Bayesian Probabilistic Population Projections for France

Vianney Costemalle*

Abstract – Population projections are performed regularly by national statistics institutes. In France, the most recent projections were produced by Insee in 2016 using a deterministic approach based on 27 different scenarios. In this article, we propose a new approach, which combines probabilistic population projections and a greater use of the Bayesian paradigm in order to quantify the uncertainty of future population levels without resorting to scenarios. Using the components method, the mortality rate, fertility rate and net migration are projected independently by sex and age. These three components are modelled, taking account of registry data (number of births and deaths) and net migration data series. The results reveal that the population of metropolitan France will continue to grow, reaching a level of between 66.1 million and 77.2 million inhabitants in 2070, with a probability of 95%.

JEL Classification: C11, C53, J11, J13, F22
Keywords: probabilistic projections, Bayesian inference, time series, population, mortality, fertility, migration

*Insee at the time of writing (vianney.costemalle@sante.gouv.fr)
Acknowledgments – The author would like to thank Julian Arbel, Junni Zhang, John Bryant, Marie Reynaud, Isabelle Robert-Bobée, Nathalie Blanpain, Guillemette Buisson, Vanessa Bellami, Catherine Beaumel as well as two anonymous referees.

Received November 2017, accepted July 2018.
Translated from: "Projections probabilistes bayésiennes de population pour la France"
Citation: Costemalle, V. (2020). Bayesian Probabilistic Population Projections for France. *Economie et Statistique / Economics and Statistics*, 520-521, 29–47. https://doi.org/10.24187/ecostat.2020.520d.2031

Population projections are performed regularly by statistics institutes around the world, as well as certain international organisations such as the United Nations (UN), which has published *World Population Prospects* (UN, 2017) every two or three years since 1951. Population projections offer many benefits and they have numerous users. They are primarily used to predict the possible future population of a region, a country or the entire world under certain assumptions, in terms of both number of inhabitants and structure. In the short to medium term, these projections form the basis for economic and social planning, such as pension funding (COR, 2017) or the construction of public infrastructure. They also form an essential part of certain other exercises, such as economic, climate or environmental projections.

In the case of France, the most recent official projections date from 2016 (Blanpain & Buisson, 2016a ; 2016b) and provide an indication of what the population will be in 2070 if past trends continue, with different variants on these assumptions (see Blanpain in this issue). Details of the projections by region, and in particular those for metropolitan France, are only available for the period from 2013 to 2050. This article aims to explore a new method for projecting the population of France: probabilistic projections. The proposed approach is said to be probabilistic since it allows the uncertainty surrounding future population levels to be quantified. This is where it differs from the traditional approach, which is a set of deterministic projections based on different scenarios. The fundamental difference between these two approaches is not so much the results themselves, but the way in which they are interpreted and used.

Probabilistic projections are based on statistical models, the majority of which are parametric. The uncertainty surrounding some elements making up the population can be captured by error terms, as is the case with time series, but it can also come from Bayesian inference of the model's parameters. The aim is to quantify the level of uncertainty surrounding the future population. This can be achieved using the stochastic approach, the Bayesian approach or even a combination of the two. In this article, we use stochastic models with Bayesian inference of the parameters.

In a letter to the editor of the *Journal of Official Statistics*, a group of demographers and academics from various countries highlighted the contributions and challenges of probabilistic projections in demography and called for more research and practice in this area by statistics institutes (Bijak *et al.*, 2015). They highlighted the fact that probabilistic projections have already been developed and used successfully in other disciplines, such as meteorology, climatology and even aviation. Bayesian statistics are also taking their time in breaking through into the field of demography. Although Bayes' theorem was established more than 250 years ago, it is only recently, with the appearance of MCMC (Markov Chains Monte-Carlo) algorithms in the 1980s and the explosive increase in computer processing power, that Bayesian inference has been used (Bijak & Bryant, 2016).

Some statistical institutes have already adopted the approach aimed at producing probabilistic population projections for their official statistics. This is the case in the Netherlands and New Zealand in particular. The Netherlands started producing probabilistic projections based on stochastic methods in 1998. New Zealand has also been reporting probabilistic population projection results since 2012 (MacPherson, 2016; Dunstan & Ball, 2016). The UN, which develops projections for all countries, eventually switched from a deterministic to a probabilistic method in 2014 (Costemalle, 2015). Furthermore, some elements of its projections are based on Bayesian inference.

The overwhelming majority of population projections are based on the component method, which consists of producing separate projections for the three key components of population dynamics, namely fertility, mortality and migration. The population at a given time is broken down into sex and age categories and is equal to the population during the previous period plus births and immigrants and minus deaths and emigrants. In this way, it is possible to chart the development of the population and its structure by sex and age category from one period to the next. In order to achieve this, the number of births by sex must be determined for each period, along with the number of deaths and the net migration by sex and age group. As regards births and deaths, the most common methods are based on projected fertility and mortality rates. However, probabilistic population projections remain an active area of research: there is no single method; on the contrary, there are almost as many approaches as there are types of data and they differ from one country to the next.

In the first part of this article, we will highlight the key differences between deterministic and probabilistic projections before going on to describe some of the different approaches that have been developed in demography with respect to probabilistic population projections. The second part is dedicated to describing French mortality, fertility and net migration data, and the third part looks at the presentation and validation of the models used for each of the three components. We finish by presenting the results of the probabilistic projections obtained in this manner for France, before going on to discuss the assumptions within the models.

1. Deterministic and Probabilistic Projections and Developments in Demography

Predicting the future is a difficult task and has given rise to the development of many different methods over the centuries. The most recent and sophisticated methods are based on mathematical models that attempt to detect certain patterns or invariants in the data and to extrapolate the trends observed, while also respecting certain constraints that may be imposed. Both deterministic and probabilistic projections require a certain degree of modelling of observed data and differ only in the nature of the forecasts made.

1.1. Deterministic and Probabilistic Approaches: Different Ways of Addressing the Future

In the first instance, what we are looking to project depends, from a deterministic standpoint, on certain parameters. The selection of these parameters represents a hypothesis that is also referred to here as a scenario. A scenario is then given detailing the way in which these parameters are considered most likely to develop on the basis of accumulated knowledge, expert opinions and intuition. A given scenario corresponds to one single possible projection, and the relationship between the two is deterministic. In cases where the scenario plays out as expected, the projection will be certain. Deterministic projections answer the question: "What would happen in the future if such a scenario were to occur?". Extreme scenarios can therefore be created to see how the future would pan out if they were to come true. Deterministic projections are thus a formidable tool when it comes to exploring the future on the basis of predefined scenarios. Any uncertainty in the projection then relies on the scenario coming true. Possible scenarios are formulated, but it is impossible to know how likely they are to occur. It could even be argued that the probability of them coming true is zero (if the values are continuous) or very low (if the values are discrete). The degree of probability is estimated intuitively and is reflected in the terms used to describe these scenarios: demographers refer to the "central" scenario, which is the scenario considered the most plausible based on current knowledge, and "extreme" scenarios.

Conversely, probabilistic projections are based on models that attempt to take account of the uncertainty stemming from a lack of knowledge of certain aspects of the projections. These models are based on assumptions made on the basis of expert judgement and intuition. The underlying assumptions on which models for probabilistic projections are based are the equivalent of the scenarios used for deterministic projections. The advantage of probabilistic projections is that they make it possible to quantify the uncertainty based on past developments and to extrapolate it into the future to provide confidence intervals for the projections. The interpretation and use of probabilistic projections therefore differs from that of deterministic projections.

By way of an example, weather forecasts have long been making use of probabilistic projections: we are not only told whether or not it will rain the next day, but also the probability that rain will fall (Raftery, 2014). Since future events are inherently uncertain, indicating the probability of their occurrence in view of current knowledge provides more information than a deterministic projection based on a scenario. In economics in particular, time series are used as a means of producing probabilistic projections: in the case of a simple random sampling method, for example, we know that the variance increases with the square root of time.

By adding error terms to the models, it is therefore possible to create stochastic probabilistic projections. Another method for quantifying uncertainty is to use the Bayesian paradigm. Under this method, the model parameters are viewed as random variables, in the same way as error terms in stochastic models. Bayesian inference then involves estimating the *a posteriori* distribution of these parameters, i.e. after the data have been observed. This distribution gives possible values for the parameters, together with their degree of probability. It differs from the *a priori* distribution, which is the distribution given by the modeller and which is intended to reflect the knowledge of the problem before any data has been observed.

1.2. Probabilistic Projections in Demography: A Wide Variety of Models in Practice

Population projection techniques can be divided into three categories (Booth, 2006). The first group includes methods based on the extrapolation of trends, which seek to extend the trends identified in the past, in most cases in a linear fashion. They are based solely on past data and do not attempt to explain the mechanisms underlying the developments. They often prove to be effective. The second set of methods used for population projections involves establishing long-term trends. These methods are based on the expectation that the future will unfold in a certain way. This may be backed up by expert opinions, which assess what could be expected to happen in the future on the basis of current knowledge, or on people's intentions, such as those measured by fertility intention surveys (Régnier-Loilier & Vignoli, 2011). Finally, the last category of projections is made up of the structural models, which attempt to explain the mecanisms of population changes using exogeneous variables. These exogeneous variables must then be projected in accordance with one of the three projection categories. The approaches often combine several of these techniques and the techniques used differ according to the components (mortality, fertility and migration) that are to be projected.

A classic method of projecting mortality was developed by Lee & Carter (1992) and consists in decomposing the change in the logarithm of mortality rates into an age effect and a time effect, specific to each age. The time effect is then considered as a time series for which the parameters are estimated. By calculating or simulating the future values of this time effect on the basis of the models used a very large number of times, it is possible to obtain a probabilistic projection. The basic idea of this approach is to capture the regular changes in the data and to extrapolate these regularities. The Lee-Carter method has since been used very frequently to project mortality, as well as to project fertility and migration. Wiśniowski *et al.* (2015) put forward a more extended version of this, adding a generation effect, which can be applied to all three components of population change. In addition, these authors have proposed that these projections be carried out in an entirely Bayesian framework. The Lee-Carter model has also been generalised by Hyndman & Ullah (2007), who break down the logarithm for mortality rates or fertility rates into key components before extending the coefficients of each of those components using time series. Furthermore, Hyndman & Booth (2006) suggest performing a Box and Jenkins transformation on the rates studied with a view to generalising the log transformation. This approach is entirely stochastic.

The whole point of probabilistic projections is to allow the degree of probability of future projections to be quantified. In 2001, Lutz *et al.* (2001) announced that the world population is likely to stop growing by the end of the century. More specifically, their stochastic models and calculations predict that there is an 85% probability that the world population will begin to decline by the end of the century. The UN, which regularly publishes population projections, began using a probabilistic and Bayesian method in 2014. The results give a different view of the development of the population in the long term. In fact, they show that the world population is unlikely to have stopped growing by 2100 (Gerland *et al.*, 2014). The methodology used differs from that applied by Lutz *et al.* (2001): the aggregated values, which are life expectancy at birth and the total fertility rate (TFR), are projected directly in a first step. These indicators are then decomposed in sex-specific and age-specific mortality rates and age-specific fertility rates. In order to project life expectancy, the amount by which life expectancy increases every five years is modelled by a double logit function on the basis of actual life expectancy and a large number of parameters. These parameters are estimated by Bayesian inference, which leads to an *a posteriori* distribution of increases in life expectancy and therefore an *a posteriori* distribution of life expectancy itself by 2100 (Raftery *et al.*, 2013). This is an example of a probabilistic projection that does not use stochastic terms, but is instead based solely on parametric modelling and Bayesian inference. For its part, the TFR indicator is modelled according to a three-phase process of development: a phase of high fertility rates, a phase of rapid fertility decline to below the generation replacement level, and a phase of stagnation of the fertility rate with a long-term convergence towards a level of 2.1 children per woman (Alkema *et al.*, 2010).

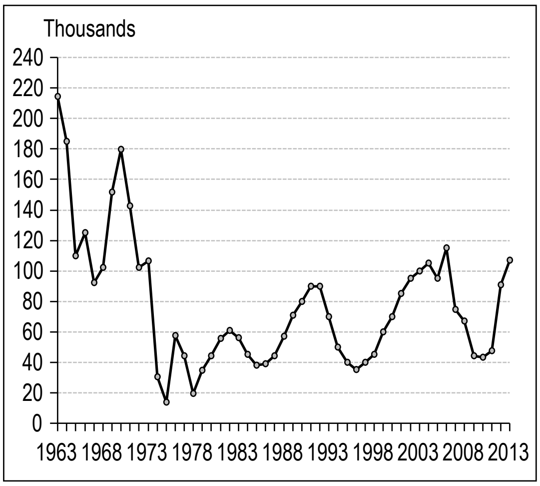
It therefore appears that there are numerous models available to project each of the three components. Working on the assumption that no single model can capture the full range of possible assumptions about mortality trends, especially when these assumptions are not consistent with one another, Kontis *et al.* (2017) made use of 21 different probabilistic projection models, the results of which were then weighted in accordance with the performance of each of the models, in order to ultimately obtain a single probability distribution for the desired indicators.

2. Data for France

To ensure that we have long series, we will restrict ourselves to the area of metropolitan France. We therefore have, for the years 1962 to 2013, the total population on 1 January of each year, the annual net migration, the number of deaths and the number of births by the age of the mother, all detailed by sex and age.[1] We have selected the same projection horizon as that used for the most recent official projections for France (Blanpain & Buisson, 2016b). The aim is therefore to project the 2014 population to 2070. Between 1962 and 1998, the data are not broken down by age beyond 100 years. From 1999 onwards they are broken down in detail up to 110 years. We then chose to retain the one-year age categories since the data are available, and we created a higher age category representing people 100 years of age and older. In the remainder of this section, we will describe the net migration, mortality and fertility data, highlighting invariants, trends and irregularities.

Net migration is the number of people in a given year who come to live in France from outside of metropolitan France, regardless of their nationality, minus the number of people living in metropolitan France who move abroad. It is undoubtedly the most difficult component to measure, because, although the number of people entering the country can be estimated using the population census (Brutel, 2014), we do not know how many have left. Net migration can therefore be calculated as the difference between the changes in the population and the natural balance. Unlike many other European countries, France does not have a population register and must therefore rely on the population census to estimate migration flows. As the census was only conducted once every 7-8 years or so until 1999, it was not possible to directly calculate the change in the population from one year to the next. In 1962, net migration was exceptionally high as a result of approximately 860,000 French nationals returning from Algeria; from 1963 onwards, net migration has been consistently positive, but the numbers have been much lower: it averaged 64,000 over the period from 1963 to 2013. Net migration appears to have remained stable on average from the 1990s onwards, although there have been some large fluctuations (Figure I), largely due to the various policies pursued, but also as a result of the economic and international context. For the period from 1990 to 2013, net migration was, on average, 72,000 and 79,000 for the last ten years available (2004-2013).

Figure I – **Changes in net migration between 1963 and 2013**



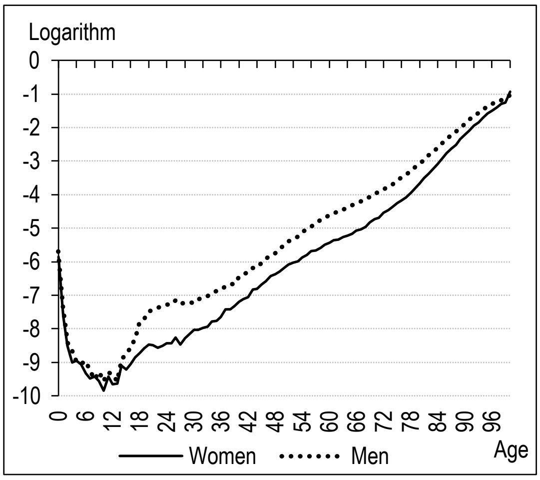
Sources and coverage: Insee, population estimates and civil registry statistics; Metropolitan France.

In order to describe mortality, the number of deaths must be related to the corresponding population at risk. This population is counted in person-years and takes account of the total time spent by all persons residing in France. It is approximately equal to the population present on 1 January, plus half of the net migration. By relating the number of deaths to this population, we then obtain the mortality rates, which can be broken down by sex, age and year. Mortality rates grow quasi-exponentially from the age of 25 upwards (Figure II). Before the age of 25, the profile is different due to infant mortality, which is higher for newborns. Mortality rates decline from birth until around 10 years of age, before rising steadily. At around the age of 18, the mortality of men becomes significantly higher than that of women, and the gap remains present throughout life, with a greater or lesser magnitude depending on age.

The logarithm of mortality rates, for a fixed age and sex, decreases in an almost linear manner over time (Figure III). This is especially true for older age groups, but does not seem to be quite the case for younger age groups. For example, the logarithm of the mortality rate at 10 years of age decreases faster and faster. Conversely, at age 30, the logarithm of the mortality rate slows its decline until it stagnates for males from the early 1980s to the mid-1990s, at which point

1. 2013 is the latest year for which all these data were final when the projections presented in this article were carried out, in 2017. In particular, the figure for net migration was not yet available for 2014. We have not used the provisional data, then available up until 2016, but which are revised from one year to the next before being final and therefore of a different nature from the final data.

Figure II – **Logarithm of mortality rates in 2013 by sex and age**



Sources and coverage: Insee, population estimates and civil registry statistics; Metropolitan France.

mortality declines sharply for that age group and has continued to decline steadily and in an apparently linear fashion since. This stagnation in mortality among young adults in the 1980s and 1990s, when the general trend was towards a steady decrease in mortality, is linked to the AIDS epidemic, which reached France in the early 1980s. In general, as mortality rates are steadily declining, life expectancy at birth is increasing each year, and more rapidly for men than for women (Blanpain, 2016), although life expectancy sometimes decreases from one year to the next, as was the case in 2015 for cyclical reasons (Bellamy & Beaumel, 2016).

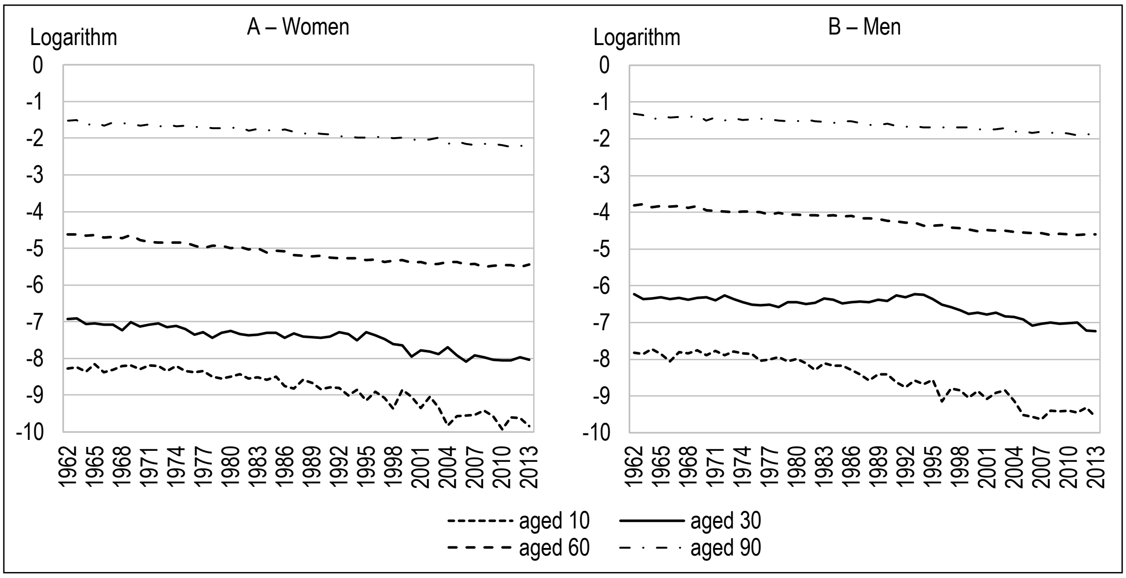
Since the early 1970s, the TFR[2] has declined sharply from 2.9 children per woman in 1964 to 1.8 children per woman in 1976 (Figure IV). It has since stabilised at an average of around 1.85 children per woman. Nevertheless, an upward trend has been observed in the TFR since the mid-1990s.

The fertility rate at a given age is defined as the ratio of the number of babies born to mothers of that age to the number of women of the same age in the year in question. This number corresponds to the number of women on 1 January of the year plus half of the corresponding net migration and minus half of the deaths recorded for this population. The profile of age-specific fertility rates follows a bell curve: the probability of having a child in a given year increases with age from 15 years until it peaks, after which it declines continuously, reaching zero or close to zero around 50 years.

Over time, this age distribution tends to shift to the right: the age at which peak fertility is reached increases (Figure V). In 1970, the fertility rate was at its highest at 24 years of age, while in 2013, the peak was reached at the age of 30. The maximum level of fertility reached during the year has barely changed since the mid-1970s: it fluctuates around 0.15. As the fertility peak moves to the right, the distribution of age-specific rates becomes increasingly symmetrical, as evidenced by the measure of skewness, which is rapidly decreasing towards 0 (Figure VI).

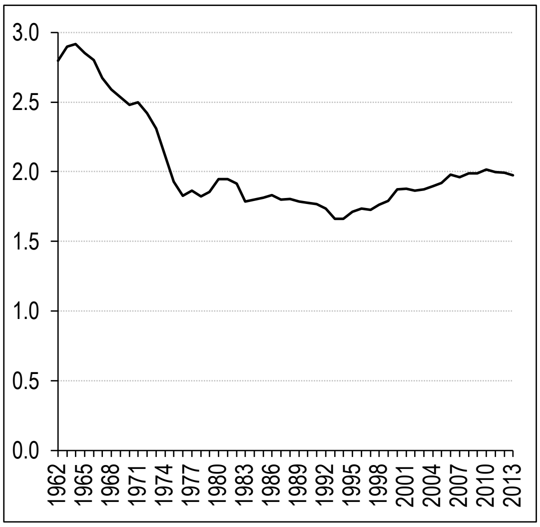
2. The total fertility rate (TFR) is calculated as the sum of the age-specific fertility rates. It corresponds to the average number of children that a woman would have during her lifetime if the probability of giving birth at a given age corresponded to the fertility rate at that age.

Figure III – **Changes in the logarithms of mortality rates from 1962 to 2013 for different ages**



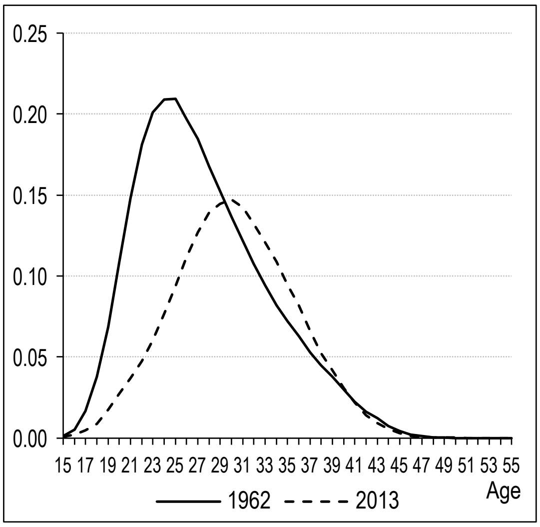
Sources and coverage: Insee, population estimates and civil registry statistics; Metropolitan France.

Figure IV – **Changes in the total fertility rate from 1962 to 2013**



Sources and coverage: Insee, population estimates and civil registry statistics; Metropolitan France.

Figure V – **Age-specific fertility rate in 1962 and in 2013**

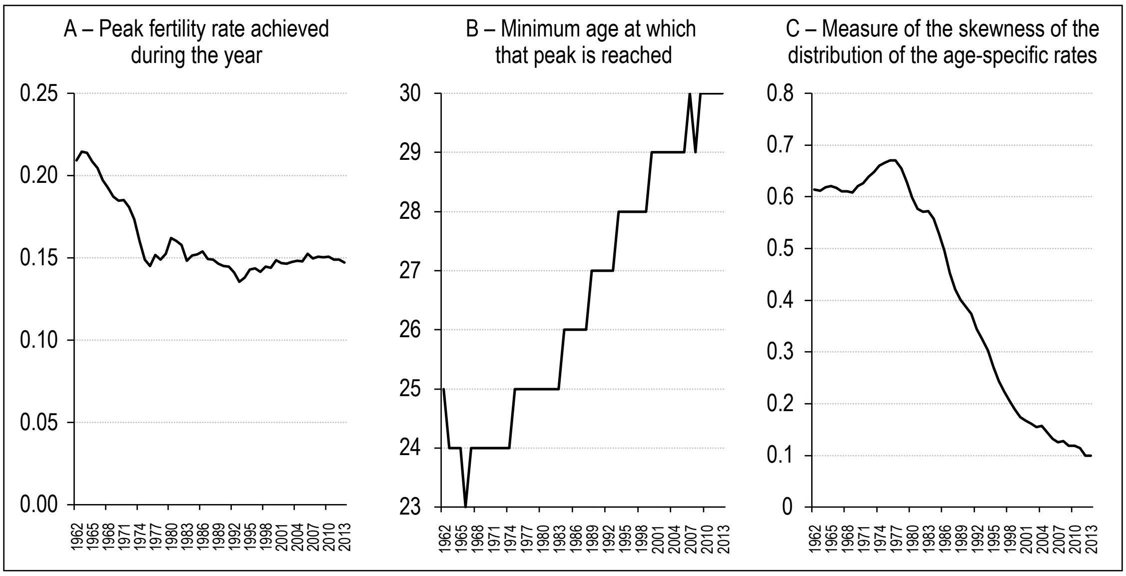


Sources and coverage: Insee, population estimates and civil registry statistics; Metropolitan France.

Unlike mortality rates, changes in fertility rates do not occur in a regular manner over time. For example, the fertility rate at 30 years of age fell between the early 1960s and the mid-1970s; however, it has been increasing since then, albeit with a slowdown from the 2000s onwards. The fertility rate at 20 years of age had been declining since the 1970s, but in the late 1990s rebounded slightly for a few years before declining again, but at a much slower pace than in previous decades. The changes are neither monotonous nor linear, which highlights how difficult it is to extend these curves into the future.

To summarise, net migration in metropolitan France appears stable over the long term, but with significant fluctuations that seem difficult to predict. Mortality has been moving in the same direction for several decades, with an almost linear decrease in the logarithm of mortality rates at all ages and a narrowing of the gap in life expectancy between women and men. Recent fertility trends are more complex to identify, but the evidence suggests that the TFR has stabilised at an average level of just under 2 children per woman and that the distribution of age-specific fertility is changing continuously with a shift in

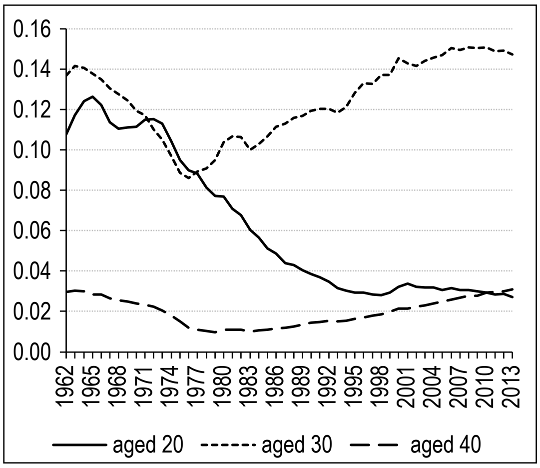
Figure VI – **Changes in peak fertility, age at which peak fertility is reached and skewness of the age distribution of fertility rates between 1962 and 2013**



Sources and coverage: Insee, population estimates and civil registry statistics; Metropolitan France.

peak fertility to higher ages and an increasingly symmetrical distribution (Figure VII). In the next section, we propose a model for each of the three components of population change, taking account of these observations and drawing upon models that have already been developed internationally and which we will describe briefly in the third section.

Figure VII – **Changes in fertility rates at different ages between 1962 and 2013**



Sources and coverage: Insee, population estimates and civil registry statistics; Metropolitan France.

3. Methods and Models

In the remainder of this article, we will use the following notations:

P(a,n,s): the number of people on 1 January of year *n*, of sex *s* born in year *n-a*;

D(a,n,s): the number of deaths in year *n*, of people of sex *s* and born in year *n-a*;

N(a,n,s): the number of live births of infants of sex *s* during year *n* and whose mother was born in year *n-a*;

M(a,n,s): the number of persons entering metropolitan France minus the number of people leaving metropolitan France during year *n*, of sex *s* and born in year *n-a*. This is the net migration in year *n*, for persons of sex *s* and born in year *n-a*.

To simplify the subsequent notations, we define *P(0,n,s)* as the number of live births in year *n* of babies of sex *s*. Furthermore, *D(0,n,s)* and *M(0,n,s)* are well defined by the above description and correspond respectively, for each year *n* and sex *s*, to the number of deaths of babies born in year *n* and to the number of newborns entering the country minus the number of newborns leaving the country. It will be assumed that the ages of women at childbirth are between 15 and 55 years inclusive, meaning that *N(a,n,s)*=0 for $a \leq 14$ and $a \geq 56$.

In addition, populations at risk are defined for deaths and births. Populations at risk are counted in person-years and depend on the number of people observed, but also on the period of time over which these people are present. For deaths, this corresponds to the population on 1 January for the year in question, plus half of net migration (assuming that inflows and outflows are evenly distributed throughout the year).

$$R_D(a,n,s) = P(a,n,s) + 0.5M(a,n,s), \text{ if } a \geq 1$$

$$R_D(0,n,s) = 0.5P(0,n,s) + 0.5M(0,n,s),$$

where $a = 0$.

For births, the number of person-years at risk is the average number of women during the year in question, assuming that migration flows and deaths remain uniform:

$$R_N(a,n) = P(a,n,women) + 0.5M(a,n,women) - 0.5D(a,n,women).$$

We also note $M(n) = \prod_{a,s} M(a,n,s)$, $N(a,n) = N(a,n,girls) + N(a,n,boys)$ and $N(a,n) = \prod_{a} N(a,n,s)$.

When noting normal distributions, we will indicate the standard deviation (rather than the variance).

3.1. Migration

The total net migration is directly projected using a first-order autoregressive model, where M_{lt} represents the long-term net migration and ε_M represents white noise:

$$M(n) = M_{lt} + \rho_M(M(n-1) - M_{lt}) + \varepsilon_M(n)$$

$$\varepsilon_M(n) \overset{i.i.d.}{\sim} N(0, \sigma_M)$$

In order to ensure a stationary process, the constraint $|\rho_M| \leq 1$ is imposed. This modelling reflects the fact that it is estimated that net migration will continue to be stable on average and will oscillate around a long-term trend. The amplitude of possible future oscillations is determined by past amplitudes. Furthermore, a very informative *a priori* is set with regard to the long-term trend by assuming, as was the case in the work of Blanpain & Buisson (2016a), that this can be estimated from the average net migration over the recent period, i.e. 80,000 persons. The *a priori* distribution for the long-term trend is therefore $M_{lt} \sim N$(80,000;

10,000). The parameters M_{lt}, ε_M, ρ_M and σ_M are estimated by means of Bayesian inference based on the net migration for the period 1995-2013.

To project the total net migration, the model parameters are randomly drawn 1,000 times according to their *a posteriori* distribution and for each set of parameters, the development of net migration is simulated according to the first-order autoregressive process. Once the net migration has been projected, it is broken down by sex and age in accordance with fixed rates calculated on the basis of the distribution of net migration by sex and age over the recent period and smoothed, as described in Blanpain & Buisson (2016a).

3.2. Mortality

As has already been mentioned, the logarithm of age-specific mortality rates appears to develop in a linear manner over time. Nevertheless, mortality rates develop at a different rate for each age over time. The number of deaths observed is directly modelled in accordance with Poisson's law, which is based on the mortality rate and the population at risk. The latter corresponds to the number of person-years present in metropolitan France in the year in question. Poisson's law is currently used to model a number of events occurring over a given period of time. It is often used to model the number of deaths in demographic work. The following model (developed by Bryant & Zhang, 2014) is used, where $\mu_D(a,n,s)$ corresponds to the mortality rate for year n for persons of sex s and age a:

$$D(a,n,s) \sim Poisson\left(\mu_D(a,n,s)\ R_D(a,n,s)\right)$$

$$\log\left(\mu_D(a,n,s)\right) = \beta^0 + \beta_a^{age} + \beta_{a,s}^{age:sex} + \beta_{a,n}^{age:year} + \varepsilon_{D,1}(a,n,s)$$

$\varepsilon_{D,1}$ are independent and identically distributed error terms according to centred normal distribution and standard deviation $\sigma_{D,1}$. The parameter β_0 is a constant, the parameter β^{age} gives the average age distribution of the logarithm of mortality rates. Finally, there are two terms that cross two dimensions: $\beta^{age:sex}$, which allows the specific effect of sex to be estimated for each age and $\beta^{age:year}$, which is a time effect specific to each age. It should therefore be noted that the development of the logarithm of age-specific mortality rates over time is the same for both women and men, since no term that crosses the dimensions of year and sex has been specified. This is because we wanted to limit the number of parameters to be estimated. When a term crossing the year-sex dimension was introduced, it was found that the *a posteriori* distribution was not correctly estimated due to a non-convergence of the Markov chains. At a third level, some of the parameters are modelled by means of dynamic linear models. For the $\beta^{age:year}$ parameter, this allows the development over time to be broken down, by age, into a level ($\theta^{age:year}$) and a trend ($\delta^{age:year}$):

$$\beta_{a,n}^{age:year} = \theta_{a,n}^{age:year} + \eta(a,n)$$
$$\theta_{a,n}^{age:year} = \theta_{a,n-1}^{age:year} + \delta_{a,n}^{age:year} + \upsilon(a,n)$$
$$\delta_{a,n}^{age:year} = \delta_{a,n-1}^{age:year} + \omega(a,n)$$

The terms η, v and ω are independent error terms that follow centred normal distribution.

To project age-specific mortality rates into the future, once the *a posteriori* distribution of all the model parameters has been estimated, it is sufficient to generate new trend terms, followed by new level terms and finally new $\beta^{age:year}$ parameters, up to the desired horizon.

3.3. Fertility

For fertility, we chose to proceed in three stages. First of all, the TFR is projected according to a first-order autoregressive model. The UN uses the same method for its third stage of fertility change, on the assumption that the TFR tends towards 2.1 in all countries (Alkema *et al.*, 2010). When compared with the method used by the UN, we have chosen to also estimate the parameters of the model by means of Bayesian inference rather than by maximum likelihood. We therefore remain within an entirely Bayesian framework for all our estimates and projections. The model is as follows:

$$ICF(n) = ICF_{lt} + \rho_F\left(ICF(n-1) - ICF_{lt}\right) + \varepsilon_F(n)$$

where $ICF(n) = \prod_{a=15}^{55} \frac{N(a,n,girls) + N(a,n,boys)}{R_F(a,n)}$ is the total fertility rate in year n. As was the case for net migration, after estimating the Bayesian inference, we simulate 1,000 possible trajectories for the development of this index up to the desired horizon.

The second stage consists of projecting the age-specific fertility rates μ_F, independently of the projection of the TFR. As is the case with mortality, these are defined by modelling births by means of a Poisson process:

$$N(a,n) \sim Poisson\left(\mu_F(a,n)\ R_F(a,n)\right)$$

by way of a reminder, $N(a,n)$ corresponds to the number of births in year n, given by mothers born in year n-a. Following the method proposed by Bijak *et al.* (2015), which is based on the

Lee-Carter method, we then modelled the logarithm of the fertility rate as the sum of a fixed age effect, a time effect for which the intensity and direction are different for each age, and a generation effect:

$$\log\left(\mu_F\left(a,n\right)\right)=\alpha_a+\beta_a\kappa_n+\gamma_{n-a}+\varepsilon_{F,1}\left(a,n\right)$$

The time effect κ and the generation effect γ change in accordance with the first-order autoregressive processes:

$$\kappa_n=\varphi_0+\varphi_1\kappa_{n-1}+\xi\left(n\right)$$

$$\gamma_{n-a}=\Psi_0+\Psi_1\gamma_{n-a-1}+\zeta\left(n\right)$$

where the error terms ξ and ζ follow normal laws of zero expectation. Once again, all parameters are estimated by Bayesian inference in order to subsequently produce 1,000 fertility rate simulations for each age and each future year. These projected rates extend linear trends, although the parameters φ_1 and Ψ_1 may, if they are strictly smaller than 1, cause the time effect or the generation effect to cancel out in the long term. The estimates give an *a posteriori* distribution of φ_1 and Ψ_1, which are very close to 1. This results in the fertility rates becoming abnormally high for certain ages, which leads to TFRs that are much higher than those projected in the first stage.

The third stage involves then correcting the age-specific fertility rate for each year and aligning it to the TFR initially projected. In order to do so, we simply multiply all of the rates in a given year by a constant. Note that no constraint was added for the average age at childbirth, whereas Insee's projections retain a ceiling at 32 years old based on experts' opinion (see Blanpain, this issue).

Lifetime fertility is based on the fertility rates of a given generation of women. Like the TFR, it is a synthesis of fertility rates at different ages. However, unlike the TFR, which is a cross-sectional indicator, this is a longitudinal indicator and therefore requires the fertile life of an entire generation to be observed before it can be calculated. This therefore limits the number of observation points in the past. This is why we, like many other authors, decided to model and project the total fertility rate. Life expectancy is also a cross-sectional indicator.

3.4. Projections Using the Components Method

The components method makes it possible to develop the population from one year to the next by noting that the population on 1 January of a given year is equal to the population on 1 January of the previous year, plus the number of births that took place during the previous year, minus the number of deaths and plus net migration. This translates into the following equations:

$$P\left(a,n,s\right)=P\left(a-1,n-1,s\right)-D\left(a-1,n-1,s\right)+M\left(a-1,n-1,s\right)$$

if $a\geq 1$ and $P\left(0,n,s\right)=N\left(n,s\right)$.

The number of deaths and births are obtained each year by means of random sampling in accordance with Poisson's law (see models). In order to do this, the persons at risk must be identified for deaths and the women at risk in the case of births. We begin by calculating deaths for each age, with the exception of deaths among newborns. We then deduce the women at risk for each age between 15 and 55 years (in order to do so, we need to know the figures for net migration and the number of deaths). Finally, we calculate the number of deaths among newborns. The distribution of the number of births in a given year between male and female is determined by the sex ratio, which is set at 1.05 in accordance with past observations.

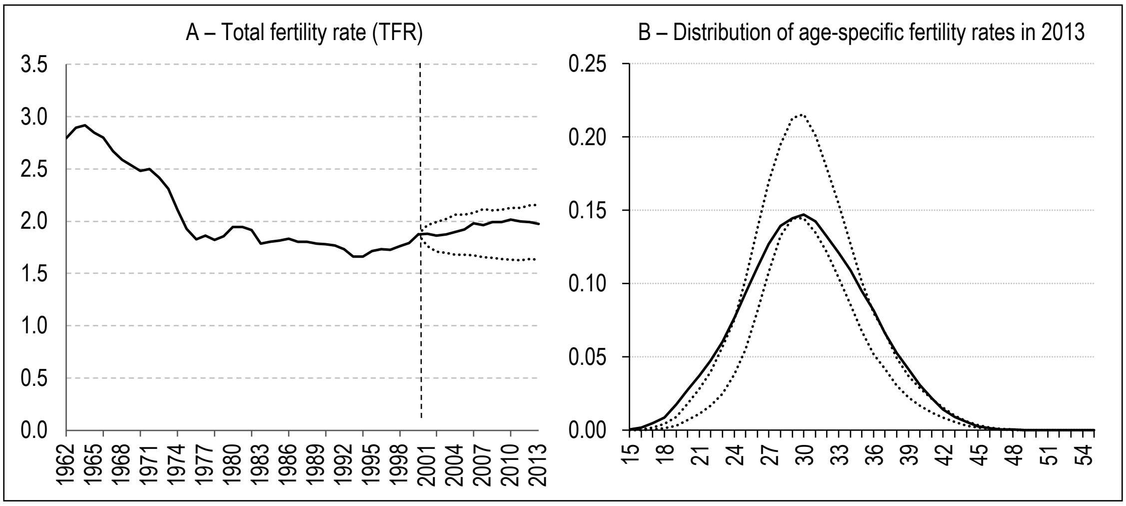
3.5. Validation of the Models

One way of testing the models used is to separate the data relating to the past into two categories: one part, approximately two-thirds, is used to estimate the models and the remaining part, approximately one-third, is used to compare the model estimates with reality.

In the case of mortality, we decided to estimate the model for the period from 1962 to 1995 and to compare the results during the period from 1996 to 2013. For fertility, we estimated the models over the period from 1975 to 2000 and we compared the results from the period between 2001 and 2013. It is clear that the logarithm of mortality rates is projected adequately at older ages (from around 35-40 years of age), but that the model used presents decreases in these rates that are much slower than what is actually observed. This is because, at very young ages, the logarithm of mortality rates is not linear, but is instead slightly concave. Moreover, mortality rates for young adults more or less stagnated in the 1980s and 1990s, before falling sharply. The model was not able to predict this sudden drop.

As regards fertility, the TFR observed is well within the 95% confidence interval of the probabilistic projections of the TFR. However, when we look at the distribution of age-specific fertility rates, it becomes apparent that the method used leads to a tighter distribution than is actually observed (Figure VIII). The deformation of the distribution of age-specific fertility

Figure VIII – **TFR and age-specific fertility rates, observed (1962-2013) and projected (2001-2013)**



Notes: The dotted lines indicate the 2.5% and 97.5% quantiles of the probabilistic projections and the solid line indicates the actual TFR and fertility rates (from 1962 to 2013).
Sources and coverage: Insee, population estimates and civil registry statistics (fertility rates), Metropolitan France. Author's calculations (probabilistic projections)

rates is therefore a little too pronounced in our projections.

4. Results of Bayesian Probabilistic Projections for France up to 2070

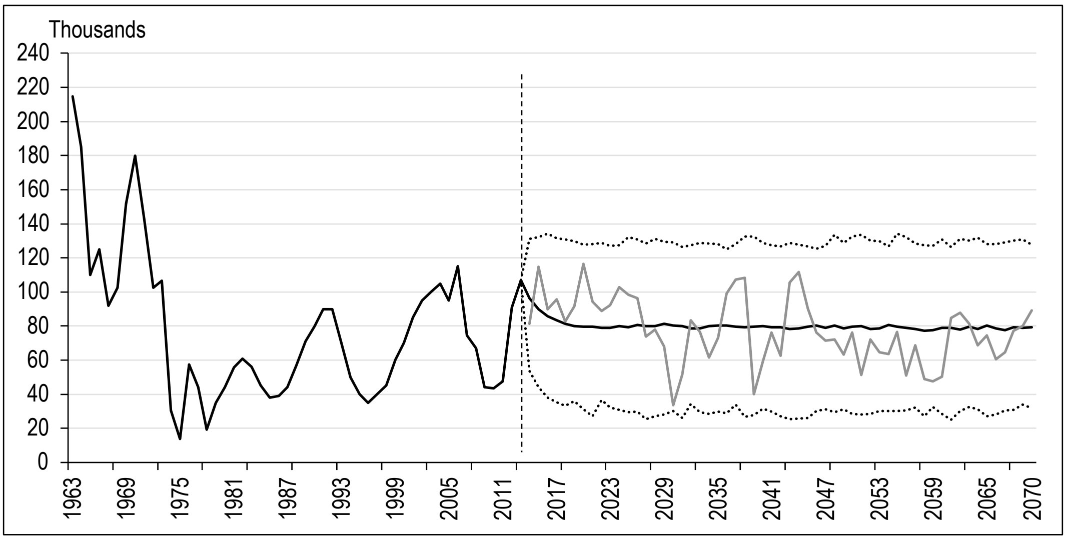
The parameters of the models for net migration, mortality and fertility were estimated by Bayesian inference using the open source software, *Stan* and the R *demest* package published by the Statistical Institute of New Zealand.[3] We simulated 1,000 values for each of these parameters according to their *a posteriori* law. We then generated 1,000 possible evolution trajectories for net migration, sex-specific and age-specific mortality rates and age-specific fertility rates. In the end, 1,000 estimates can be obtained for any demographic indicator derived from these three components, including the size of the total population. Confidence intervals of 95% or 80% are then derived from these, which contain 95% or 80% of the estimates, respectively.

4.1. Migration Projections: A Strong and Constant Uncertainty

Projected net migration follows a stable trajectory as this was specified in the model. The median of the 1,000 possible trajectories decreases in

3. *https://github.com/StatisticsNZ/demest*

Figure IX – **Net migration, past and projected**



Note: The dotted lines indicate the 2.5% and 97.5% quantiles and the solid line indicates the median of the *a posteriori* distributions. The light grey curve represents one of the 1,000 simulations.
Sources and coverage: Insee, population estimates and civil registry statistics (1962-2013), author's calculations (2013-2070); Metropolitan France.

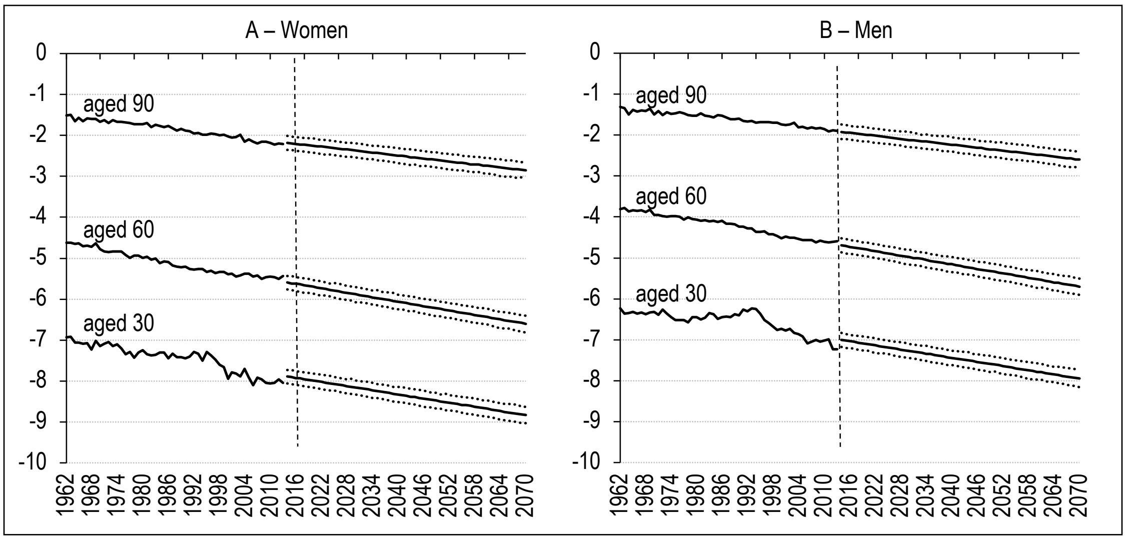
the first few years of projections before rapidly stabilising at 79,000 (Figure IX). The confidence interval also remains constant over time: at a probability of 95%, net migration will remain at between 29,000 and 129,000 each year. This amplitude is due to the significant fluctuations observed in the past and slightly exceeds the minimum and maximum observed in 1996 and 2006 respectively.

4.2. Mortality Projections: Little Uncertainty Given Past Developments

The model for mortality predicts that age-specific mortality rates will continue to decline in a linear manner, following the same trend for both males and females (Figure X). The uncertainty in the projected mortality rates does not increase over time. This is because the variance in the level and trend errors v and ω is very small compared with the variance in the error term η. Errors therefore do not accumulate over time. This is due to the fact that the trends observed are highly linear.

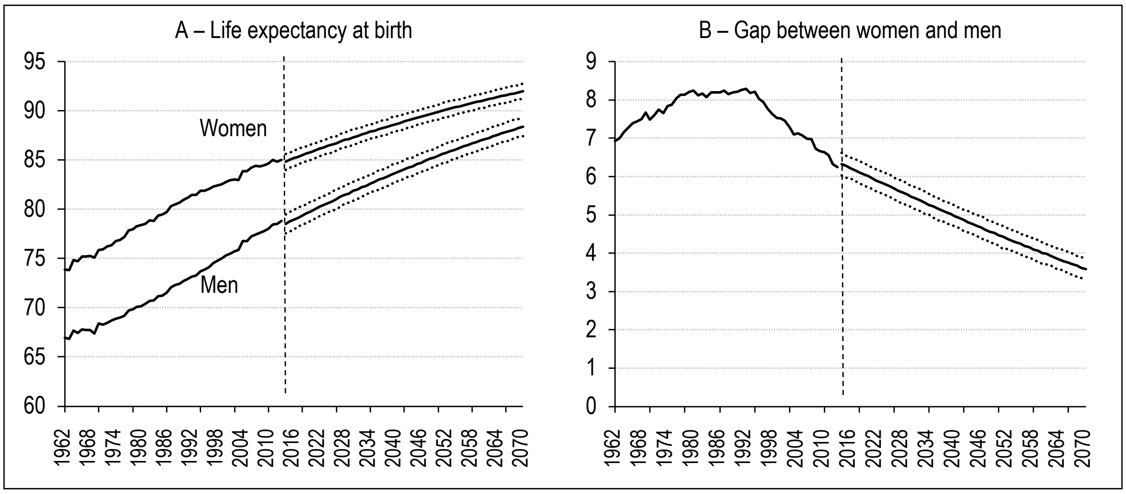
Due to the constant reduction in mortality rates, life expectancy will continue to increase in the coming years for men and women alike. The results of the model indicate that, with a probability of 95%, life expectancy at birth in 2070 will be between 91.2 and 92.8 years for women and between 87.4 and 89.4 years for men (Figure XI). The gap in life expectancy between

Figure X – **Changes in the logarithm of age-specific mortality rates, estimated and projected**



Note: The dotted lines indicate the 2.5% and 97.5% quantiles and the solid line indicates the median of the *a posteriori* distributions.
Sources and coverage: Insee, Metropolitan France. Author's calculations.

Figure XI – **Estimated and projected changes in women's and men's life expectancy and gender gap in life expectancy**



Note: The dotted lines indicate the 2.5% and 97.5% quantiles and the solid line indicates the median of the *a posteriori* distributions.
Sources and coverage: Insee, population estimates and civil registry statistics (1962-2013), author's calculations (2013-2070). Metropolitan France.

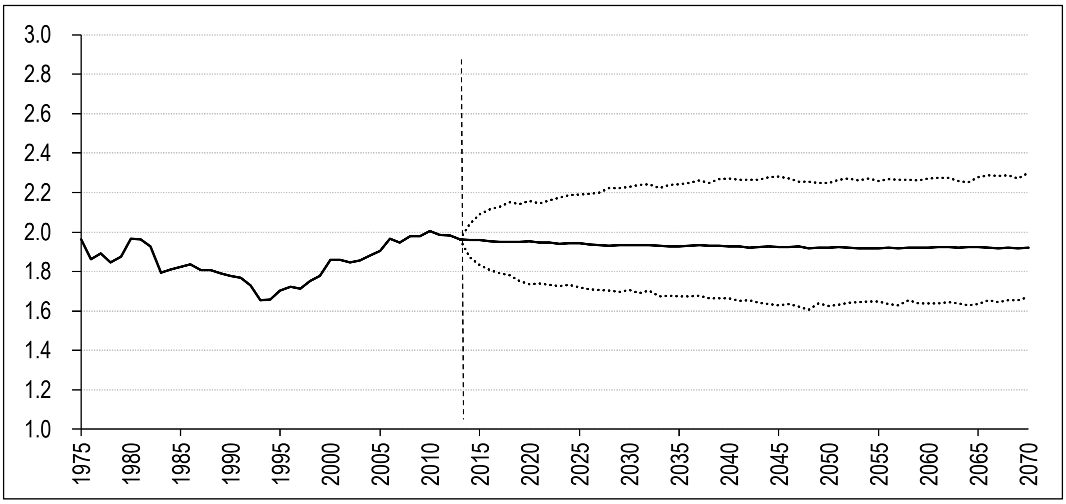
women and men will likely continue to narrow to reach 3.6 years in 2070 (between 3.3 and 3.9 years with a probability of 95%).

4.3. Fertility Projections: Births to Older Mothers and More Symmetrically Distributed Around the Modal Age

The median long-term TFR is 1.93, slightly below the mean of the *a priori* distribution, which is set at 1.95 (Figure XII). According to the model used, the TFR will be between 1.63 and 2.26 children per woman in 2070 at a probability of 95%. Unlike the projections for net migration and mortality rates, the confidence interval at 95% becomes wider over time. The uncertainty with regard to future fertility therefore becomes higher, in spite of having set a long-term TFR in the model.

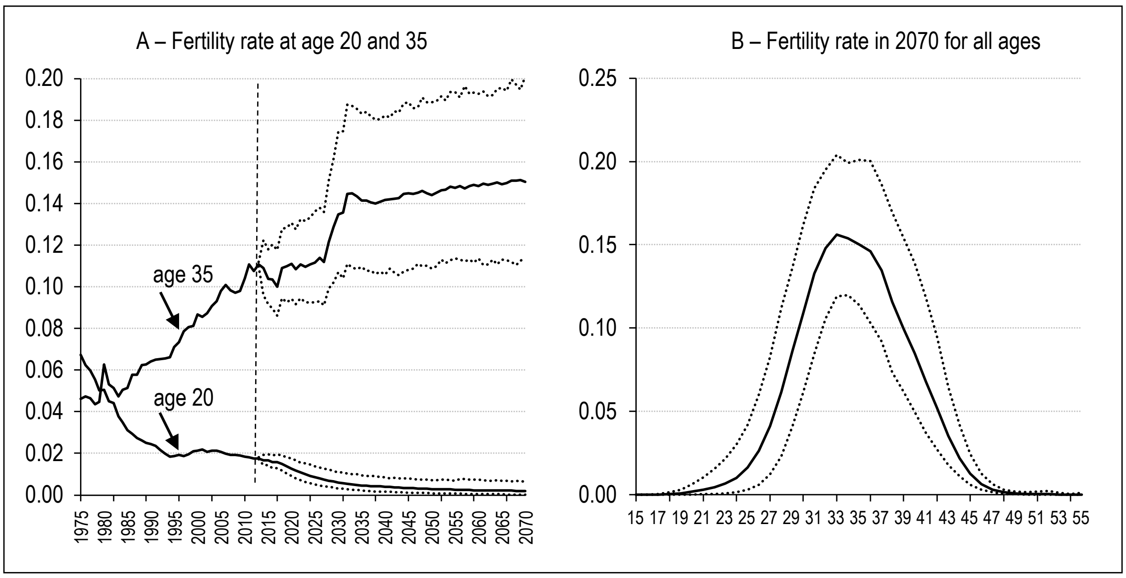
The age-specific fertility rates begin to stabilise from 2050 onwards (Figure XIII). The average age at childbirth rises rather quickly until around 2040, after which the increase continues but at a slower and slower rate until it reaches a value of between 32.2 and 35.9 years in 2070 (confidence interval of 95%). The distribution of age-specific fertility rates therefore shifts more and more to the right and becomes increasingly symmetrical, as evidenced by the changes in the measure of skewness, the median of which is tending towards 0 (Figure XIV).

Figure XII – **Changes in the total fertility rate, estimated and projected**



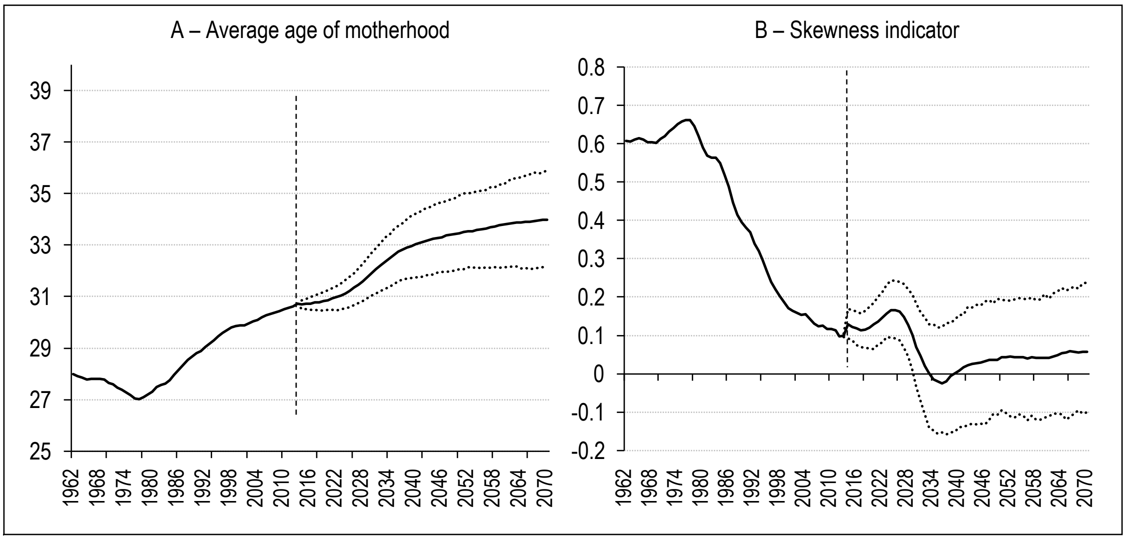
Note: The dotted lines indicate the 2.5% and 97.5% quantiles and the solid line indicates the median of the *a posteriori* distributions.
Sources and coverage: Insee, population estimates and civil registry statistics (1962-2013), author's calculations (2013-2070); Metropolitan France.

Figure XIII – **Changes in the fertility rates, estimated and projected**



Note: The dotted lines indicate the 2.5% and 97.5% quantiles and the solid line indicates the median of the *a posteriori* distributions.
Sources and coverage: Insee, Metropolitan France. Author's calculations.

Figure XIV – **Changes in the average age of motherhood and skewness of the age-specific distribution of fertility rates**



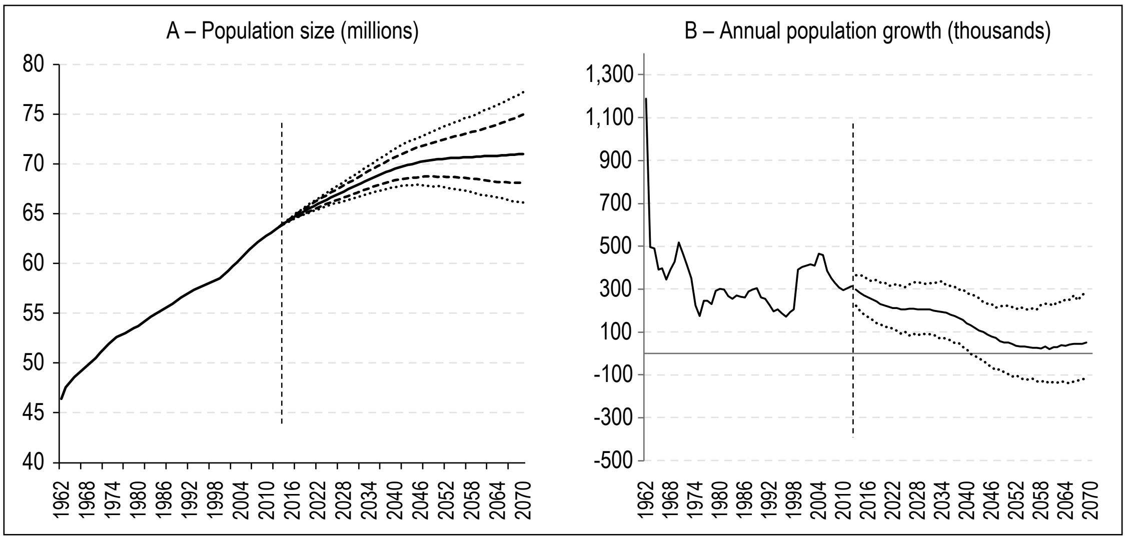
Note: The dotted lines indicate the 2.5% and 97.5% quantiles and the solid line indicates the median of the *a posteriori* distributions.
Sources and coverage: Insee, population estimates and civil registry statistics (1962-2013), author's calculations (2013-2070); Metropolitan France.

4.4. Total Population Projections: Growth Likely to Be Strong Until 2040 and Much Weaker Thereafter

The total population of metropolitan France will continue to grow until it reaches a level of between 66.1 million and 77.2 million in 2070 with a probability of 95%, and between 68.1 million and 75.0 million with a probability of 80% (see Figure XV). The median projection corresponds to a level of 71.0 million inhabitants in 2070. The population of metropolitan France could therefore increase continuously throughout the next fifty years, or it could increase before beginning to decline around 2050. According to the model used here, there is a 1% probability that the population will start to decrease from 2040 onwards (i.e. the population will reach its peak in 2040) and a 19% probability that this will occur in 2050. The uncertainty regarding the size of the population according to the model used is relatively minor until around 2040-2050, after which it increases more rapidly in the years that follow.

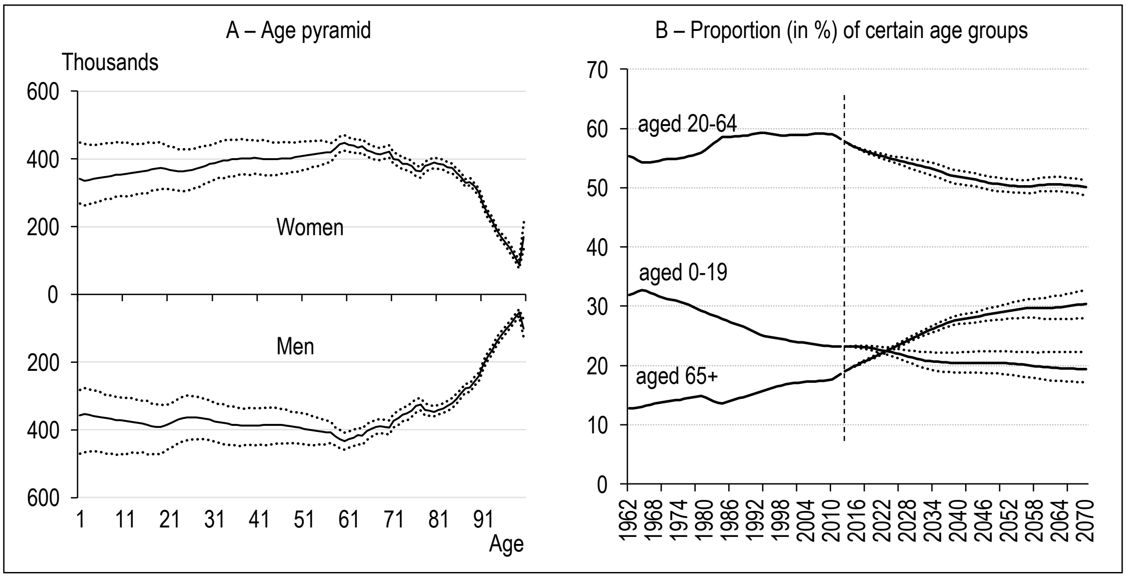
The structure of the population will also change, as can be seen in the population pyramid for 2070, the base of which is much straighter and thinner than the pyramid depicting current ages. The proportions of certain age groups will therefore decrease, particularly the youngest

Figure XV – **Past and projected changes in the total size of the population and annual population growth**



Note: The dotted lines indicate the 2.5% and 97.5% quantiles, the dashed lines indicate the 10% and 90% quantiles and the solid line indicates the median of the *a posteriori* distributions.
Sources and coverage: Insee population estimates and civil registry statistics (1962-2013), author's calculations (2013-2070); Metropolitan France.

Figure XVI – **Age pyramid for 2070 and changes in the proportion of certain age groups**

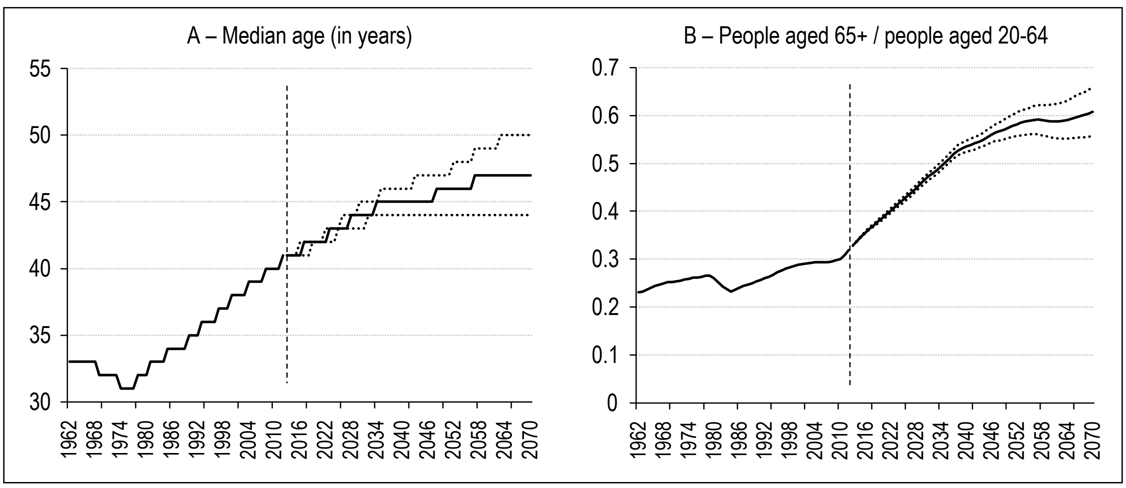


Note: The dotted lines indicate the 2.5% and 97.5% quantiles and the solid line indicates the median of the *a posteriori* distributions.
Sources and coverage: Insee population estimates and civil registry statistics (1962-2013), author's calculations (2013-2070); Metropolitan France.

(see Figure XVI): the proportion of people aged 0-19 years will continue to decrease slowly until it reaches a median level of 19% in 2070; 20-64-year-olds will follow the same pattern, with a median level of 50% in 2070. Conversely, the proportion of the population aged 65 and over will probably continue to increase until it exceeds the share of people aged under 20 in 2070. This figure increased from 13% in 1962 to 19% in 2013 and has the potential, with a probability of 95%, of making up between 28% and 33% of the population in 2070.

The population will therefore continue to age. The median age of the population, which was 41 years in 2013, could, with a probability of 95%, be between 44 and 50 years in 2070. As a result, the ratio of people aged 65 and over to people aged 20 to 64 years is likely to rise sharply in the coming years. The rapid and linear increase in this ratio between now and the early 2040s is largely due to the ageing of the large generations born during the baby boom. In fact, people born at the start of the baby boom in 1946 turned 65 in 2011 and those born at the end of the baby boom in 1975 will turn 65 in 2040. According to the models used, the ratio of those aged 65 and over to those aged 20-64 years, which today stands at 0.33, will reach a value of between 0.56 and 0.67 in 2070 with a probability of 95% (see Figure XVII).

Figure XVII – **Changes in the median age of the population and the ratio of people aged 65 and over to people aged 20-64 years**



Note: The dotted lines indicate the 2.5% and 97.5% quantiles and the solid line indicates the median of the *a posteriori* distributions.
Sources and coverage: Insee population estimates and civil registry statistics (1962-2013), author's calculations (2013-2070); Metropolitan France.

These probabilistic projections can be compared with the deterministic projections made by Insee. The deterministic projections concerning metropolitan France only cover the period from 2013 to 2050.[4] The central scenario adopted leads to a population that is slightly larger than the median of our probabilistic projections: according to the first projection, the population of metropolitan France would reach 71.7 million inhabitants in 2050 and 70.5 million according to the second. Furthermore, the confidence interval estimated by the probabilistic projections is much lower than the interval between the high and low population scenarios, which are the extreme scenarios of the deterministic projections. The difference between the two extreme deterministic scenarios is 11.1 million inhabitants in 2050, whereas the confidence interval of the deterministic projections for that same year is 5.7 million for the 95% confidence interval and 3.6 million for the 80% confidence interval.

4.5. Discussion

According to the models described in this article and the simulations carried out, the population of metropolitan France is expected to continue increasing in the coming decades. However, there is a non-negligible probability that it will start to decline before 2070, although this is less likely than an increase or stabilisation. The structure of this population is also likely to change: a general ageing of the population is expected due to increased life expectancy, a stagnating trend in the TFR and the continued arrival of baby boomers at retirement ages. The model used to project net migration is the simplest of the three models used. The lack of age-specific data on people entering and leaving the country precludes the use of Poisson modelling to obtain the rates, as we did for the number of deaths and the number of births. In general, models for projecting net migration are less sophisticated and have been the subject of less research effort than those for mortality and fertility, since the available data are less rich. Nevertheless, it is worth noting that some countries, such as New Zealand in particular, which have detailed data concerning people entering and leaving the country, are starting to offer advanced modelling of migration phenomena, taking account of a large number of parameters, such as the level of education attained by the population (Bryant & Zhang, 2014). Since our modelling is fairly simple, it follows that most of the past changes in net migration are considered noise. Since this noise is then propagated into the future, the confidence intervals of the projected net migration are very wide and therefore reflect our level of uncertainty about the future evolution of migration. This is why we have restricted the estimation of the parameters of the model (and therefore the variance of the error term in particular) to the period 1995-2013, to ensure that we do not take account of large fluctuations in the migratory balance that are too old. Estimating the model over a longer period would have led to even greater uncertainty about the future development of net migration.

Unlike net migration, mortality trends are very stable and the model used is able to take account of these trends without considering them to be predominantly noise. As a result, the confidence intervals of projected mortality rates and life expectancy are very small. This may seem misleading, as one could be led to believe that we are almost certain of what will happen. In reality, it is important to remember that the confidence intervals on future mortality levels are conditionally determined by the model taking the correct approach to reality. Indeed, such levels of confidence can only be attained for future mortality rates by assuming that the observed trends will continue. In spite of this, the model used does not take account of certain peculiarities of mortality in France. Firstly, it does not allow gender-specific changes in the logarithm of the mortality rate to be projected at a given age. Furthermore, it appears that generations born after the Second World War have very little gain in terms of mortality at a given age when compared with previous generations, regardless of the age in question (Blanpain & Buisson, 2016a). The model used does not allow such generation effects to be taken into account: deviations from the general trend are therefore treated as noise and included in the error terms rather than being seen as a well-identified effect. The resulting projected life expectancies are therefore somewhat lower than those obtained by the projections made by Blanpain & Buisson (2016a).

Fertility is modelled differently from net migration and mortality. In fact, unlike mortality rates, fertility rates do not evolve in a regular manner over time. They can increase and then decrease or vice-versa, and therefore intersect. Extending fertility rates in accordance with linear trends also leads to situations that appear implausible in the light of other fertility indicators, such as the TFR and the peak fertility rate attained during the year, which have remained more or less stable

4. see *https://www.insee.fr/fr/statistiques/2859843*

since 1975. The idea was to initially extend the TFR, which is an indicator that reflects the level of fertility, using the same method as was used to project net migration. We then extended the age-specific fertility rates in accordance with the method described by Wiśniowski *et al.* (2015), and we modified these rates to bring them back down to the TFR projected in the first place. This provides a fairly realistic trend of age-specific fertility rates, with the distribution shifting towards older ages and becoming more symmetrical. This approach (projecting an aggregated indicator and then breaking it down into detailed categories) is not new and is also the approach adopted by the UN. The disadvantage is that a long-term TRF must be set and the level chosen obviously affects the results.

* *
*

Probabilistic population projections provide new insights into possible population change. They make it possible, under certain modelling assumptions, to quantify the level of uncertainty concerning the future development of demographic indicators and in particular the evolution of total population size. They therefore offer a clear advantage over deterministic projections based on scenarios for which the probability of occurrence is not quantified. Any demographic indicator, whether it be life expectancy, the average age of motherhood or the proportion of people aged 65 and over, can be determined with some degree of probability. One of the potential difficulties in interpreting the results stems from the fact that one should not think in terms of a single point, but rather in terms of the probability distribution, just as a dice cannot be defined by just one of its six sides, even if it is loaded. Instead, it is by giving the probability that each number will appear that we will have a good description of the dice and what we can expect when it is rolled. Once this difficulty has been overcome, the interpretation and use of probabilistic population projections offers a great deal of freedom and flexibility. Conversely, the results of deterministic projections become complicated to use and disseminate when the number of scenarios under consideration is multiplied by the effect of several hypotheses intersecting.

There are a number of ways in which the methods used in this article, and therefore the results, can be improved. The first step is to better understand the migration phenomena by performing a detailed analysis of persons entering the country. It would also be interesting to look more thoroughly at estimates of flows of persons leaving the country, both now and in the past, which are relatively new in France, taking account of the available data. When projecting mortality, it would be useful to incorporate a generation effect and to allow mortality rates to develop differently for women and men. Several models are possible for this; however, the difficulty still remains that if there are too many parameters, there is a high risk that the model will not be identified or that the convergence of the Markov chains used to estimate the *a posteriori* distributions will be poor. In order to improve the projection of age-specific fertility, one could, as has already been done in several studies, find a parametric model of the distribution of age-specific rates. Although it would not necessarily be easy, it would then suffice to extend these parameters, as in the case of Lee-Carter modelling, by detecting regularities and trends in the development of these parameters. Beta distribution is a possible model, but its rounded shape would not represent the data well. Gamma distribution would better reflect the distribution of fertility rates, but it is defined on a support that is open to the right. It must therefore be truncated to ensure that there are no unrealistic results. The Hadwiger number presents a third option, as it seems better suited to modelling the distribution of fertility. The downside is that it can take a long time to estimate its parameters and their interpretation is not necessarily obvious. So why not propose an *ad hoc* function that faithfully reflects the observed data? It could be tempting to estimate the distribution of fertility rates in a non-parametric way, i.e. in reality by using a very large number of parameters. The difficulty then lies in the projection of these very many parameters. We could also consider developing structural models for the three components of population change that would make it possible to explain past change according to more detailed mechanisms and based on external variables; however, this would also require to have a sufficient number of elements to allow us to project the evolution of the variables. It would also be very informative to conduct sensitivity analyses, which would allow to test how the results vary when certain assumptions in the models are changed slightly. This would help to better understand and quantify the precise role of each component in population change.

As can be seen, there is undoubtedly much room for improvement, and this will require significant investments in research into understanding and modelling migration, mortality and fertility. This would only be beneficial for probabilistic population projections, the degree of uncertainty of which depends above all on our knowledge (or ignorance) of these topics. Finally, it is important that we do not compare probabilistic population projections with deterministic population projections. The latter remain extremely useful and allow us to test what would happen in the future in a given scenario. The general conclusions reached are also very consistent with those reached via deterministic projections of changes in population size and age structure. However, it is primarily up to the users of population projections to choose the approach that best suits them, depending on what they are using them for. Probabilistic and deterministic projections are two different ways of tackling uncertainty and trying to shed light on the future. □

BIBLIOGRAPHY

Alkema, L., Raftery, A. E., Gerland, P., Clark, S. J., Pelletier, F. & Buettner, T. (2010). Probabilistic projections of the total fertility rate for all countries. Center for statistics and the social sciences, University of Washington. *Working paper* N° 97. https://doi.org/10.1007/s13524-011-0040-5

Bellamy, V. & Beaumel, C. (2016). Bilan démographique 2015, le nombre de décès au plus haut depuis l'après-guerre. *Insee Première* N° 1581.
https://www.insee.fr/fr/statistiques/fichier/version-html/1908103/ip1581.pdf

Bijak, J., Alberts, I., Alho, J., Bryant, J., Buettner, T., Falkingham, J., Forster, J. J., Gerland, G., King, T., Onorante, L., Keilman, N., O'Hagan, A., Owens, D., Raftery, A. & Ševčíková, H. (2015). Probabilistic population forecasts for informed decision-making. *Journal of Official Statistics*, 31(4), 537–544.
https://doi.org/10.1515/jos-2015-0033

Bijak, J. & Bryant, J. (2016). Bayesian demography 250 years after Bayes. *Population Studies*, 70(1), 1–19.
https://doi.org/10.1080/00324728.2015.1122826

Blanpain, N. (2016). Les hommes cadres vivent toujours 6 ans de plus que les hommes ouvriers. *Insee Première* N° 1584. https://www.insee.fr/fr/statistiques/fichier/version-html/1908110/ip1584.pdf

Blanpain, N. (2020). Is the Ageing of the French Population Unavoidable? *Economie et Statistique / Economics and Statistics*, this issue.

Blanpain, N. & Buisson, G. (2016a). Projections de population 2013-2070 pour la France : méthode et principaux résultats. Insee, *Document de travail* N° F1606.
https://www.insee.fr/fr/statistiques/fichier/2400057/F1606_version_revue_septembre2018.pdf

Blanpain, N. & Buisson, G. (2016b). Projections de population à l'horizon 2070. *Insee Première* N° 1619.
https://www.insee.fr/fr/statistiques/fichier/version-html/2496228/ip1619.pdf

Booth, H. (2006). Demographic forecasting: 1980 to 2005 in review. Australian National University, *Working Papers in Demography* N° 100.
https://demography.cass.anu.edu.au/sites/default/files/sod/publications/working-papers/100.pdf

Brutel, C. (2014). Estimer les flux d'entrées sur le territoire à partir des enquêtes annuelles de recensement. Insee, *Document de travail* N° F1403. https://www.insee.fr/fr/statistiques/fichier/1381143/F1403.pdf

Bryant, J. & Zhang, J. L. (2014). Bayesian forecasting of demographics rates for small areas: emigration rates by age, sex and region in New Zeland. *Statistica Sinica Preprint* N° SS-14-200tR3.
https://www.jstor.org/stable/44114337

Conseil d'orientation des retraites (2017). Évolutions et perspectives des retraites en France. Rapport annuel du COR. https://www.cor-retraites.fr/sites/default/files/2019-06/doc-4055.pdf

Costemalle, V. (2015). Projections de populations : l'ONU adopte une méthode bayésienne. *Statistique et Société*, 3(3), 9–14. http://statistique-et-societe.fr/article/view/500/474

Dunstan, K. & Ball, C. (2016). Demographic projections: user and producer experiences of adopting a stochastic approach. *Journal of Officials Statistics*, 32(4), 947–962. https://doi.org/10.1515/jos-2016-0050

Gerland, P., Raftery, A. E., Ševčíková, H., Li, N., Gu, D., Spoorenberg, T., Alkema, L., Fosdick, B. K., Chunn, J., Lalic, N., Bay, G., Buettner, T., Heilig, G. K. & Wilmoth, J. (2014). World population stabilization unlikely this century. *Science*, 346(6206), 234–237. https://doi.org/10.1126/science.1257469

Hyndman, R. J. & Booth, H. (2006). Stochastic population forecasts using functional data models for mortality, fertility and migration. Australian National University, *Working Papers in Demography* N° 99. https://demography.cass.anu.edu.au/sites/default/files/sod/publications/working-papers/99.pdf

Hyndman, R. J. & Ullah, S. (2007). Robust forecasting of mortality and fertility rates: a functional data approach. *Computational Statistics & Data Analysis*, 51(10), 4942–4956. https://doi.org/10.1016/j.csda.2006.07.028

Kontis, V., Bennett, J. E., Mathers, C. D., Li, G., Foreman, K. & Ezzati, M. (2017). Future life expectancy in 35 industrialised countries: projections with a Bayesian model ensemble. *The Lancet*, 389(10076), 1323–1335. https://doi.org/10.1016/S0140-6736(16)32381-9

Lee, R. D. & Carter, L. R. (1992). Modeling and forecasting U.S. mortality. *Journal of the American Statistical Association*, 87(419), 659–671. https://doi.org/10.2307/2290201

Lutz, W., Sanderson, W. & Scherbov, S. (2001). The end of world population growth. *Nature*, 412, 543–545. https://doi.org/10.1038/35087589

MacPherson, L. (2016). National population projections : 2016(base)-2068. Statistics New Zeland. https://www.stats.govt.nz/information-releases/national-population-projections-2016base2068

Raftery, A. E., Chunn, J. L., Gerland, P. & Ševčíková, H. (2013). Bayesian probabilistic projections of life expectancy for all countries. *Demography*, 50, 777–801. https://doi.org/10.1007/s13524-012-0193-x

Raftery, A. E. (2014). Use and communication of probabilistic forecasts. University of Washington. arXiv:1408.4812v1. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5403155/

Régnier-Loilier, A. & Vignoli, D. (2011). Intentions de fécondité et obstables à leur réalisation en France et en Italie. *Population-F*, 66(2), 401–432. https://doi.org/10.3917/popu.1102.0401

United Nations, Population Division (2017). World Population Prospects: The 2017 Revision, key findings and advance tables. *Working Paper* N° ESA/P/WP/248.

Wiśniowski, A., Smith, P. W. F., Bijak, J., Raymer, J. & Forster, J. J. (2015). Bayesian population forecasting: extending the Lee-Carter method. *Demography*, 52, 1035–1059. https://doi.org/10.1007/s13524-015-0389-y