"Global population projections by the United Nations"
John Wilmoth, Joint Statistical Meetings, Seattle, 10 August 2015

[Slide 1] Let me begin by thanking the Southhampton social statistics group for proposing an invited session on demographic forecasting during the 2015 Joint Statistical Meetings, and for inviting me to take part and to present the latest global population projections by the United Nations, which, as you may know, were released very recently at the end of July.

In the time that I have, I will briefly describe some key aspects of the global population projections produced by the United Nations. The UN's work in this area has changed and, I think, improved over time. I hope that by the end of my presentation you will understand better the background and the meaning of the UN's projections, as well as some of the key findings from this latest round.

[Slide 2] I will start with a bit of history. The UN has produced 24 sets of global population projections, beginning in 1951. The early projections were for the world as a whole and for large regions. Projections for individual countries began in 1968. The latest set, the 2015 Revision, includes projections from 2015 until 2100 for 201 countries or areas.

[Slide 3] My main purpose today is to talk about some methodological issues that we face in producing the UN's population projections. However, since the new data were just released, it seems appropriate also to present some of the key findings that emerge from our most recent analysis of global population trends. First, the world's population continues to grow, and even though the annual rate of growth has been falling for the past four decades, we expect the world's population to keep growing throughout the 21st century. Currently, we are adding more than 80 million inhabitants to the planet every year, with the result that we expect today's population of around 7.3 billion to increase to 8.5 billion in 2030, then 9.7 billion in 2050 and onward to 11.2 billion in 2100.

Second, most of this growth is now occurring in Africa, and that will continue to be the case over the next several decades. In fact, between now and 2050, of the total projected increment of 2.4 billion for the world as a whole, more than half, around 1.3 billion, will be added to the population of Africa. Asia will account for another large share of the total, followed by other regions that are growing much more slowly, and one region, Europe, whose population is projected to shrink at least slightly over this time period.

Third, we expect some noteworthy shifts in the relative size of the population giants in future years, in particular for China and India. It has been predicted for a long time that the population of India would eventually overtake that of China, but in our latest assessment of the evidence, we have moved forward the date at which the crossover is expected to occur. We now expect that India will become the largest country in the world starting in 2022. Similarly, Nigeria, which now ranks 7th in terms of total population, is expected to continue growing very rapidly. In fact, it is projected to overtake the United States and become the 3rd largest country in the world by 2050.

Fourth, I would like to highlight the considerable demographic diversity that exists in the world today. On one side, there are a large number of countries, including many of the least developed countries, many of them located in sub-Saharan Africa, where high levels of fertility continue to drive rapid population growth, and that growth is associated with a very young age structure. The challenges are those of providing adequate health care to an ever-expanding population of children and young people, and of empowering youth through education and opportunities for employment that will give them a sense of hope for the future. On the other side, there are a large number of countries, including many in Europe, North America and East Asia, where historically low levels of fertility have nearly brought an end to population growth, leading to slower growth and decline in more extreme cases. Even if populations do not decline but merely stabilize their growth by bringing the birth rate into alignment with the reduced death rate, they face the prospect of a substantial ageing of the population, with attendant concerns about the fiscal challenges this will present to the working-age population as it seeks to support programmes to finance old-age pensions for the elderly as well as medical and personal care. In these situations, I like to remind people that demography is not destiny. It is, however, a powerful force that must be acknowledged and understood, so that governments can make the appropriate policy response.

Lastly, I will mention a piece of good news from these latest projections, which is that during the period from around 2000 until 2015 there has been an unusually rapid fall in levels of child mortality for several countries, including some of the poorest and least developed. Moreover, the increase of life expectancy at birth has been especially rapid in the least developed countries, which saw an average increase of life expectancy at birth of around 6 years over a 10-year period from the early 2000s to the early 2010s. This compares to an increase in life expectancy at birth of only 3 years for the world as a whole. Thus, the gap in this critical measure of human well-being, while still large, appears to have shrunk in size.

[Slide 4] Now, with that quick summary in mind, let's turn to the issue of how we make these calculations. This next slide offers a quick overview of the method used by the United Nations to project the populations of countries and the world. All calculations are made within a cohort-component framework. When projecting the population of an individual country, assumptions about future trends in fertility and mortality are derived mostly from information about historical trends for the country itself. Nevertheless, the model of demographic change for an individual country is a specific case of a more general global model, and the available data for a single country are often supplemented by information from other countries, as necessary.

Let me explain quickly that I will focus here on how we model future trends in fertility and mortality. For international migration, the UN assumes, quite simply, that current levels will be maintained for several decades, and then gradually decline toward zero in the distant future, even beyond the projection horizon of 2100. This assumption is the topic of the next presentation by Adrian Raftery. For the rest of my talk, I will focus on fertility and mortality as the main components of the UN's projections.

Traditionally, the UN's medium projection has been accompanied by a small number of alternative future trends, derived by varying key assumptions, in particular fertility, or by constructing hypothetical scenarios, such as an instant shift to replacement fertility or to zero migration. More recently, a new method based on a formal statistical model has made it possible to assess the probability of observing a range of plausible future trends.

The median trajectory of this probabilistic model is derived using the same assumptions about future trends in fertility, mortality and migration as the traditional medium-variant projection. Thus, the difference between the two approaches involves only the way of describing alternative future trends, not the central trend itself.

[Slide 5] Before we go deeper into the details of the UN's projection model, let us be clear about two important pieces of terminology. Levels and trends in human fertility and mortality can be described using a variety of measures. Two of the most common measures are essential components of the projection model used by the United Nations:

- The total fertility rate, or TFR, is expressed as the number of births per woman over her lifetime. It equals the sum of age-specific fertility rates for a given time period and thus represents the average number of births that women would bear if the intensity of childbearing throughout their lives matched the current age-specific rates.
- Life expectancy at birth, or e_0, is expressed as the average duration of life in a population. It equals the mean of a hypothetical distribution of ages at death, if a cohort of individuals were to experience throughout life the age-specific death rates of the period in question.

[Slide 6] Projected trends in fertility and mortality are derived from historical information about trends in those same variables. As you may know, observed regularities in historical trends have led to various theories of demographic change. In building a projection model, the UN used existing theories of demographic change to guide the specification of a general model, which was then fitted to the experience of individual countries in the framework of the Bayesian hierarchical model (BHM).

A major advantage of the BHM in this context is that the model for a given country can be derived using a combination of data for the country itself and for other countries as well. When data are plentiful for a given country, the estimation procedure gives most of the weight to data for that country when choosing the parameters of the country-specific model. When data are sparse, the method gives more weight to data for other countries.

[Slide 7] Although the BHM provides a useful framework that allows us to consider jointly the data for an individual country and for its neighbors and the world, the exact specification of the model reflects also some essential elements of the existing theoretical understanding of temporal changes in levels and patterns of fertility and mortality during the modern era. Despite differences in the details, most existing theories about long-term trends in fertility and mortality share certain common features, which are reflected in the UN's projection model.

For example, in the case of fertility, the model depicts a transition from high to low levels of the total fertility rate, while for mortality the model depicts a transition from low to high levels of life expectancy at birth. As the famous demographer, Paul Demeny, wrote in 1968, "In traditional societies fertility and mortality are high. In modern societies fertility and mortality are low. In between, there is a demographic transition."[1] This is the most basic and uncontroversial aspect of demographic transition theory: that a transition *has* occurred, *is* occurring or *will* occur in all human populations. The models of fertility and mortality change that underlie the UN's projection method give form to a general theory of transition, using a double-logistic curve to describe rates of change in these two variables as a function of their current level.

There is less theoretical consensus about what happens after these major transitions have occurred. Will life expectancy continue to increase? Will fertility

---

[1] Demeny, P. (1968). "Early fertility decline in Austria-Hungary: A lesson in demographic transition." *Daedalus*, pp. 502-522.

remain well below replacement level for a long time? In each case, the UN projections reflect a view of the future that is widespread but probably not universal in the same way that there is an almost universal recognition of transition theory. For fertility, the UN assumes that post-transition trends will fluctuate and that countries with very low levels of fertility will see at least a modest recovery toward higher levels — not necessarily to the replacement level of 2.1 children per woman, but still in that direction. For mortality, we assume that the future increase of life expectancy at birth will be roughly linear, with a rate of change that converges toward the observed increase of the maximum age at death.

[Slide 8] The next slide illustrates the complete fertility model in three phases, corresponding to the periods before, during and after the transition. Note that the rapid drop of the total fertility rate during the transition is followed by a period of fluctuation and modest increase.

[Slide 9] The next slide illustrates the shape of the trend in life expectancy at birth that is embedded in the model. It is a kind of S-curve, with a future increase that decelerates and eventually becomes linear, with a slope that must be chosen based on a reasoned argument. In this case, it was argued that the rate of increase in life expectancy at birth is likely to decline in the future, since in the long run the rate of increase in the mean of a distribution cannot exceed the rate of increase in its maximum, which has been moving up more slowly.

[Slide 10] The next three slides focus on phase II of the fertility model, or the period of major decline. The model of decline depicts the rate of change in the total fertility rate as a function of the TFR itself. The general form of this relationship is an inverted U, which is modeled using a double-logistic curve. The inverted U-shape implies that the transition begins slowly, proceeds more quickly in the middle, and then slows down at the end. We use the Bayesian hierarchical model to estimate the parameters of the double-logistic model for each country, yielding information about both the country-specific parameter estimates and their uncertainty.

[Slide 11] Here is an illustration of the fitted model for India. In the left panel, the solid red line depicts the estimated curve, depicting the rate of change in the total fertility rate as a function of its current level. There is uncertainty around that estimate, as depicted both by the individual simulated curves (in grey) that approximate the posterior distribution, and by the percentile bands (in dashed and dotted red) corresponding to four percentiles of the posterior distribution, at 2.5, 10, 90 and 97.5, which form 80- and 95-percent prediction intervals for the estimated double-logistic model of fertility decline for India.

In the panel on the right, rates of change from the panel on the left are transformed into projected future changes in the total fertility rate. Again, the uncertainty bands (in red) correspond to 80- and 95-percent prediction intervals. The individual trajectories of the future TFR (in grey) were generated by MCMC sampling, taking into account both the uncertainty related to model estimation (on the left) and the uncertainty of future annual changes, as described by a random walk. In short, this is a model of a random walk with variable drift, which takes into account the uncertainty about the drift parameter in assessing the uncertainty of future trends.

[Slide 12] Briefly, this slide illustrates how the Bayesian hierarchical model yields estimates of the rate of decline in the total fertility rate for countries with different historical patterns, and also for countries that differ greatly in terms of data availability and their current stage of transition.

[Slide 13] Again briefly, this slide illustrates the third phase of the fertility model, which starts by definition after there has been an increase over two consecutive 5-year time intervals for a country with a TFR that is already below 2. The empirical base used for estimating this portion of the fertility model now includes 38 countries or areas, mostly in Europe but also, increasingly, in Asia, North America and elsewhere.

[Slide 14] It goes without saying that there is considerable uncertainty about future population trends. A major challenge for demographers has been how to describe and communicate that uncertainty to everyday consumers of population projections. Although the traditional variants and scenarios have no probabilistic interpretation, by documenting the sensitivity of results to changes in assumptions, they remind users of the possibility that future trends may differ from the central forecast. Thus they emphasize the *existence* of uncertainty, even if they do not help very much in understanding the *extent* or the *magnitude* of that uncertainty. For a quantitative assessment of the uncertainty of population projections, one needs a statistical model.

[Slide 15] It is instructive to compare the two means of depicting alternative future population trends in some specific cases. Here, for Nigeria, the medium variant is shown as a solid red line. The other red lines depict the distribution of future trends according to the probabilistic model, while the blue lines depict the high- and low-fertility assumptions, which are exactly half a child higher or lower than the medium variant.

In the case of a high-fertility country like Nigeria, the high- and low-fertility assumptions are not particularly extreme and do not cover the full range of uncertainty revealed by the probabilistic analysis. In short, if our notion of a "high" and "low" scenario derives from a total fertility rate that varies by plus or minus half a child, we will understate the uncertainty of future trends in this context, as seen on the left for the TFR and on the right for total population.

[Slide 16] The situation is just the opposite for low-fertility countries, such as the Russian Federation. Although future trends are expected to fluctuate, those changes are likely to take place within a fairly small range, which is easily covered by plus or minus half a child. In this case, defining a "high" and "low" scenario by this traditional method leads to a modest overestimation of the future uncertainty of population trends.

[Slide 17] For large regions or the world as a whole, the high- and low-fertility variants grossly overstate the uncertainty of future trends. This is because the high- and low-fertility assumptions are applied uniformly across *all* countries and *all* future time periods. The uniform application of this assumption yields an exaggerated depiction of the range of plausible future trends. Although some countries will have higher levels of fertility than implied by the medium scenario, others will have lower levels. Likewise, some countries will have fertility levels that are higher than expected in some time periods but lower in others. The impact of these differences will tend to cancel out.

For the UN's projection of world population in 2100, the high- and low-fertility variants differ by around 9 billion, spanning a range from roughly 7½ to 16½ billion. By contrast, the 95-percent prediction interval for the global population in 2100 based on the probabilistic analysis extends from around 9½ to around 13½ billion, covering a range of around 4 billion.

[Slide 18] What have we learned from the new probabilistic projections? We have learned, as mentioned already, that the traditional high- and low-fertility variants tend to exaggerate the uncertainty of future trends in some cases while understating the uncertainty in others. We have also learned that stabilization of the world's population within this century seems somewhat unlikely. However, stabilization in that time frame is not impossible or even implausible according to this analysis: we have estimated that there is a 23-percent chance that the world population will either level off or begin to decline by the end of the 21st century.

[Slide 19] The Population Division acknowledges that the assessment of the uncertainty of our global population projections using the probabilistic approach is still a work in progress. The current method does not incorporate several forms of uncertainty as listed here, which may be of major or minor importance compared to the uncertainty already reflected in the model.

I will briefly mention just two items on this list. First, the current method does not incorporate uncertainty related to future trends in international migration. This omission may lead to a substantial understatement of the uncertainty of future population trends for a country like Germany, where international migration is now the major driver of population change. Second, the current method does not take account of uncertainty related to our knowledge of the past and present demographic situation of a country.

[Slide 20] This graph, for example, shows available data on levels of the total fertility rate in Nigeria since 1970. Obviously, the various sources of information do not provide a clear and consistent signal. Between the 2010 and the 2012 Revisions, the UN revised upward its estimates of the total fertility rate for Nigeria in the latest time periods, since newly available data suggested a higher level. However, this upward revision was soon contradicted by the next available data point, which became known within a year of publishing the 2012 Revision and suggested a TFR value clearly below the revised estimate.

[Slide 21] Therefore, in the latest revision, the current fertility estimate for Nigeria has been revised downward by roughly one quarter of a child. As a result, the 2015 estimate lies in between the estimates from the two previous revisions.

[Slide 22] The example of Nigeria illustrates the impact of current fertility estimates on the projected future trend. This slide shows the fertility projections of the 2012 and 2015 Revisions in blue and red, respectively, including the median values as well as the 80- and 95-percent prediction intervals. We see that the downward revision of around one quarter of a child in the current fertility estimate is maintained in the projected trend of fertility over the next three decades, after which the difference diminishes and effectively disappears.

[Slide 23] The immediate impact of the reduced fertility level during the first decades of the projection period is hardly noticeable in this graph of the projected population size. However, the difference in population size grows over future decades, so that the cumulative impact of the revised fertility estimate is quite substantial by the end of the century: Nigeria’s projected population of 914 million for 2100 in the 2012 Revision was lowered to 752 million in the 2015 Revision, which is a downward adjustment of about 18%.

[Slide 24] To summarize my key points:

- Population projections by the United Nations are derived from models of future trends in the components of change, in particular fertility and mortality.
- The UN's method of population projection is informed by demographic theories of the change over time in fertility and mortality. For each country, the models are calibrated using data from the country itself and, especially when data are sparse, from other countries as well.
- The uncertainty of future population trends has traditionally been depicted using variants and scenarios. In the two latest revisions, we have also employed a new method based on the Bayesian hierarchical model that yields probabilistic statements about plausible future trends.
- Lastly, work on the probabilistic assessment of uncertainty by the UN is ongoing and could benefit from further efforts to incorporate additional sources of uncertainty, including the uncertainty in future projections that derives from uncertainty in current estimates of key demographic parameters such as fertility.

[Slide 25] I would like to acknowledge the efforts of our fine colleagues in the United Nations Population Division, who do the hard work of producing the estimates and projections of world population that so many of you use in your own work. I also want to thank Professor Adrian Raftery and his collaborators for an extremely fruitful and also very enjoyable collaboration over the past decade.

[Slide 26] Finally, I want to point out that there are many freely available R packages for implementing the UN's probabilistic projection method.

[Slide 27-30] And for those of you who would like more detailed information on the probabilistic projection method, the last four slides of my presentation contain further information about the suite of R programs, as well as an extensive bibliography.

You can obtain a copy of this presentation either from the website of the United Nations Population Division or by sending me an email.

Thank you for your attention.