Probabilistic demographic forecasts

*Nico Keilman**

1 Why use probabilistic forecasts?

The Norwegian parliament ('Stortinget') has decided to buy 52 new JSF jet fighters type F35. When the decision was taken, the price for one F35 was unknown because the plane was still under construction. Therefore, the Stortinget demanded a total cost forecast that had a 50% chance of being accurate. In 2012, the price estimate was NOK 61 billion (approximately EUR 6 billion). In addition, the Stortinget wanted to have a cost estimate that had an 85% chance of being accurate. This estimate was NOK 72 billion; see https://www.regjeringen.no/no/dokumenter/prop-73-s-20112012/id676029/sec9, Section 2.4.4.

Requesting a probabilistic cost forecast with a 50% or an 85% chance of being accurate is established practice for large (>NOK 500 million) public projects in Norway. Probabilistic forecasts are necessary because the future is uncertain. There are many different possible futures, and some are more likely to come to pass than others. A probabilistic forecast, as opposed to a deterministic forecast, quantifies the uncertainty about future developments. While a probabilistic forecast is not necessarily more accurate than a deterministic forecast, the former contains more information, which is useful for planning purposes.

Let us assume that in a specific area, only a deterministic forecast of the relevant variable(s) is available. This may leave room for political decisions to be made. An example is the 'shutdown' of the U.S. government in December 1995. President Clinton proposed a seven-year budget plan that would produce a $115 billion deficit over the seven-year period, according to the economic forecasts of the Congressional Budget Office (CBO). However, the Office of Management and Budgets (OMB) estimated rather more optimistically that the budget would be balanced at the end of the period. Republican leaders demanded that Clinton propose a plan that would be balanced when using the CBO numbers, rather than the OMB's; see https://www.washingtonpost.com/news/wonk/wp/2013/09/25/here-is-every-previous-government-shutdown-why-they-happened-and-how-they-ended/?utm_term=.4bb10057f0d9, shutdown # 17.

---

* Nico Keilman, Department of Economics, P.O. Box 1095 Blindern, N 0317 Oslo, Norway
E-mail: nico.keilman@econ.uio.no

The following example illustrates why probabilistic forecasts are much more useful than deterministic forecasts. Let us assume that we are casting a die. The outcome is random, and has yet to occur. Before we cast the die, someone asks us to predict the outcome. This can be any value between one and six – all outcomes are equally likely. The expected value (equal to the median value in this particular case) is 3.5, which is our best guess. However, the probability that 3.5 will come up is zero. The probability of a result in the interval $[2, 5]$ – i.e., a two, a three, a four, or a five – is two out of three. A die with *five* faces numbered one, two, ..., five has an expected value of three. This is our best guess if we have to predict the outcome. The probability that three will come up is only 20%, while the probability of a result in the interval $[2, 4]$ is 60%.

Now let us assume that we are predicting the population. The outcome is random. Why should we give only one (best guess/expected value, the probability of which is near zero) or three (high, medium, low) numbers, and not the whole set of possible results, like when casting a die? Many would argue that we cannot give the whole set because we do not know the probability distribution. However, we can estimate this distribution. Users should be told what the chance is of the real population falling between the high and the low value: is it 30%, 60%, or even 90%?

In the field of population forecasting, it is established practice to express uncertainty by means of variant predictions for fertility, mortality, and migration (UNECE 2017, 25). This leads to logical flaws in high and low variants of forecast results (Alho and Spencer 2005). Assume that $[B_L, B_H]$ and $[D_L, D_H]$ are high-low intervals for births B and deaths D. There is no guarantee that the two intervals will cover an equal probability range. Now assume that the intervals are given a probabilistic interpretation, such as 80% prediction intervals. $I = B - D$ is natural population growth. $[B_L - D_H, B_H - D_L]$ then has a higher than 80% probability of occurring (unless B and D are perfectly negatively correlated, which is unrealistic).

Another potential complication is that two variants that may be extreme for a certain variable need not be extreme for a different variable.

To illustrate additional problems that typically arise when forecast variants are used to reflect uncertainty, Table 1 shows the results for the 2014-based population forecast for Norway. The predicted number of people aged 20–66 varies from 3.42 to 4.05 million. The interval between these two extremes is 17.2% wide compared to the number in the variant with medium population growth. For the age group 67+, the interval is slightly wider (24.3%). However, the dependency ratio (defined as the ratio of the number of elderly people to the number of people aged 20–66) has a relative interval width of no more than 7.5%, which is much smaller than we would intuitively expect to observe. The explanation for this surprising finding is that the strong population growth variant combines high life expectancy with high fertility. These two variables move in the same direction for the population numbers in the first and second rows, but they counterbalance each other in the dependency ratio.

Table 1:
Number of people aged 20–66 and 67+, and the dependency ratio of elderly people, 2040, according to three forecast variants of population growth

	Weak growth (W)	Medium growth (M)	Strong growth (S)	(S-W)/M (%)
20–66 (mln)	3.42	3.63	4.05	17.2
67+ (mln)	1.10	1.26	1.41	24.3
Dependency ratio (67+/20–66)	0.32	0.35	0.35	7.5

Source: Statistics Norway.

2 Probabilistic demographic forecasting: general principles

Traditional *deterministic* population forecasts are computed by means of the so-called cohort-component model $K_{t+1} = P_t K_t$, used recursively. Here, K_t is a vector of population numbers broken down by age and sex as of time t; and P_t is a matrix of transition probabilities during the time interval $(t, t + 1)$ that account for fertility, mortality, and emigration. Immigration by age and sex is added as a vector. A forecast for the first year is obtained by combining the jump-off population K_0 with assumed values of the matrix P_0; for later years, the model is used recursively. This forecasting approach was first developed by Edwin Cannan in 1895. By the 1930s, it had become generally accepted by the statistical agencies of many countries (De Gans 1997). In the 1970s, the model was extended to include a regional breakdown of the population vector (multiregional model; see, e.g., Rogers 1995), or an extra dimension in general, such as educational level, labour market status, or household status (multistate/multidimensional model).

Thus, a deterministic forecast can be derived by combining specific assumptions about fertility, mortality, and migration. Different projections or scenarios can be produced by systematically combining different assumptions. Collectively, those different projections can give some impression of the degree of uncertainty, but not in any quantified way. The probability that an outcome will be above or below a given scenario is unknown (Dunstan and Ball 2016).

Various methods for *probabilistic* population forecasting have been developed since the 1960s, although Törnquist (1949) was probably the first to integrate probabilistic thinking into population forecasting. In this approach, the elements of the matrix P_t are random variables. Early contributions were made by Pollard (1966), Sykes (1969), Schweder (1971), and Cohen (1986). The initial aim was to find analytical solutions for the predictive distributions of the variables of interest. Due to correlations between components, ages and sexes, and autocorrelations in all variables, approximations were necessary. Later work (e.g., Keyfitz 1985; Kuijsten 1988; Lee and Tuljapurkar 1994) was based on Monte Carlo simulation.

In the past two decades, Alho's Model for Scaled Error has received some general acceptance. The simulation model is extremely flexible in specifying various correlation structures. The corresponding computer program PEP ('Program for Error Propagation') is available free of charge. The model assumes normal distributions (in the log scale) for the input parameters of the cohort-component model. See Alho (1998) and Alho and Spencer (2005). In the UPE (Uncertain Population of Europe) project, it was assumed that the expected values of the normal distributions would coincide with the mean variant of deterministic projections. The numerous (co-)variances and (auto-)correlations came from time-series analyses, errors in historical forecasts, and expert judgements. The volatility in the UPE forecast results was constructed to be similar to historical patterns of volatility. See http://www.stat.fi/tup/euupe/upe_final_report.pdf for the final project report and a general project description, and http://www.stat.fi/tup/euupe/index_en.html for stochastic forecast results for 18 countries in Europe.

3 Probabilistic demographic forecasting: the expert-based approach

During the second half of the 1990s, Wolfgang Lutz became one of the pioneers (together with Sergei Scherbov, Warren Sanderson, and Alexander Hanika) of probabilistic demographic forecasting. For example, applications were developed for Austria by Hanika et al. (1997) and Lutz and Scherbov (1998a); for Germany by Lutz and Scherbov (1998b); and for the global population by Lutz (1996) and Lutz et al. (1997, 1998, 2001).

Lutz and colleagues developed the so-called expert-based approach. Its main idea can be summarised as follows. Given a deterministic population forecast with high-low intervals for key forecast indicators (such as the total fertility rate, the life expectancy, or the level of net migration) together with point forecasts for these indicators (medium variant values), all in a future year t, experts can be asked to define the probability content for these intervals, such as 90%. Let us assume that the probability distribution at time t for the variables is normal, and perform random draws from that distribution by Monte Carlo simulation. Variables for intermediate years are found by straight-line interpolation between the value in the jump-off year and the drawn value for year t. In constructing such 'random slopes', perfect autocorrelation is assumed. This unrealistic assumption has been relaxed in later work to a time-series approach based on a moving average model (e.g., Lutz et al. 2001; Sanderson et al. 2004). The advantage of relying on the expert-based approach is that it is simple, and is thus easy to explain to demographers and forecast users with little or no statistical experience. However, a major drawback of using this approach is that there is no empirical basis for the (co-)variances implicitly selected by the experts. At the same time, experts tend to be overconfident in their opinions, which could result in the generation of forecasts that have a lower degree

of variability than would be desirable; and in the selection of prediction intervals that are too narrow (Alho and Spencer 2005). Such overconfidence is problematic, given that sensitivity analyses by Lutz et al. (1996) have shown that assumptions of perfect or zero correlation between fertility and mortality might have a big impact on prediction intervals for total population size and broad age groups in the long run.

The expert-based approach stimulated further refinements. For instance, Billari et al. (2012, 2014) addressed some of the weaknesses with respect to autocorrelation. They proposed applying a Bayesian interpretation to expert-based predictions, and treating expert inputs as data. While there are good reasons to be sceptical about the use of a method that depends heavily on the opinions of experts (overconfidence, selection), the Bayesian approach is emerging as the state of the art in probabilistic population forecasting (Bijak and Bryant 2016).

4 Examples of official probabilistic demographic forecasts

Official probabilistic population forecasts have been published by the statistical offices of Italy, the Netherlands (since 1999), and New Zealand.[1] The statistical agencies in a few other countries, such as Norway and Sweden, have experimented with probabilistic population forecasts. An important example of the use of probabilistic forecasting based on Bayesian methods can be found in the United Nations' World Population Prospects; see https://population.un.org/wpp/ and Raftery et al. (2012).

The case of Statistics New Zealand is worth a closer look. As Dunstan and Ball (2016) reported, the shift from using a deterministic approach to using a probabilistic approach is less difficult to make than might be expected. Uncertainty in the different components can be modelled simply or with more complexity. The authors mentioned a number of benefits that the producers and users of New Zealand's probabilistic forecasts have noticed since such forecasts were first published in 2012. First, while the probability distribution of a certain forecast result is often skewed, this message is not conveyed in a deterministic forecast. For example, the share of elderly people in the future population is usually around 15–20% in developed countries (depending on the definition of 'elderly'). Since any share is restricted to the interval [0, 1], an outcome that is five percentage points below the expected value is less likely than an outcome of equal magnitude above the expected value. Second, the fan chart for a particular forecast result tells us how

---

[1]   See ISTAT (2018),
http://statline.cbs.nl/Statweb/publication/?DM=SLEN&PA=82683eng&D1=0,7,10-12,14-15,21&D2=0-4&D3=0-1,6,11,16,21,26,31,36,41,l&LA=EN&VW=T and
http://www.stats.govt.nz/browse_for_stats/population/estimates_and_projections/
NationalPopulationProjections_HOTP2016.aspx.

much more likely small or large deviations from the expected value or the median value are. Third, any user who wants one number can use the median value. Fourth, the stochastic results indicate how far into the future we can reasonably provide a precise forecast, and the horizon is different for each variable. The New Zealand Treasury is developing a stochastic version of its fiscal model in order to get a better idea of how uncertain their variables of interest are.

Indeed, ". . . a probabilistic projection clearly gives more useful information than the usual high and low variants. . ." (Lutz and Scherbov 1998a, p. 13).

Finally, as UN Population Division director John Wilmoth stated in 2013: ". . . I expect that demographers will continue to be surprised by trends that do not follow our prior expectations. It is for this reason that the Population Division has worked hard in recent years to be more explicit and precise about the degree of uncertainty affecting projections of future population trends". http://www.un.org/en/development/desa/news/population/population-division-director.html

5 Quantifying uncertainty by means of the scenario approach

The book 'World population and human capital in the 21st century', edited by Lutz, Butz, and KC and published in 2014, represents an impressive summary of the work by Wolfgang Lutz and his team in the field of forecasting the world population broken down by age, sex, and educational level. It uses techniques of multistate modelling, but the forecasts are entirely deterministic. Here I would like to comment on two points: namely, the role of education and the use of scenarios.
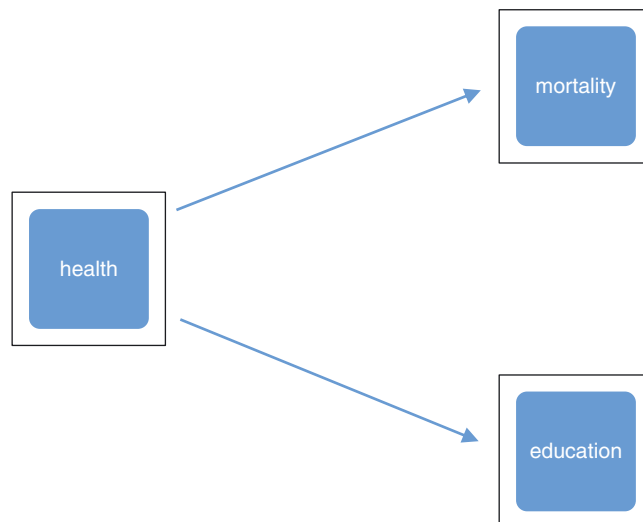
5.1 Education

Does adding the dimension of education to the variables of age and sex completely fill the gaps in our knowledge of fertility, mortality, and migration behaviour? As Samir KC, one of the book's editors, states: "We have shown in the past that knowing the education level of the population can help us make better projections" (Samir KC in https://blog.iiasa.ac.at/2016/11/25/interview-defining-the-futures/). However, uncertainty remains, because the links between education and components of demographic change are not well modelled. In the analyses and forecasts by Lutz et al., education is seen as having an impact on fertility, mortality, and migration. This is certainly the case, but causality runs in the opposite direction as well. First, the number of children a woman has can have a large effect on her daughter's education: if the mother has more children than she can afford to educate (given the school fees and the need for the children's labour), the daughter's education is likely to suffer. Second, adolescent girls and young women with one or more children have fewer opportunities to attend school than childless women. Reverse causality of this type has been demonstrated for women in a number of countries, including Norway (Cohen et al. 2011) and Denmark (Gerster

et al. 2014). An effect of this kind has also been observed in developing countries. For instance, a Tanzanian law dating back to the 1960s allows all state schools in the country to ban young mothers from attending. Over the past decade, more than 55,000 Tanzanian pregnant schoolgirls have been expelled from school (Center for Reproductive Rights 2013). For an early discussion of this issue, see Rindfuss et al. (1984). A population with high young adult mortality will have lower educational levels than a population with low mortality in this age group, all other things being the same. Brain drain affects the average educational levels of both the sending and the receiving country. Thus, the causal link runs not only from education to fertility, mortality, and migration; but in the opposite direction as well.
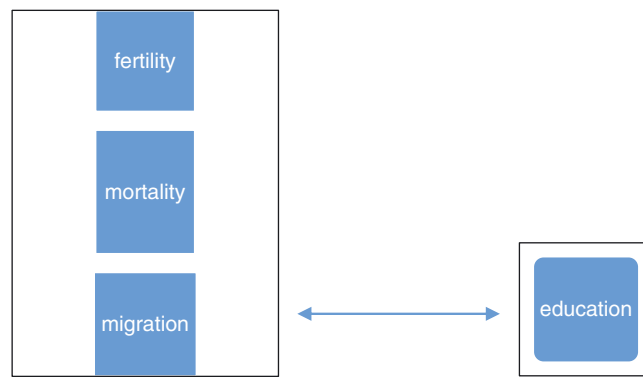
Alternatively, it can be argued that there are common underlying factors. For example, it is reasonable to assume that the health of the population aged 15–30 has an impact on both the mortality and the educational levels of the population as a whole, as being in better health leads to lower mortality and more years spent in education.

Another example is family background. Tropf and Mandemakers (2017) found a spurious correlation between education and fertility in twin data from the United Kingdom. The association between years of schooling and age at first birth depends strongly on family background factors (genetic and family environment).

In reality, both reverse causality and common underlying factors are likely to be at work. There is some discussion of causality issues in Chapter 2 of 'World Population and Human Capital in the Twenty-First Century'. This discussion is, however, restricted to cases in which education is thought to have an effect on demographic behaviour. The authors do not mention reverse causality or common underlying factors.

It therefore appears that the link between education and demographic behaviour was poorly modelled by Lutz and colleagues, and that uncertainty remains. Hence, there is a strong need to create probabilistic forecasts for populations broken down by age, sex and, education. We hope that developing such forecasts is on the research agenda at the Wittgenstein Centre.

5.2 Scenarios

'World Population and Human Capital in the Twenty-First Century' contains projections that contributed to an ongoing international effort to develop a new generation of global change scenarios related to the Intergovernmental Panel on Climate Change (IPCC). These scenarios (called Shared Socioeconomic Pathways, or SSPs) have been designed to capture how socioeconomic challenges will affect future climate mitigation and adaptation trends. Each SSP is based on a set of assumptions about the components of change, and background stories for those assumptions. The book presents the 'human core' of the broader SSPs, which also include many other dimensions, such as energy and economic variables. Of the five SSPs, let us look at the following three scenarios:

- SSP1 assumes that large investments will be made in education and health.
- SSP2 assumes that the current demographic and education trends will continue.
- SSP3 assumes that current progress in increasing school enrolment rates and decreasing fertility and mortality will stall.

Imagine a government official whose task it is to formulate policies that prepare his country for the challenges associated with climate change. He wonders which set of projections among the three mentioned above is more likely to be proven correct. Here is the answer from the demographer:

"Pick SSP1 if you believe its assumptions are likely; but
pick SSP2 if you believe its assumptions are likely; and
pick SSP3 if you believe its assumptions are likely".

This example illustrates the point that scenarios have limited relevance for the user. They are useful as a basis for discussion about consistent patterns between input

variables. Moreover, since they are examples of sensitivity analysis, they are relevant for modellers (and a few skilled users) who want to distinguish between important and less important input variables. Scenarios are less suitable for a general user, because they answer a question the user did **not** ask ('What . . . if. . .?'). Scenarios do **not** answer the question that the user asked ('What can we expect . . . and how probable is a different path?'); see Bijak et al. (2015).

Indeed, as Lutz and Goldstein (2004) argued, "people interested in uncertainty should either use completely probability-free scenarios for sensitivity testing or fully probabilistic projections for a more comprehensive view of the uncertainty involved".

Scenarios are also suitable for 'users' who are engaged in developing policies and want to reach a certain goal in the future, as they can use a certain scenario to check whether they are 'on track'. However, many different paths can be taken to achieve the same goal.

6 Conclusions

Wolfgang Lutz and colleagues have made major contributions to the field of probabilistic population forecasting. While some of the details can be criticised, the pioneering work of Wolfgang Lutz et al. in the 1990s was important because it showed that probabilistic population forecasting is feasible, and does not have to be very complicated. At the same time, it should be stressed that the current world population forecasts (including education) made by Wolfgang Lutz et al. and the IPCC forecasts based on SSP scenarios will be much more useful and relevant for the user once they are formulated in probabilistic terms.

7 Acknowledgements

References

Alho, J. M. 1998. A stochastic forecast of the population of Finland. Reviews 1998/4. Helsinki: Statistics Finland.

Alho, J. and B. Spencer 2005. *Statistical demography and forecasting*. New York: Springer.

Bijak, J., I. Alberts, J. Alho, J. Bryant, T. Buettner, J. Falkingham, J. Forster, P. Gerland, T. King, L. Onorante, N. Keilman, A. O'Hagan, D. Owens, A. Raftery, H. Ševčíková and P. W. F. Smith 2015. Letter to the Editor: Probabilistic population forecasts for informed decision-making. *Journal of Official Statistics* 31(4): 537–544.

Bijak, J. and J. Bryant 2016. Bayesian demography 250 years after Bayes. *Population Studies* 70(1): 1–19.

Billari, F. C., R. Graziani and E. Melilli 2012. Stochastic population forecasts based on conditional expert opinions. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 2(175): 491–511.

Billari, F. C., R. Graziani and E. Melilli 2014. Stochastic population forecasting based on combinations of expert evaluations within the Bayesian paradigm. *Demography* 51(5): 1933–1954.

Centre for Reproductive Rights 2013. New Report: Tanzania Forces Girls to Undergo Invasive Pregnancy Tests, Expels Pregnant Students. https://www.reproductiverights.org/press-room/new-report-tanzania-forces-girls-to-undergo-invasive-pregnancy-tests-expels-pregnant-stud.

Cohen, J. E. 1986. Population forecasts and confidence intervals for Sweden: A comparison of model-based and empirical approaches. *Demography* 23(1): 105–126.

Cohen, J. E., Ø. Kravdal and N. Keilman 2011. Childbearing impeded education more than education impeded childbearing among Norwegian women. *Proceedings of the National Academy of Science of the United States of America* 108(29): 11830–11835.

De Gans, H. 1999. *Population forecasting 1895–1945*. Dordrecht: Kluwer.

Dunstan, K. and C. Ball 2016. Demographic projections: User and producer experiences of adopting a stochastic approach. *Journal of Official Statistics* 32(6): 947–962.

Gerster, M., M. Ejrnæs and N. Keiding 2014. The causal effect of educational attainment on completed fertility for a cohort of Danish women: Does feedback play a role? *Statistics in Biosciences* 6(2): 204–222.

Hanika, A., W. Lutz and S. Scherbov 1997. Ein probabilistischer Ansatz zur Bevölkerungsvoraus-schätzung für Österreich. *Statistische Nachrichten* 12: 984–988.

ISTAT 2018. Il futuro demografico del paese: Previsioni regionali della popolazione residente al 2065 (base 1.1.2017). Report Statistiche 3 maggio 2018. Roma: ISTAT.

Keyfitz, N. 1985. A probability representation of future population. *Zeitschrift für Bevölkerungswissenschaft* 11(2): 179–191.

Kuijsten, A. C. 1988. Demografische toekomstbeelden van Nederland. Verkenningen no. 49 Planologisch en Demografisch Instituut, Universiteit van Amsterdam.

Lee, R. and S. Tuljapurkar 1994. Stochastic population forecasts for the United States: Beyond high, medium and low. *Journal of the American Statistical Association* 89: 1175–1189.

Lutz, W. ed. 1996. *The future population of the world: what can we assume today?*, Revised Edition. London: Earthscan.

Lutz, W., W. C. Sanderson and S. Scherbov 1996. Probabilistic population projections based on expert opinion. In *The future population of the world: what can we assume today?*, Revised Edition, ed. W. Lutz, 397–428. London: Earthscan.

Lutz, W. and S. Scherbov 1998a. An expert-based framework for probabilistic national population projections: The example of Austria. *European Journal of Population* 14: 1–17.

Lutz, W. and S. Scherbov 1998b. Probabilistische Bevölkerungsprognosen für Deutschland. *Zeitschrift für Bevölkerungswissenschaft* 23(2): 83–109.

Lutz, W., W. C. Sanderson and S. Scherbov 1997. Doubling of world population unlikely. *Nature* 387: 803–805.

Lutz, W., W. C. Sanderson and S. Scherbov 1998. Expert-based probabilistic population projections. In *Frontiers of population forecasting, supplement to population and development review*, eds W. Lutz, J. W. Vaupel and D. A. Ahlburg, 139–155.

Lutz, W., W. C. Sanderson and S. Scherbov 2001. The end of world population growth. *Nature* 412: 543–545.

Lutz, W. and J. Goldstein 2004. Introduction: How to deal with uncertainty in population forecasting? *International Statistical Review* 72(1): 1–4.

Pollard, J. H. 1966. On the use of the direct matrix product in analysing certain stochastic population models. *Biometrika* 53: 397–415.

Raftery, A., N. Li, H. Ševčíková, P. Gerland, G. Heilig 2012. Bayesian probabilistic population projections for all countries. *PNAS* 109(35): 13915–13921.

Rindfuss R. R., C. St John and L. L. Bumpass 1984. Education and the timing of motherhood: Disentangling causation. *Journal of Marriage and the Family* 46(4): 981–984.

Rogers, A. 1995. *Multiregional demography*. Chichester: Wiley & Sons.

Sanderson, W., S. Scherbov, B. C. O'Neill and W. Lutz 2004. Conditional probabilistic population forecasting. *International Statistical Review* 72(2): 157–166.

Schweder, T. 1971. The precision of population projections studied by multiple prediction methods. *Demography* 8(4): 441–450.

Sykes, Z. M. 1969. Some stochastic versions of the matrix model for population dynamics. *Journal of the American Statistical Association* 64: 111–130.

Törnquist, L. 1949. Om de synspunkter, som bestämt valet av de primära prognos-antaganda. *Beräkningar rörande Finlands befolkning, dess reproduktion och framtida utveckling. Statistiska Meddalanden 38*, eds J. Hyppolä, A. Tunkelo and L. Törnquist, 69–75. Helsinki: Statistiska Centralbyrån (in Finnish, Swedish and French).

Tropf, F. C. and J. J. Mandemakers 2017. Is the association between education and fertility postponement causal? The role of family background factors. *Demography* 54: 71–91.

United Nations Economic Commission for Europe – UNECE 2017. Recommendations on communicating population projections: Prepared by the Task Force on Population Projections. Document ECE/CES/2017/7.