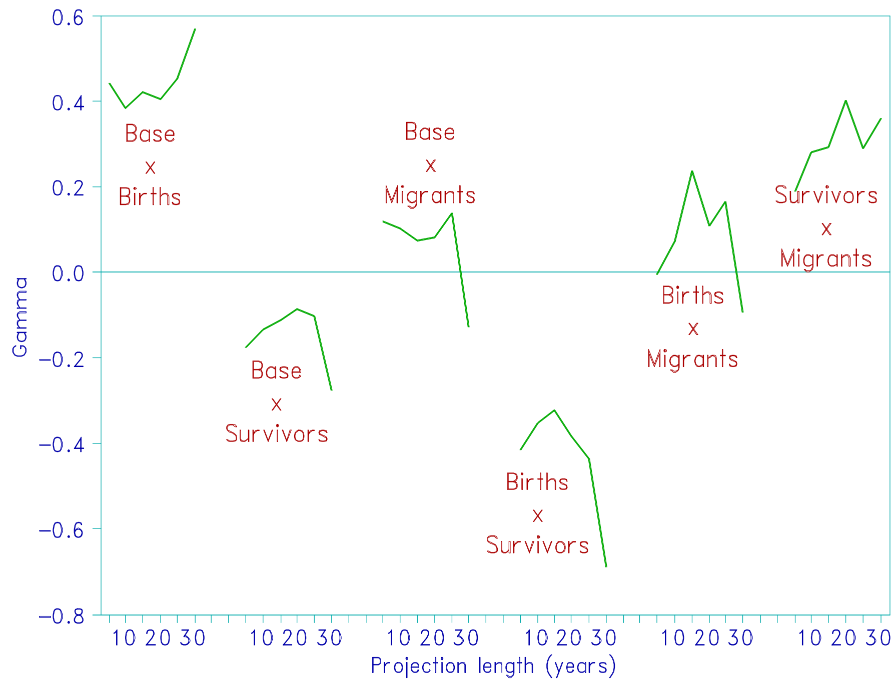


Figure 2. Association between error in projection components, by projection length



Note: Gamma coefficients are for crosstabulations similar to the subtables in Table 4.

Figure 3. Ratio of V (visible) error to U (unsigned) error, by component and projection length

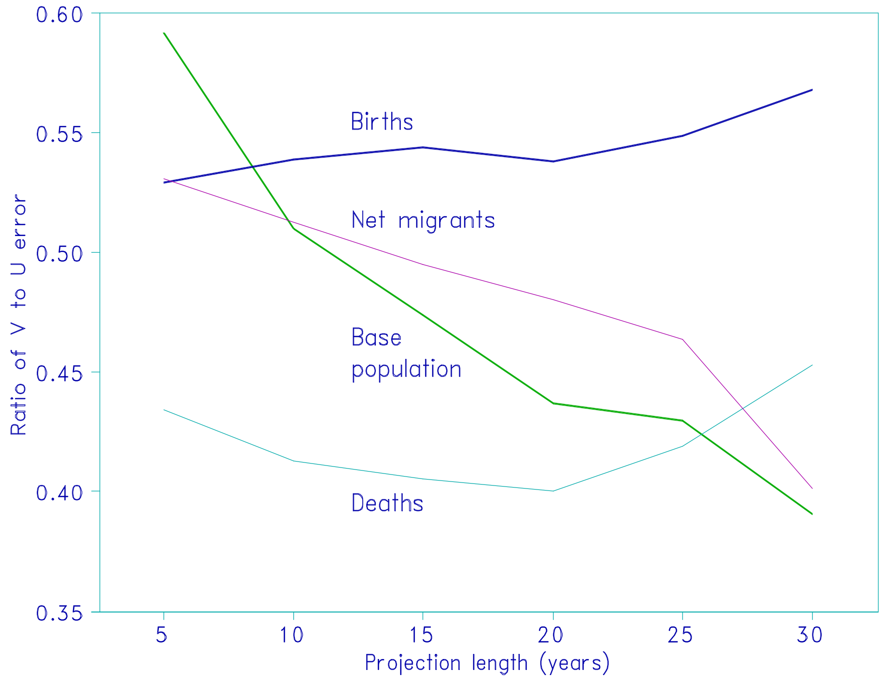
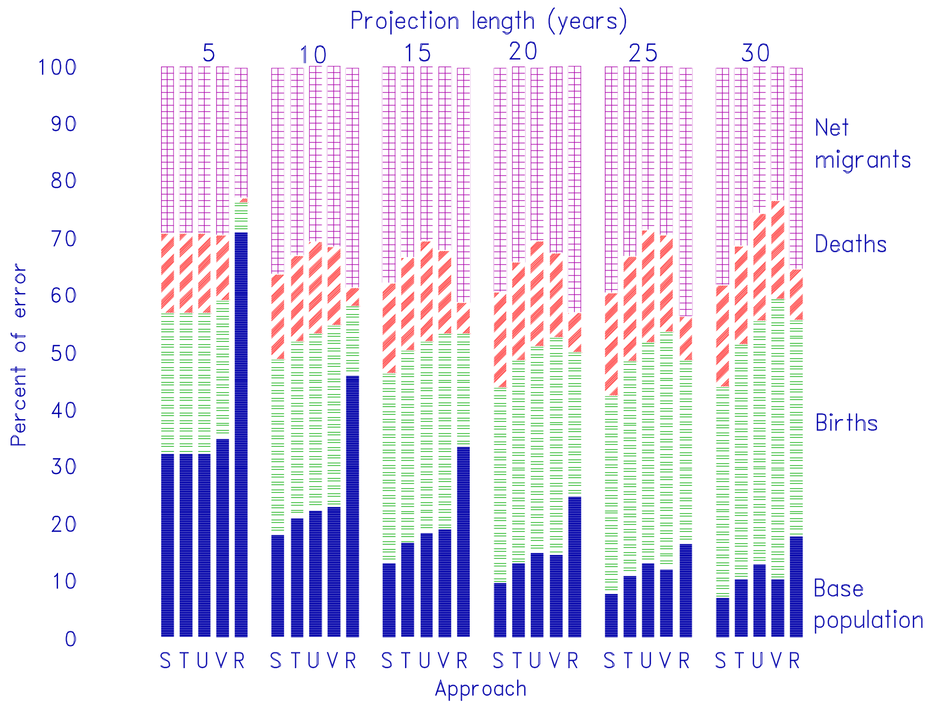
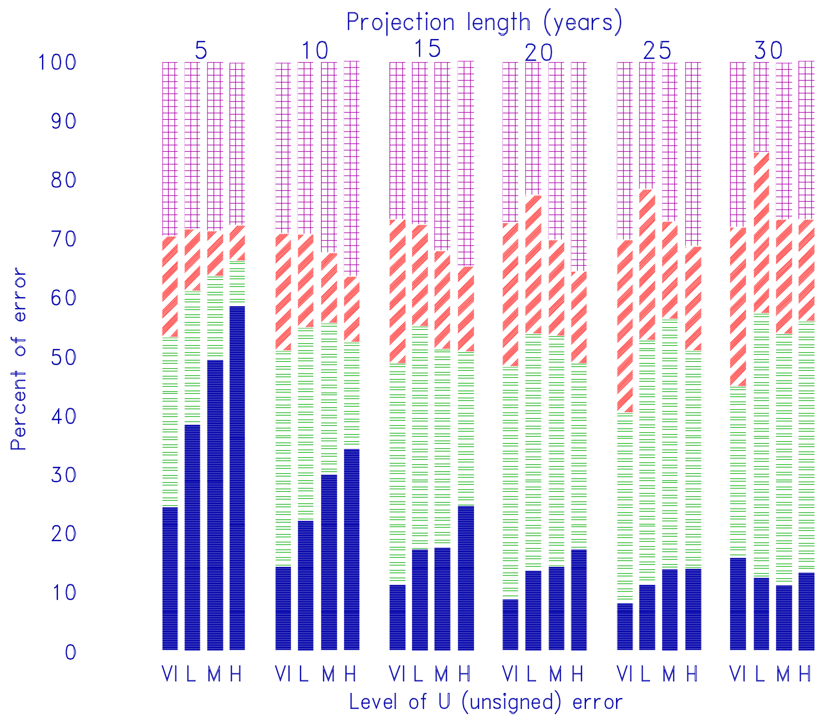


Figure 4. Attribution of projection error to different components, by projection length, using alternative approaches



Note: S (summed), T (total trend), U (unsigned), and V (visible) error are defined in the text. R is the partitioning of variance (R^2) from the National Research Council (2000), with interactions and unexplained variance left out.

Figure 5. Decomposition of U (unsigned) projection error, by projection length and level of U error



Note: Vl (very low) error is .000 to .049, L (low) .050 to .099, M (medium) .100 to .149, and H (high) .150 and up.

Figure 6. U (unsigned) error attributed to different components: Means for countries by initial population size, initial total fertility, and demographic history

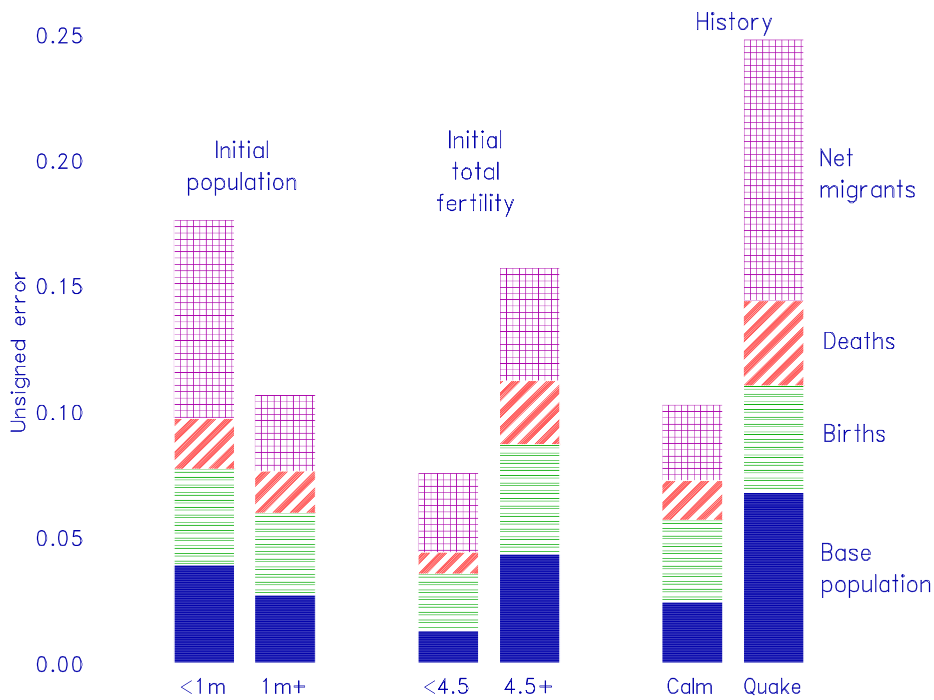


Figure 7. U (unsigned) error attributed to different components: Means by world region

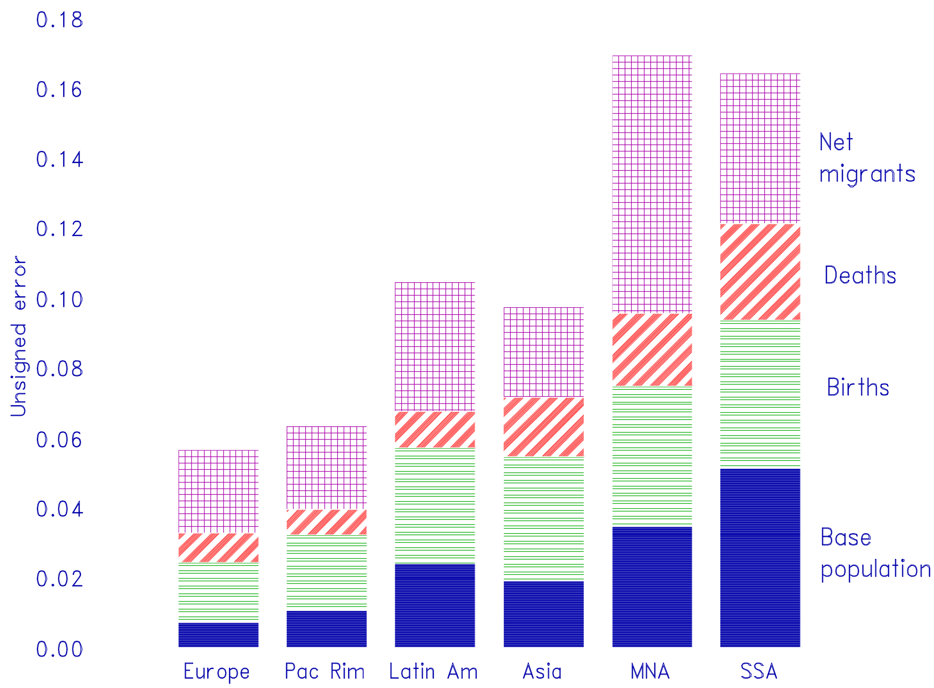


Figure 8. Alternative decomposition of projection error into starting-level and slope error, by projection length

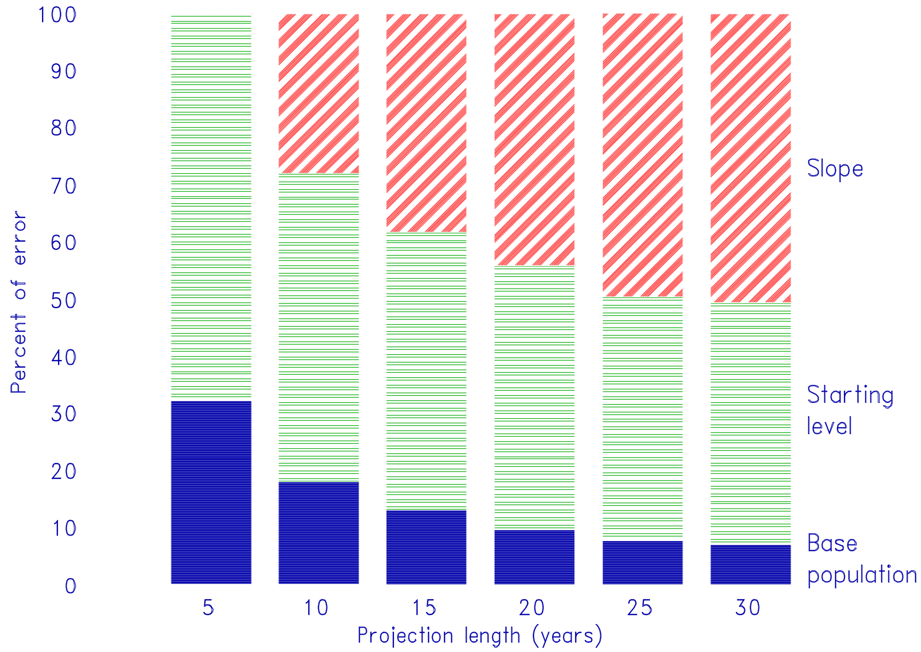


Figure 9. Positive and negative error, and resulting visible error, in world projections, by forecast and projection length

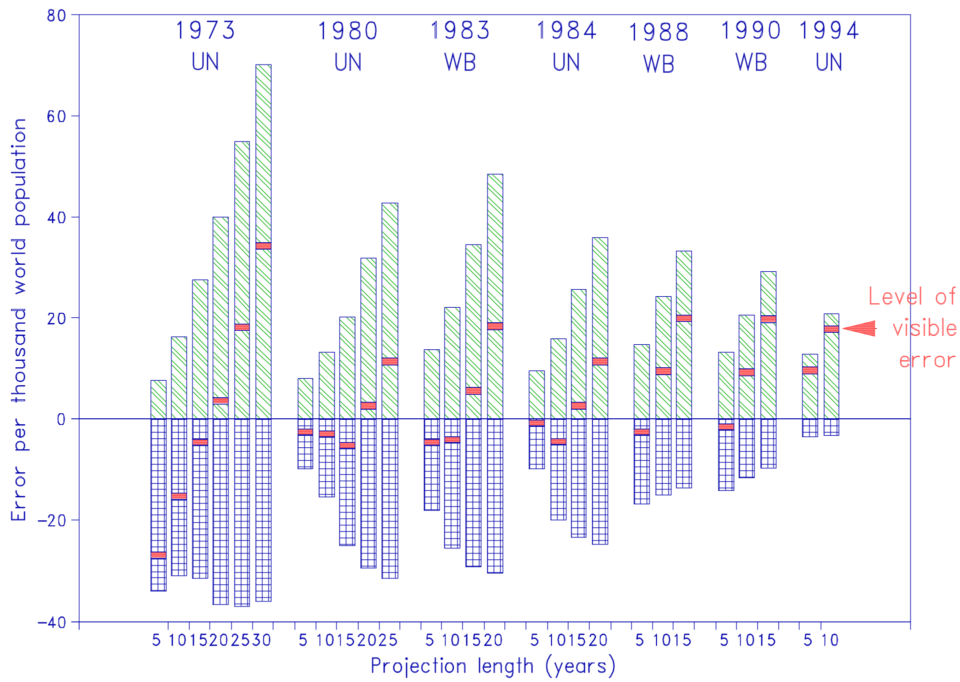


Figure 10. Contributions to world projection error from the five countries contributing the most positive and negative error and from all other countries, by source of error: Means across forecasts

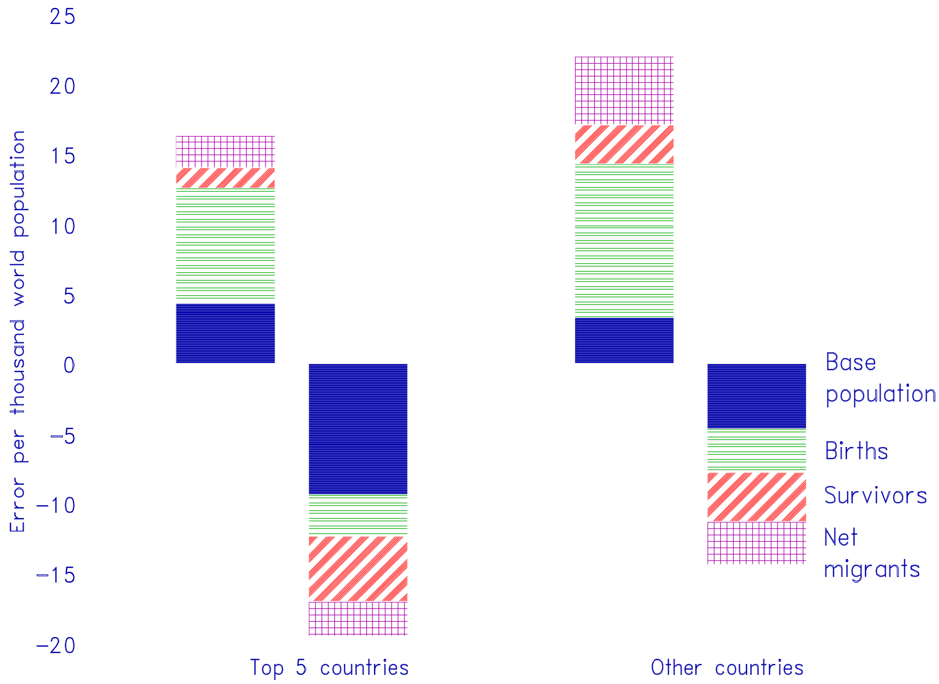


Figure 11. Contributions to world projection error from countries grouped by population size, by source of error: Means across forecasts

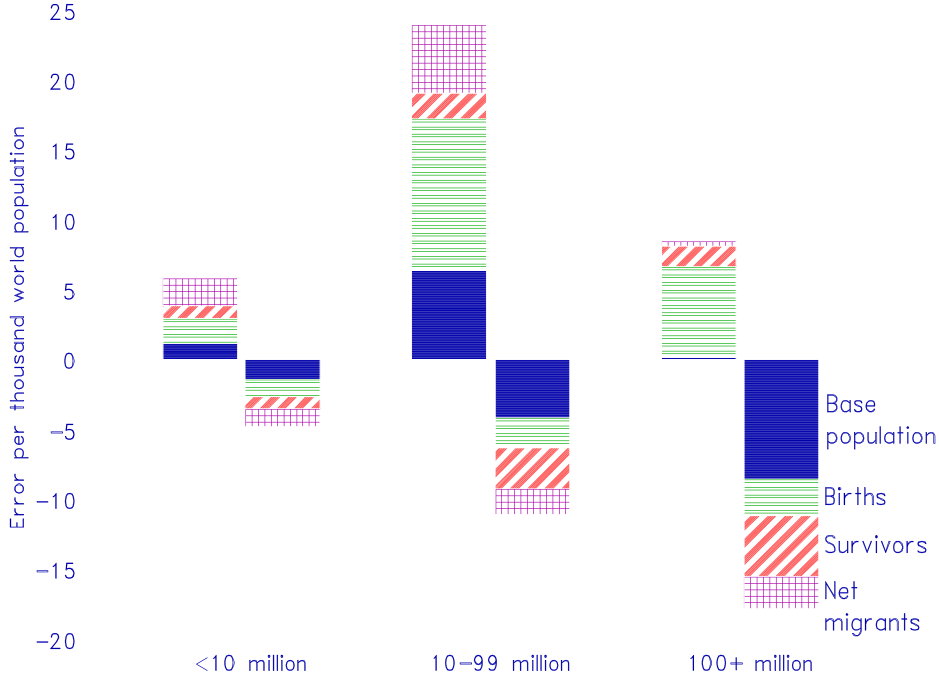


Figure 12. Contributions to world projection error from six regions, by source of error: Means across forecasts