

**Global population projections: A critical
comparison of key methods and assumptions**

John Wilmoth, Director, Population Division

African Population Conference, Entebbe, Uganda, 18 November 2019

- [Slide 1]** Let me begin by thanking Rachel Snow for proposing this special session on demographic forecasting during the eighth African Population Conference. I am grateful for the opportunity to take part in a discussion focusing not only on the global population projections by the United Nations but also on the projections produced by IIASA and the Wittgenstein Centre in Vienna. I will begin by describing briefly some key aspects of the methods and assumptions that underlie the UN projections. I will then focus on differences compared to the IIASA projections and the causes of those differences. After a critical assessment of the two methodologies, I will conclude with a brief discussion of the implications of different perspectives on future trends, and a proposal on how the two teams could work together to deepen our understanding.
- [Slide 2]** I start with a bit of history. The United Nations has produced 26 sets of global population projections, every two or three years since 1951. Early projections were for the world and large regions, while projections for individual countries began in 1968. The latest set, released earlier this year, includes information for 235 countries or areas.
- [Slide 3]** Allow me quickly to summarize the methods and assumptions that underlie the population projections of the United Nations. Three key elements are shown here, including: (1) the role of demographic transition theory, (2) functional forms of UN models of fertility and mortality trends, and (3) recent improvements thanks to re-casting the UN's traditional approach in the framework of a Bayesian hierarchical model. Let's take a closer look at these three points.
- [Slide 4]** At its core, the projection model used by the United Nations derives from theories of the demographic transition. Despite the lack of a single theory that enjoys universal acceptance as an explanation of historical demographic change, the UN projection model is based on the core premise that underlies all such theories: that transitions from high to low levels of mortality and fertility have occurred, are occurring or will occur in all populations of the world.

[Slide 5] There are three phases in the model of fertility change used to derive assumptions about future trends in total fertility. Phase I is the period before the transition from high to low levels of fertility. Phase II is the decline itself, and Phase III is the period after the decline.

All countries have begun their fertility transitions and exited Phase I. Therefore, the fertility model used by the UN covers only Phase II and Phase III. I will return later to the post-decline model of Phase III. This slide focuses attention on the model for Phase II.

On the left side, we see a curve in the shape of an inverted U – a double logistic curve. This functional form has been used by United Nations demographers for many years to depict the speed of decline in fertility as a function of its current level. On the right side, we see the trend in the total fertility rate implied by the curve on the left. A key feature of the inverted U shape is that it produces a typical transition pattern, in which change is slow at first, then more rapid in the middle, and then slow again at the end.

[Slide 6] The same functional form, the double logistic model, is used also to depict the rate of increase in life expectancy at birth. Like the fertility model, the model of increasing life expectancy is characterized by slower change at first, then more rapid, and then slower changes at the end. The red line on the left corresponds to the assumed long-run rate of increase in life expectancy for populations at very low levels of mortality, which was derived from an analysis of trends in record longevity.

The United Nations uses these models to depict plausible future changes in total fertility and life expectancy for all countries. Both models are estimated using a large body of historical data and are used to derive assumptions about future trends in mortality and fertility for all countries. These assumptions are combined with simple assumptions about international migration to generate population projections for all countries and areas.

[Slide 7] In the past, these projections were derived in a deterministic manner. In recent years, however, the United Nations and its collaborators have developed a technique of probabilistic projection by estimating the double logistic curves of the fertility and mortality models in the framework of a Bayesian hierarchical model, which provides a means of sharing experiences

between countries with similar demographic conditions. The Bayesian model also generates probabilistic prediction intervals that depict the likelihood of alternative future trajectories.

[Slide 8] Before the UN developed its probabilistic projection model, it followed the common practice of generating high and low scenarios. Specifically, these were high- and low-fertility variants, computed in the same manner as the medium variant with only a small change in the fertility assumption: the TFR was assumed to be higher or lower by half of a live birth for every country or area and for every year in the future.

This graph compares the UN's traditional high- and low-fertility variants to the prediction intervals of the probabilistic approach. We see that these high and low scenarios are quite extreme at the global level. The explanation is simple: while it may be plausible to assume a TFR that is higher or lower by half of a live birth for a given country in a given year, the assumption that all countries in all years will have a fertility level that is higher or lower by this amount is not so plausible. Some countries may experience fertility levels above the medium trend, while others may be below. Some countries may be higher in some years and lower in others. In the real world, these sorts of differences tend to cancel out.

Therefore, at the global level, the traditional high and low scenarios of the UN tend to vastly overstate the range of plausible future population trajectories, because they assume a perfect correlation of deviations from the medium variant across countries and over time. This "aggregation fallacy" of deterministic projection scenarios is one of the most important insights to emerge from the probabilistic analysis of the UN population projections.

[Slide 9] Now let us turn to a comparison of the two sets of projections. I begin by examining the difference in the medium projections. How large is the difference, and what is the cause?

[Slide 10] Earlier I showed you the UN medium-variant projection with prediction intervals from the probabilistic model. Where does the IIASA medium projection lie in comparison?

[Slide 11] Here is the answer: the IIASA medium projection lies close to, or slightly below, the lower limit of the 95-percent prediction interval. Thus, a first conclusion is that one cannot dismiss the IIASA medium trajectory as implausible, based on the UN's probabilistic model.

Nevertheless, the two medium projections offer rather different visions of the most likely future trend of global population. In each case, the medium trajectory is accompanied by alternative scenarios that surround the medium, purporting to offer a range of more-or-less plausible trajectories. We will come back to this point later.

For now, let us keep our focus on the two medium projections and try to understand the difference between them.

[Slide 12] Three explanations are worth considering, as shown here. All three have been mentioned in earlier discussions or writings on this topic. However, upon closer examination, it seems that these explanations are not equally plausible or relevant.

[Slide 13] We begin with the assertion by Lutz and colleagues in 2018 that the difference between the two sets of projections is due to different methods of deriving long-term fertility trends, one based primarily on statistical extrapolation and the other giving more weight to scientific reasoning.

[Slide 14] It is true that the projection method used by the United Nations includes statistical models that extrapolate historical trends of fertility and mortality into the future. But was the development of these models lacking in expert argumentation and scientific reasoning?

As noted earlier, the functional form of these models was motivated by theories of the demographic transition. In addition, I will mention three other examples of how both the specification and estimation of the statistical models used by the United Nations have benefitted from quite serious expert argumentation and scientific reasoning.

The three examples concern: (1) the treatment of observations that pre-date modern contraception, (2) the means of constraining the long-run rate of increase in life expectancy at birth based on trends in record longevity, and (3) the justification for the UN's post-decline fertility model.

Since there is not enough time in today's presentation to elaborate on all three examples, I will skip the first two and move immediately to the third example.

[Slide 15] As mentioned earlier, the UN fertility model has two parts, corresponding to the historical decline from high to low levels (Phase II), and the period following that decline (Phase III). A country is considered to have entered the post-decline period only after an upturn is observed in the TFR over two consecutive 5-year periods.

During the post-decline period, it is assumed that a country's fertility level will move toward a country-specific average value and then fluctuate around that value. The post-decline model does not impose a specific value for fertility levels in the long run. Rather these emerge only after fitting the model to the available data for all countries and areas. The typical result of this exercise is a trend for each low-fertility country that rises somewhat above its current level but usually remains below the replacement level of 2.1 live births per woman.

[Slide 16] By 2019, the post-decline period had been observed in the fertility trends of 40 countries or areas. Moreover, this number has been increasing over time, rising from 21 countries or areas having entered Phase III at the time of the 2010 revision to almost twice that number today.

[Slide 17] Moreover, the plausibility of an upturn in fertility trends following a decline to levels well below replacement is supported also by recent research showing that at high levels of the Human Development Index (HDI), increasing development can reverse the declining trend in total fertility.

[Slide 18] Now, let us turn to another important matter: the oft-repeated claim that the difference in the two sets of projections is attributable to differences in attention to education as a predictor of fertility levels and trends.

[Slide 19] An example of this claim can be found in the 2018 volume presenting the most recent population projections by IIASA and its collaborators. On page 117, we find the dual claim that the UN projections are based on an extrapolative model that does not consider the changing educational structure, whereas the medium scenario from IIASA is based on the assumption that improvements in female education will continue, resulting in a more rapid fertility decline.

Ladies and gentlemen, it is important to be clear about this point: both of these claims are demonstrably false. Let me explain why.

[Slide 20] First, let us review another statement in the same 2018 volume, which describes the definition and substantive reasoning of the assumptions underlying the IIASA projections more accurately, pointing to “comprehensive reviews of the scientific literature” and “the largest ever expert survey” that was used to assess the validity of alternative assumptions.

Note that, in this explanation, there is no mention of education as a component of the model used to derive fertility trends for the medium scenario of the IIASA projections. And indeed, if one reads closely the methodology that underlies the medium projection, the long-term trends of both fertility and mortality were chosen not by building a model of the relationship between education and these variables, but rather by eliciting opinions from experts about the most likely future values of key demographic indicators for individual countries.

[Slide 21] I will return later to the use of expert opinion for setting long-range demographic assumptions. For now, I propose this concise summary of the role of educational change in the projection methodologies used by the UN and IIASA teams.

With both methodologies, there is no explicit model of educational change as a driver of fertility or mortality trends in the medium scenario. Instead, each one accounts implicitly for educational change, either through historical experience or through expert judgement.

For the UN projections, future trends are calibrated according to historical experience. That historical experience included substantial improvements in the educational level of populations undergoing the demographic transition. Just as educational change was a key driver of fertility decline in the past, so it will be in the future. Even though there is no explicit modeling of education, it is simply wrong to assert that the UN’s approach ignores the role of education as a driver of fertility change.

Likewise, it is incorrect to suggest that IIASA’s modeling strategy for the medium scenario includes education as an explicit driver of change. No doubt, educational change is present implicitly in the opinions elicited from experts, and it is reflected in the changing population composition by levels of education that is built into the IIASA projection model *after* the fertility trend of the medium scenario has been chosen based on expert opinion.

[Slide 22] This misunderstanding about the methods used for the two sets of projections is widespread, and it has permeated the popular discussion, as illustrated by this quote from a recent book claiming that advancing education would drive a more rapid decline in fertility than anticipated in the UN projections. In fact, almost every time I have heard a comparison of the UN and IIASA projections, this same false narrative has been used to explain the difference.

In short, ladies and gentlemen, I encourage you not to accept a baseless generalization about a lack of scientific reasoning by United Nations demographers, or false statements about the explicit driving force of education in the IIASA medium projection. Instead, let us try to understand what really underlies the difference between the two sets of projections.

In purely mechanical terms, the difference is due primarily to different assumptions about future rates of change in the total fertility rate. This difference in assumptions is especially important for countries of sub-Saharan Africa, where most of the world's future population growth will take place.

[Slide 23] Here is an illustration for a single country, Nigeria. On the left, we see both the underlying data points and the estimated values of the double logistic model, for the medium scenario and the upper and lower bounds of the 80-percent prediction intervals. On the right, we see the projected trend in total fertility implied by the double logistic model.

[Slide 24] And here, we overlay the medium projection from IIASA. This graph shows clearly that the fertility trend for Nigeria assumed for the IIASA medium projection is declining more rapidly than the assumed trend of the UN medium projection. Thus, it appears that the experts consulted in formulating the IIASA projections believed that there will be a marked acceleration in the pace of Nigerian fertility decline compared to what has occurred historically.

[Slide 25] These graphs, furthermore, demonstrate that there are similar differences in assumed rates of fertility decline for most countries of Africa. The differences are both systematic and persistent over time.

In addition to different assumptions for high-fertility countries in Africa and elsewhere, the IIASA medium projection does not allow for an upturn of fertility in low-fertility countries, as illustrated earlier for the UN medium projection.

[Slide 26] Let us now look critically at the two methodologies, considering both the accuracy of results and the plausibility of assumptions.

[Slide 27] How accurate have been the earlier rounds of UN population projections? A report by the U.S. National Academy of Sciences, entitled *Beyond Six Billion* and published in 2000, looked closely at this question and concluded that errors in the UN's global projections have consistently been under 4 percent. Absolute errors in projected national populations averaged 4.8 percent for 5-year projections but 17 percent for 30-year projections.

In this slide, we look back at the last 19 revisions. This simple analysis reveals a similar magnitude of error in projecting the size of the world's population in 2015.

[Slide 28] How accurate are the probabilistic methods now used by the United Nations? Out-of-sample validation tests reveal that, on average, projection errors are less than 3% for population and life expectancy and around 12% for total fertility over a 20-year horizon. Such tests also confirm that the prediction intervals describe well the range of subsequent outcomes.

[Slide 29] Now turning our attention to the methodology that underlies the IASA medium projection, we observe that expert-based methods have been used successfully to create future scenarios, which can be instructive for various purposes. However, there is considerable evidence that experts often fail to foresee the future accurately, as they tend to overlook trends that are contrary to their pre-existing biases.

[Slide 30] Concerning the alternative scenarios of the IASA projections, SSP1 and SSP3: these are supposed to represent, on the one hand, a future world of rapid social development that is moving toward sustainability; or, on the other hand, a future world that is increasingly fragmented with stalled social development and rising inequality. Obviously, the specification of coherent alternative scenarios presents a considerable challenge. There are, however, two aspects of these scenarios that I find especially implausible.

First, under the SSP1 scenario, there is an assumption that fertility will be lower in all countries, even those that are currently at very low levels. This is contrary to the existing evidence that people living in low-fertility countries often aspire to have more children than they are

currently having. It is also at odds with data presented earlier showing that an upturn in fertility has been quite common in such countries.

Second, under the SSP3 scenario, it is assumed that there will be relatively low levels of international migration. Yet, in a fragmented world with rising inequality, the pressure for cross-border migration will remain high. It is unclear in this context whether increased security would be sufficient to limit migratory flows.

[Slide 31] As noted earlier, an implicit assumption of the IIASA medium projection is that there will be a substantial acceleration in the pace of fertility decline in sub-Saharan Africa. Yet fertility decline in this region has, to date, been substantially slower than in other world regions at comparable stages of fertility transition. Although the experience of other regions shows that a faster decline in African fertility may be possible, convincing evidence of an accelerated transition is still missing.

[Slide 32] Lastly, let us consider the alternative scenarios of the IIASA projections. Earlier, I have shown you the probabilistic projections of the United Nations.

[Slide 33] And in that context, I noted that the IIASA medium projection cannot be dismissed as implausible according to that analysis.

[Slide 34] The same cannot be said, however, of the two alternative scenarios by IIASA, as illustrated here. This is another example of the aggregation fallacy mentioned earlier, which affects all deterministic projection scenarios, including the UN high- and low-fertility variants. While the assumptions of these alternative scenarios may be plausible when applied to a single country, the idea that all countries in all future years will deviate in the same way from the assumptions underlying the medium scenario is highly implausible.

[Slide 35] By way of conclusion, I emphasize that differences between the UN and IIASA projections over the next few decades are relatively small and mostly reflect differences in baseline values. Such differences have little relevance for expectations concerning the Sustainable Development Goals or other issues within a time frame of two or three decades.

Larger differences in the long run, however, are consequential for discussions of climate change and other environmental impacts of human activities. On the one hand, overstating likely future trends could be used to justify coercive policies. On the other hand, understating likely future trends could lead to complacency about the growth of human consumption and diminish the sense of urgency about limiting environmental impacts per capita.

Moving forward, I wish to propose that the UN and IIASA teams should work more closely together to understand better the sources of difference in their projections of global population, to provide accurate explanations of differences in their methods and results, and to encourage frank and open discussions of the implications of population projections for the future well-being of the world and its inhabitants.