Research Material

Computations of French lifetables by department, 1901–2014

Florian Bonnet

Contents

Computations of French lifetables by department, 1901–2014

Florian Bonnet[1]

Abstract

BACKGROUND

Debates concerning the French territorial divide are deep in France. However, few historical data on French demography at the local level for the 20th century are available for the study of these territorial disputes.

OBJECTIVE

The main objective is to present a new demographic database built at the departmental level for the 20th century, as well as the methods used to build this database.

METHODS

The database was constructed by digitizing a large part of the raw data available in the French archives and by using new statistical sources for military deaths and deportees during the two world wars. The methods used are both the methodological protocol of the Human Mortality Database and a set of methods that take into account the specificities of the French departments.

CONTRIBUTION

With this database, the French departmental lifetables by sex and year, including military deaths and deaths in deportation, will be available, as well as the departmental populations by age, sex, and year between 1901 and 2014.[2]

---

[1] École Normale Supérieure Paris-Saclay, Paris, France. Email: florian.bonnet@ens-paris-saclay.fr.
[2] These data are available from the author. A dedicated website called "French Regional Database" will be created in the second half of the year 2020. The data will be freely downloadable from this site.

1. Introduction

Life expectancy has risen sharply in France since the beginning of the 20th century. The lifetables calculated by Vallin and Meslé (2001) for the 19th and 20th centuries show that male life expectancy at birth was 33 in 1806, 44.5 in 1898, 60 in 1946, and 74.5 in 1997. This increase at the national level does not say anything about increases at the local level, and significant differences exist among the French departments. Barbieri (2013) works on departmental mortality and shows that the life expectancy at birth of men for the period 2006–2008 was 74.4 years in Nord, compared to 79.7 years in Hauts-de-Seine, a difference of more than 5 years. This situation explains why the debate on the territorial divide according to health is important in France. Indeed, departmental differences cannot be explained from a public policy point of view: The state has to reduce these inequalities. To inform public decision-makers in their choices, it is important to know the history of these departmental differences.

Consequently, I compute and present in this paper the yearly departmental lifetables by sex for all French metropolitan departments between 1901 and 2014.[3] The computation of these lifetables is based on the exhaustive collection of both population flows (deaths by age and sex, births by sex) and population stocks (population by age and sex) at each census. I exploit a unique French characteristic: Since 1789, this country has been divided into around 100 geographical units of comparable size, namely departments. This division has changed very little during two centuries, and the statistical centralizations have been carried out at this geographical level. Moreover, to take into account the two wars that affected France between 1914 and 1918 and 1939 and 1945, I have collected in two original sources the military deaths by age during the two wars as well as the deaths in deportation by age and sex during the Second World War. With these lifetables, I get life expectancies and mortality rates at each age for more than 100 years. In addition, I get populations by age and sex at each January 1 of each year.

These lifetables at the subnational level complete a still incomplete literature. Bonneuil (1997) works on departmental mortality in the 19th century: He computes lifetables for women by five-year period and for five-year age groups. He follows Van de Walle (1974), who computes similar lifetables with a different methodology. These two authors have not studied men's mortality in the same way because of strong fluctuations due to the wars. Daguet (2006) groups lifetables, established at the department level, for 1954 to 1999, but only for the census years. Barbieri (2013) uses in her study departmental lifetables calculated by the National Institute of Statistics and Economic Studies (IN-SEE) for the period 1975–2008. However, the data were provided exceptionally. Vallin and Meslé (2005) use departmental life expectancies for the period 1906–1954. However, both reconstruction methods and data have never been published. Lastly, various

---

[3] These data will be yearly updated in order to include the most recent raw data.

mortality indicators are available in official publications, namely *Statistique Annuelle du Mouvement de la Population*. However, these indicators are relatively scarce: They relate only to infant mortality rates or standardized mortality rates.

In addition, the lifetables I compute are based on a unified methodological protocol for the whole period 1901–2014, which is not the case with the papers previously cited. This methodological protocol is available in Wilwoth et al. (2007). Many researchers are using this protocol to compute national lifetables for a large number of countries. It is also used to compute lifetables at the local level in four OECD countries. The results according to Canadian provinces for the period 1921-2011 are available in the Canadian Human Mortality Database; those according to the Japanese provinces since 1975 are available in the Japan Mortality Database; those according to States in USA since 1959 are available in the US Mortality Database; those according to australian territories since 1971 are available in the Australian Human Mortality Database. This paper therefore complements a still limited supply of local mortality data freely available by adopting an internationally recognized protocol; this allows international comparisons without methodological bias.

The rest of this paper is organized as follows. Section 2 presents the statistical sources used to compute lifetables. Section 3 details the methods used; I distinguish the methods coming from the Human Mortality Database (HMD) protocol and the methods specific to this study. Section 4 illustrates some of the results available in this new database. Section 5 concludes.

2. Sources

Computation of departmental lifetables requires two types of raw statistics: population movement (deaths and births domiciled) and population censuses. The deaths collected are not just civilian deaths: Both military deaths during the two world wars and deportation deaths between 1939 and 1945 have been included. Online Appendix A-8 gives detailed sources in which raw statistics have been found.

2.1 Deaths

Civilian deaths of each department, sex, and year over the period 1901–2014 have been retrieved from the statistics published by Statistique Générale de la France (SGF) and then by INSEE. I have retrieved deaths by age group recorded in home departments. In addition, I have collected in Vallin and Meslé (2001) single-age and sex-specific civilian deaths at the national level for the same period.

I have retrieved deaths during the two world wars from the Defense Ministry's web-

site.[4] They are available by birth year at the departmental level and by birth year and death year at the national level.

Individuals who died during deportation in the Second World War are not included in the civilian population movement. However, they numbered nearly 100,000. I have decided to include them in my lifetables, using data from MemorialGenWeb website.[5] This database records deaths published in the Journal Officiel of those who left France and died in deportation. They are listed by birth department if they were born in France and by birth country otherwise. Among the foreign-born deportees, the Poles were the most numerous (around 40%), followed by the Spanish (around 15%). Although this database is not exhaustive, the large number of observations provides a sample close to the total of deportees.

2.2 Births

I have retrieved births by year, sex, and mother's home department for the period 1901–2014 and the period 1853–1900. I have also recovered stillbirths by mother's home department and year without distinction of sex.

2.3 Censuses

I have collected populations by birth year, home department, and sex for each census of the period 1901–1962 from hard-copy publications of SGF and INSEE. For the period 1968–2014, these statistics have been found in online sources. These populations are not available for each year because censuses were held at varying intervals. Between 1901 and 2014, censuses were made in 1901, 1906, 1911, 1921, 1926, 1931, 1936, 1946, 1954, 1962, 1968, 1975, 1982, 1990, 1999, 2008, 2013, and 2014.

3. Methods

The protocol I use to compute departmental lifetables is largely inspired by the one of the Human Mortality Database. This database gathers all national lifetables computed using these methods. However, since my database is specific for both the small numbers in each department and the time period chosen (including the two world wars), I have added specific methods.

---

[4] www.memoiredeshommes.sga.defense.gouv.fr, downloaded in February 2016.
[5] www.memorialgenweb.org/memorial3/deportes, accessed in March 2016.

3.1 HMD protocol methods

3.1.1 Raw data adjustments

Raw data adjustments according to deaths are the main issue, since they are aggregated into five-year age groups until 1967 and by single age between 1968 and 2014. To get a 1×1 format (single age, year of death) for the deaths between 1901 and 1967, I distribute deaths at unknown age among age groups and adjust the curve of cumulative deaths by cubic splines. Nevertheless, deaths estimated by cubic splines are too imprecise to be used at advanced ages since open-age interval of deaths is too low. These deaths are then adjusted by means of the Kannisto model, which assumes a survival curve of logistic form, with a zero-asymptote for very old ages. These methods are presented in an extensive way in Appendix A-1. I finally proceed to a uniform adjustment so that the sum of the departmental deaths for each age, year, and sex corresponds to the national data.

For censuses, raw statistics are generally available for five-year age groups. I use the cubic splines method in the same way to estimate populations according to birth year.

### 3.1.2 Splitting deaths into Lexis triangles

Deaths by year and age may be split into two triangles for a single year, known as Lexis triangles. Overall, if the probability of death is equiprobable over time, one might think that the distribution of annual deaths by age for half in the lower triangle and the other half in the upper triangle would be sufficient. This is not the case, for two main reasons. The first is that infant mortality, when high, is observed largely in the first days after birth and must therefore be integrated into the lower triangle. The second concerns the relative size of cohorts, which also influences the distribution between triangles. When the flow of births varies greatly from one year to the next (e.g., during the two world wars), the half-death distribution in the lower triangle is strongly biased. The HMD protocol sets a sex-specific equation allowing the distribution of deaths in Lexis triangles. I use these sex-specific equations, presented in Appendix A-1, in the case of French departments.

Note that I obtain proportions of deaths in the lower triangle greater than 1 for 28 female orbservations and for 30 male observations, all in 1918 or 1919 and for deaths under age 1. This is due to the Spanish influenza epidemic, the high infant mortality rate, and the differences in size between the cohorts born in 1918 and 1919. To tackle this issue, the death proportions in the lower triangle are set at 1, leading to zero deaths in the upper triangle for these observations.

3.1.3 Computations of populations by age at January 1 of each year

To calculate the mortality rates, I need populations by age at January 1 for each year from 1901 to 2014. I get populations by age in 2014 from official statistics, and I calculate populations by age for the 1901–2013 period. Figure 1 reveals the four methods used for various periods and ages. Online Appendix A-2 precisely specifies each method used.

Figure 1:   **Methods for computations of population at January 1**
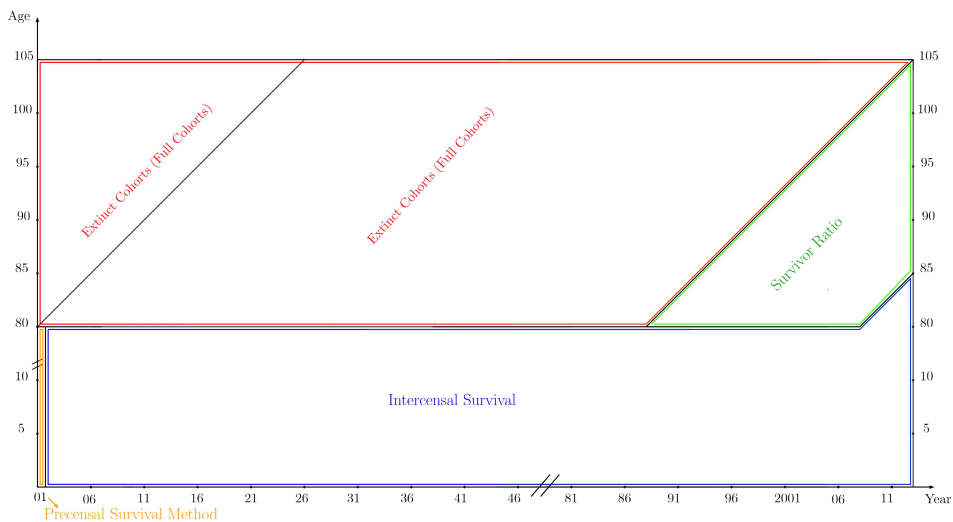
Note: Methods used to compute populations by age at each January 1. For more details, see online Appendix A-2.

First, the intercensal survival method is used to estimate the population under age 80 from 1902 to 2013. Starting from one census (say, 1901) the population by age at the following census (1906) is estimated by subtracting from the population by age in 1901 the deaths that occurred from 1901 to 1906. The difference between estimated and recorded populations in 1906, due to measurement errors and migrations, is then linearly attributed to the intercensal population figures.

Second, the precensal survival method is used to estimate the population under age 80 in 1901. Since intercensal survival estimates the population under age 80 on January 1 of each intercensal year, I cannot compute population by age on January 1 the year of the first census. To do so, I use the population on the day of the 1901 census and add the deaths occurring between January 1, 1901 and the census day. Since there is no second census available, as with intercensal survival method, I cannot correct for migrations and errors. I assume that in so short a period, these are minimal.

Third, the extinct cohorts method is used to estimate the population aged 80 and over born in the cohorts that died between 1901 and 2013. I assume that migrations after age 80 are small. I compute the population of a cohort still alive by summing its future observed deaths.

Fourth, I estimate the population aged 85 and over in 2013 with the survival ratio method. I assume that the survival ratio between two ages for the extinct cohorts can be applied to the still living cohorts to estimate their size at the last census. The 85-and-over estimated are then adjusted by the 85-and-over recorded in 2013. After this adjustment I compute the size of the intermediate populations located in the green quadrilateral of Figure 1 by substracting step by step the observed deaths.

3.1.4 Adjustment of computed mortality rates

I can compute departmental mortality rates by age and sex with deaths in Lexis triangles and populations at each January 1. Mortality rates are the ratios between the number of deaths and the number of individuals exposed to the risk:

$$M_{xt} = \frac{D_{xt}}{E_{xt}} = \frac{D_L(x,t) + D_U(x,t)}{\frac{1}{2}\left[P(x,t) + P(x,t+1)\right] + \frac{1}{6}\left[D_L(x,t) - D_U(x,t)\right]}. \tag{1}$$

Note that I do not calculate populations for 2015, although these are needed for 2014. To estimate lifetables for that year, I assume that the population at each age in 2015 is equal to the one that in 2014, and the formula becomes:

$$M_{x2014} = \frac{D_{x2014}}{E_{x2014}} = \frac{D_L(x,2014) + D_U(x,2014)}{P(x,2014) + \frac{1}{6}\left[D_L(x,2014) - D_U(x,2014)\right]}. \tag{2}$$

These rates are not used directly to calculate lifetables. I smooth mortality rates beyond age 90 to avoid erratic fluctuations due to small numbers of deaths and population at risk. The instantaneous probability of dying over age 80 in the Kannisto model can be expressed as follows (with a and $b \geq 0$):

$$\mu_x(a, b) = \frac{ae^{b(x-80)}}{1 + ae^{b(x-80)}}.$$
(3)

Mortality rates estimated with the Kannisto model $M_x(a, b)$ are:

$$M_x(a, b) = \mu_{x+0,5}(a, b).$$
(4)

If $D_x \sim Poisson\left(E_x \mu_{x+0,5}(a, b)\right)$, then parameters a and b may be calculated by minimizing the following function:

$$-logL(a, b) = \sum_{x=80}^{105} \left[D_x log\mu_{x+0,5}(a, b) - E_x \mu_{x+0,5}(a, b)\right].$$
(5)

I can calculate $\hat{M}_x(\hat{a}, \hat{b})$ for all ages above 90, with estimated parameters $\left(\hat{a}, \hat{b}\right)$. I assume that the population's mortality rates are equal to the mortality rates in the survival tables (m_x):

$$\begin{cases} m_x = M_x & x \in [0, 89] \\ m_x = \hat{M}_x & x \in [90, 105] \end{cases}.$$
(6)

3.1.5 Computations of lifetables

To convert the survival table mortality rates into probabilities of dying, one must define a_x, the mean number of years lived by people dying between ages x and $x + 1$. I assume that deaths are uniformly distributed at each age:

$$\begin{cases} a_x = \frac{1}{2} & x \in [1, 104] \\ a_x = \frac{1}{m_{105}^{\infty}} & x = 105+ \end{cases}.$$
(7)

For age 0, I follow Preston (2000), who relies on Coale and Demeny's (1983) lifetables. Thus:

$$\begin{cases} m_0 \geq 0.107 & \begin{cases} a_0 = 0,350 & \text{for women,} \\ a_0 = 0,330 & \text{for men,} \end{cases} \\ m_0 < 0.107 & \begin{cases} a_0 = 0,053 + 2.800 & \text{for women,} \\ a_0 = 0,045 + 2.684 & \text{for men.} \end{cases} \end{cases} \qquad (8)$$

The probabilities of death may be calculated as follows:

$$\begin{cases} q_x = \frac{m_x}{1+(1+a_x)m_x} & x \in [0,104] \\ q_x = 1 & x = 105+ \end{cases}. \qquad (9)$$

With values of q_x, I can compute each of the lifetable values, for each age. Two lifetables are estimated: in the format (1×1) and in the format (1×5) – that is, for each age and each group of five years. For the sake of readability, abridged lifetables in the (5×1) and (5×5) formats are also estimated. So I get values for age groups [0,1[, [1,5[, [5,10[, [10,15[, ... and so on until ages 105 and over. Values in abridged lifetables are computed from previous variables.

3.2 Specific departmental methods for the period 1901–2014

The methods presented previously come from the Human Mortality Database protocol. However, they are too general to be applied without correction to the case of French departments during the 20th century. These corrections are due to three main issues: the quality of the raw data, the two world wars, and the territorial changes in my departmental classification.

3.2.1 Specific methods due to deaths during the two world wars

In their reconstruction of the French lifetables, Vallin and Meslé (2001) show that infant mortality should be analyzed carefully. They distinguish three kinds of deaths among those aged 0 to 1: deaths occurring before age 1 and accounted for in official deaths, deaths occurring between birth and official birth declaration (wrong stillbirths), deaths occuring before birth (stillbirths). In order not to bias the lifetables and particularly the evolution of the infant mortality rate over time, Vallin and Meslé (2001) disentangle wrong stillbirths from stillbirths and include wrong stillbirths in the statistics of both

births and deaths aged 0 to 1. For the departmental lifetables, I distribute the national wrong stillbirths between departments according to the departmental stillbirths, on a pro rata basis. These departmental false stillbirths were added to both departmental births and departmental deaths aged 0 to 1. These computations are based on the hypothesis that the distinction between wrong stillbirths and false stillbirths made at the national level by Vallin and Meslé (2001) is correct. Moreover, these departmental lifetables do not allow us to study the issue of infant mortality including stillbirths.

Moreover, data retrieved from censuses are not of identical quality. Consequently, I make two adjustments. First, I distribute individuals of unknown birth year pro rata of the numbers in known birth years. Although this does not present any problem for most censuses, this is not true in 1901. At that date, these numbers were included with those of the open age interval. So I use the 1911 census, which is particularly detailed. Moreover, younger age groups did not always use the same variables, namely year of birth or age. I use a linear interpolation to compute figures per year of birth. Online Appendix A-3 presents these adjustments in a more detailed manner.

3.2.2 Specific methods due to the two world wars

The two world wars had significant demographic effects, both at national and departmental levels. The first effect was due to internal migrations caused by the conflict and France's division into occupied and unoccupied zones in 1940. Raw statistics give no direct indication for this question. The second effect concerns the heavy military losses, which had to be included in death statistics. On this particular point, this study is the first to integrate military and deportation deaths into lifetables at the subnational level.

For military deaths during the two world wars, I use three sources of information. The first provides the total of deaths by department and birth year. It comes from the Defense Ministry's database, which lists all the *Morts Pour La France* (MPLF) of the two world wars. The second provides the total of deaths at the national level by birth year and death year. It mobilizes the crowd-based indexing on the *Mémoire des Hommes* website: Anyone, using their personal research on a specific soldier, can inform both his birth year and death year. The third is the total of deaths as estimated by researchers at the national level (Pedroncini 1992, Prost 2008, Hubert 1931, Lagrou et al. 2002), so as to verify the overall consistency of the various sources. Online Appendix A-4 presents in an extensive manner the methods used to obtain military death by year of birth, year of death, and department of residence.

Deportees during the Second World War are classified in the database by birth place, which is different from home place. I build cross-matrices between birth place and home place for the deportees born in France and those born abroad. For that purpose I use two raw materials. The first is the 1936 census for the foreign-born, which provides their

distribution among departments in France. The second is the 1946 census for the French-born, which provides their distribution by birth place and home place at departmental level. Finally, I adjust these figures by the total of deportees estimated by researchers, namely 110,000, in line with Dupaquier (1988). Online Appendix A-4 presents the computations of deportees by age, sex, and home department.

3.2.3 Specific methods due to territorial changes and missing data

The main advantage of the French departments is their stability since the beginning of the 19th century. However, there were some changes during the two last centuries, especially with regard to the eastern borders and the Paris region. To take this into account, some adjustments are necessary. In this study, I use a departmental classification with 97 departments: the 95 departments of the current metropolitan France (with Corse counting as one), as well as Seine and Seine-et-Oise in their pre-1968 boundaries. Territorial breakdowns are twofold in this study: Either departmental boundaries changed because of a territorial reorganization, or the data are missing within the unified departmental classification that I use.

The departmental boundary changes are of two types for the period 1853–2014. The first concerns the pre-1901 period. Savoie and Nice's Comté were attached to France following the *plébiscite* of April 22–23, 1860. Savoie and Haute-Savoie were created ex nihilo on June 14, 1860, while Alpes-Maritimes was created by aggregating a part of Var (Grasse's canton) to the Comté. Moreover, following the war against Prussia in 1870, Meurthe and Moselle in their old form disappeared to form Moselle and Meurthe-et-Moselle.[6] In addition, the departments boundaries of Haut-Rhin,[7] Bas-Rhin, and Vosges[8] changed. For this period, I distributed births of the old-classification departments between the unified-classification departments. The second change concerns the 1901–2014 period. It follows the *Ile-de-France* reorganization in 1964, effective in 1968. This reorganization led to the dissolution of Seine and Seine-et-Oise. These departments were divided between Paris, Yvelines, Essonne, Hauts-de-Seine, Seine-Saint-Denis, Val-de-Marne, and Val d'Oise.

---

[6] Until 1870, two departments existed, namely Meurthe and Moselle. Their grouping fell within the same limits as Meurthe-et-Moselle and the new Moselle. The new Moselle includes the territories under German control in 1870, namely the districts of Château-Salins and Sarrebourg for the old Meurthe and Thionville, Metz, Forbach-Boulay Moselle, and Sarreguemines for the old Moselle. In contrast, the new Meurthe-et-Moselle includes the territories remained French at that time; the districts of Luneville, Nancy and Toul for the old Meurthe and the canton of Briey for the old Moselle.

[7] In 1870, Haut-Rhin in its former boundaries was divided between Haut-Rhin as we know today – which was under German control until the end of the Second World War – and Territoire de Belfort, which remains under French control.

[8] In 1870, the former cantons of Schirmeck and Saales (in Vosges) were attached to Bas-Rhin, under German control. The new boundaries of these two *départements* are those that we know nowadays.

The missing data in the unified departmental classification are also of two types. First are the missing data due to the two world wars: Aisne, Ardennes, Marne, Meurthe-et- Moselle, Meuse, Nord, Oise, Pas-de-Calais, Somme, and Vosges for the 1914–1918 period; Moselle, Bas-Rhin, and Haut-Rhin for the period 1939–1945; Corse for the period 1943-1944. The second category is departments temporarily under German control, namely Bas-Rhin, Haut-Rhin, and Moselle before 1919.

Births in the missing departments during the period 1853–1900 are first estimated. These births allow the distribution of deaths according to Lexis triangles. I consider that the changes were synchronized between missing departments and a neighboring department. For *Var* and *Alpes-Maritimes*, whose limits are stable since 1861, I use the ratio between births in 1861 and births in Bouches-du-Rhône to deduce births between 1853 and 1860; I assume that the ratio was the same during the period. I proceed in the same way for Savoie and Haute-Savoie, for which I use Ain as reference. Regarding Vosges, Territoire de Belfort, and Meurthe-et-Moselle, I used Haute-Saône as reference for the 1853–1869 period. As I know values for Meurthe, Moselle, Haut-Rhin, and Vosges (former *departments*), it was easy to deduce values for Moselle and Haut-Rhin in their current boundaries. For the 1870–1900 period, births in Moselle, Bas-Rhin, and Haut-Rhin were estimated using Haute-Saône as reference.

Data from the population movement for missing departments during the two world wars is also estimated. Even if the lifetables of these departments should be analyzed with caution, this allows an approximation of their current mortality conditions. For that, I go further than the method used for births by endogenizing the choice of the reference department. The general assumption used for the estimations of these missing data is that the neighbor departments have similar evolutions concerning their demographic variables because of their culture and their shared living conditions. For each set of missing departments and periods, I choose a geographically close department whose data are available. Table 1 gives these candidates for each set of missing departments. I then calculate a score based on the synchronicity of demographic variations over the period surrounding the missing period. From this score, a reference department is defined for each department with missing data and used to estimate these values. This method is used to estimate both births and deaths by age (sum of civilian, military, and deportation deaths). Online Appendix A-5 goes into detail about the choice of reference departments and the method used.

With the reorganization of Ile-de-France in 1968, I must differentiate the departments belonging to the old classification from those belonging to the new ones. The former departments are followed over the 1901–1968 period, and the new ones between 1968 and 2014. So I make several adjustments. The first concerns the distribution of births before 1968 among the departments of the new classification in order to distribute deaths in Lexis triangles. It is done pro rata of 1968's births. Then I estimate 1968's age populations for departments of the old classification by using the intercensal survival

method: I assume that the Ile-de-France migratory profile was the same for Seine and Seine-et-Oise. Online Appendix A-6 discusses these two adjustments.

Table 1: **Panel of candidate reference departments**

Period	Missing departments	Panel of reference departments
1914–1919	Aisne, Ardennes, Marne, Meurthe-et-Moselle, Meuse, Nord, Oise, Pas-de-Calais, Somme, Vosges	Aube, Eure, Haute-Marne, Haute-Saône, Seine-Inférieure, Seine-et-Marne, Seine-et-Oise
1939–1945	Moselle, Bas-Rhin, Haut-Rhin	Doubs, Meurthe-et-Moselle, Haute-Saône, Vosges
1943–1944	Corse	Alpes-Maritimes, Bouches-du-Rhône, Gard, Hérault, Var

Finally, computation periods vary by departments. I distinguish them according to four classes. Class 1 (C_1) concerns all departments outside Moselle, Bas-Rhin, Haut-Rhin, and Ile-de-France (except Seine-et-Marne). These 85 departments are tracked over the period 1901–2014. Computations of population on January 1 of each year are done as shown in Figure 1. Departments in class 2 (C_2) are the former Ile-de-France departments, namely Seine and Seine-et-Oise. Lifetables were estimated over the period 1901–1968. Class 3 (C_3) concerns the new Ile-de-France departments: Essonne, Hauts-de-Seine, Seine-Saint-Denis, Val-de-Marne , Val d'Oise, Paris, Yvelines. Lifetables are available for the period 1968–2014. Bas-Rhin, Haut-Rhin and Moselle are in class 4 (C_4): Lifetables are estimated between 1921 and 2014. Figures in Online Appendix A-7 draw the methods used to estimate the January 1 populations for each of these four classes. These are variants of Figure 1.

3.3 Reliability of the data and comparison with other studies

The raw data used in this study come from old statistical sources. I therefore verified that their use could be done without introducing bias in future analyses.

First, I tested whether departmental and national data were consistent. Vallin and Meslé (2001) calculated the national lifetables for the 19th and 20th centuries. Consequently, I verified that the departmental sums of deaths, births, and populations were equal to national values. These expectations were true, which testifies to the quality of the raw data. My results were therefore consistent with the results established at the national level.

Second, I test how my results are consistent with the works already done at the departmental level. To do so, I calculated the differences between my departmental life expectancies and those of Bonneuil (1997) and Daguet (2006). Results are presented in Table 2.

Bonneuil (1997) calculates the life expectancies of women in 1901–1905. I have calculated life expectancies for the same period. The comparison between these estimates shows that mine are on average higher: The median of the difference is 3.34%. In addition, 50% of departments have a difference between 0.49% and 6.05%, and 25% of them have a difference of more than 6.05%. The in-depth study of age-specific mortality rates reveals that these differences are largely explained by lower infant mortality rates. Nevertheless, since I cannot retrieve the death and population statistics of Bonneuil (1997), I do not know if these differences come from an underestimation of the number of deaths or an overestimation of the population at risk.

Daguet (2006) also reveals the departmental life expectancies at birth at the date of each census between 1954 and 1999. I compute the differences for both men and women. Overall, differences are much smaller. The median is around 0.2%, with no distinction for men and women and no temporal trend. The differences for 50% of the departments fall between 0% and 0.7% in 1962. These differences in 1999 for men are 0.22% and 0.73%, respectively. Although slight differences remain, one can conclude that life expectancies are reliable, even if slightly overestimated.

Table 2:   Differences between studies of departmental life expectancies at birth

	Men			Women		
	1st quartile	median	3rd quartile	1st quartile	median	3rd quartile
1901–1905	–	–	–	0.49	3.34	6.05
1954	0.18	0.65	1.00	0.54	0.84	1.34
1962	0.00	0.40	0.72	–0.01	0.37	0.68
1968	0.17	0.38	0.73	–0.02	0.33	0.78
1975	–0.17	0.15	0.50	–0.11	0.19	0.47
1982	0.01	0.27	0.59	0.04	0.21	0.50
1990	0.09	0.31	0.55	0.21	0.4	0.62
1999	0.22	0.49	0.73	0.47	0.66	0.99

Note: Differences as a percentage of my computations. Distribution of 90 or 95 departmental differences, according to the classification of the year.

4. Available results and discussion

4.1 Available results

Results are available for the 97 metropolitan departments monitored over the period 1901–2014, namely the departments of the current classification (Corse counting as one) as well as the old Seine and Seine-et-Oise. The variables available are the life expectan-

cies at each age as well as a set of lifetable variables from age 0 to age 105 and over (number of survivors, mortality rates, proportions of deaths). Yearly births and populations by age are also available.

To illustrate these results, I present in Figure 2 the departmental life expectancies at birth relative to the metropolitan average, for women. The first map shows the results for 1901. One can see that the highest life expectancies were located on an axis connecting the Southwest to the Northeast, from Ardennes to Landes. Maximums were reached in Ardennes but also in Pays de la Loire (Loir-et-Cher, Indre, Indre-et-Loire, Deux-Sèvres,... etc.) and Bourgogne (Côte d'Or, Yonne, Nièvre,... etc.), with values 10% to 20% higher than the metropolitan average. In contrast, life expectancies at birth in the Southeast, Seine, and Bretagne are significantly lower than the metropolitan average (between 5% and 20% according to the department). The second map presents these life expectancies at birth in the aftermath of the Second World War. At that time, maximums were reached in Loir-et-Cher, Creuse, and Alpes-Maritimes with life expectancies 5% to 10% higher than the metropolitan average. Central-West was still a leader region, while the regions of Bretagne and Normandie were still lagging behind.

Figure 2:   Life expectancy at birth for women (in percentage of the metropolitan mean), 1901 and 1946
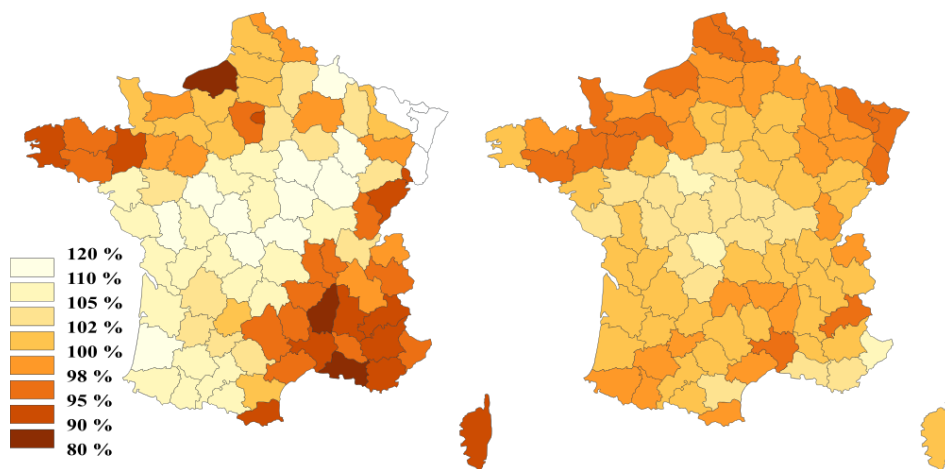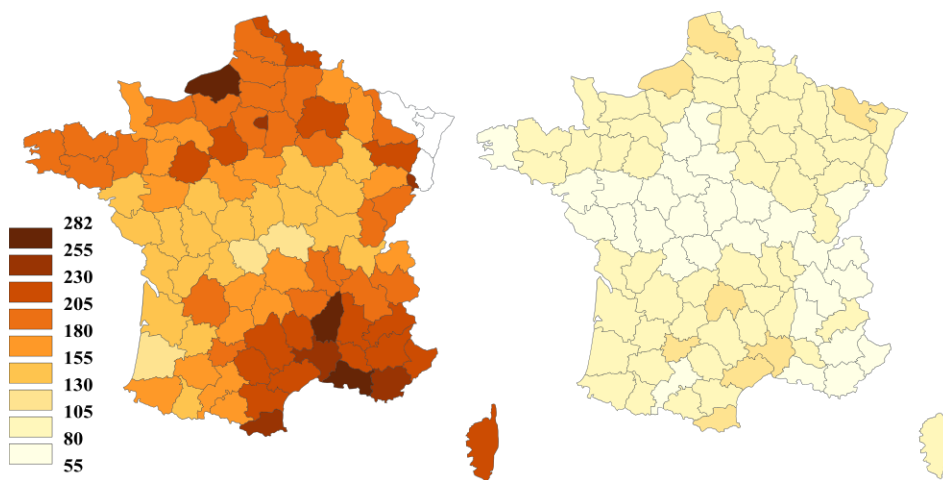

Note: Sample includes 90 departments. Moselle, Bas-Rhin and Haut-Rhin values are not available in 1901 (departments under German administration).

Rather than analyzing synthetic indicators such as life expectancy, one can look at age-specific indicators. Since mortality rates between 0 and 5 impacted strongly life expectancies at birth, I present these mortality rates for women in Figure 3. I represent

the rates per thousand, and no longer relative to the metropolitan average. The landscape in 1901 was relatively similar to the map of life expectancy, since child mortality rates were dramatically high in 1901. One can see that in extreme cases (Bouches-du-Rhône, Ardèche, Seine-Inferieure), for 1,000 children under age 5, between 255 and 282 died before their fifth birthday. Rates were generally higher in the North and the Southeast, while they were lower in a broad central band connecting the Saône-et-Loire to the Charente-Maritime and the Atlantic coast. Minimums (around 110) were reached in Creuse and Allier. The second map shows the same values in 1946. Child mortality rates decreased between the two years since they were globally around 90 per 1,000 in 1946. An under-mortality area was visible, from Eure-et-Loir to Isère via Nièvre. The Mediterranean coast presented diverse situations: Early mortality was low in the East (Var, Alpes-Maritimes) and strong in the West (Hérault, Gard, Pyrénées-Orientales).

Figure 3: **Mortality rates between 0 and 5 for women (per 1,000), 1901 and 1946**

Note: Sample includes 90 departments. Moselle, Bas-Rhin and Haut-Rhin values are not available in 1901 (departments under German administration).

Finally, one can analyze evolvements of a single department over the 1901–2014 period. Figure 4 shows female survivors at each age for different dates in Morbihan. I have chosen this *département* since it was a place of high mortality in 1901. Indeed, there was high infant mortality at that time: There were only 850 survivors in the fictitious cohort. This infant mortality almost completely disappeared in 1975. The survival curve shifted to the upper-right corner as mortality rates were globally declining. This displacement was important until 1975, mainly because of the drop in infant mortality. Subsequently, the curve moved mainly because of the decrease in mortality between 60 and 80 years, then beyond 80 years for the 1999–2014 period. This is in line with the literature about rectangularization of the survival curve (see Wilmoth and Horiuchi 1999, Fries 2002, Cheung et al. 2005, for example). This curve was in 2014 very flat until age 60 (there is almost no death below this age). Beyond this age the curve decreases dramatically, especially beyond age 80.

Figure 4: **Evolution of survivors at each age in Morbihan**
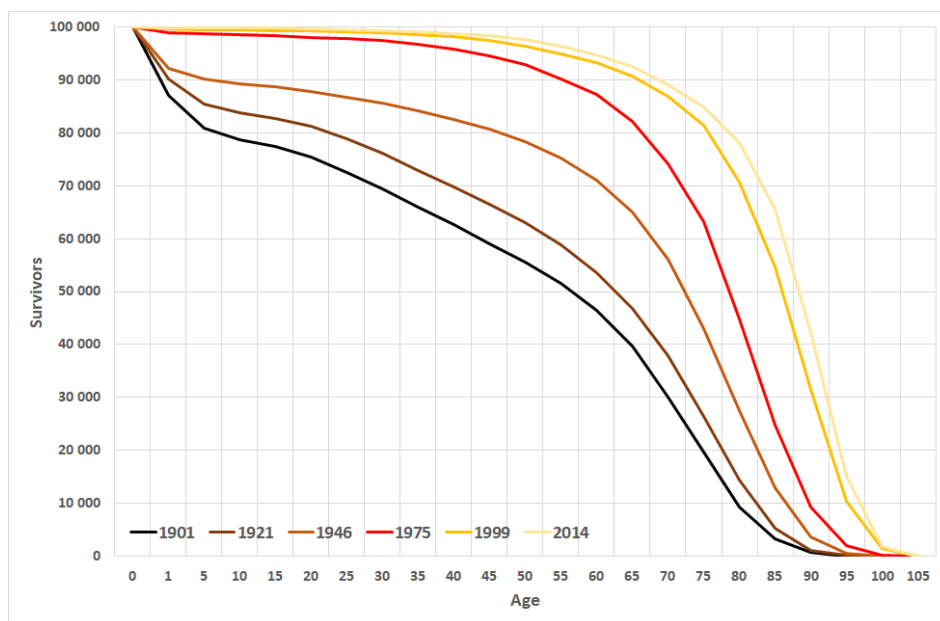
4.2 Discussion

4.2.1 Census reliability

With population censuses, one knows the spatial distribution of the population by age and sex between the French departments during the course of the 20th century. During this period, censuses served as a support for some public choices. The first concerns local budgets, since allocations coming from central administration were based on the population of each territory. These censuses therefore affected the spatial distribution of public finance. The second concerns the electoral divisions: To achieve fair representation in local and national assemblies, France created electoral divisions with roughly equal numbers of people. Censuses therefore had a very strong political impact. As a result, some regions have sought to inflate their census populations to get greater financial or electoral weight. Historians and statisticians have shown that Marseille's population was overestimated in the 1930s.[9] This was also true in Corse in 1962: Results of the exhaustive counting were not published because of inconsistencies. These censuses are, however, the basis of age–population computations. Even though ambiguous cases remain marginal over the period, they nevertheless existed.

4.2.2 Interdepartmental migrations

Methods used in this study partly take into account the issue of migrations. At each census date, the difference between estimated and recorded population can be seen as an approximation of net migration flows at each age. These flows are then distributed in proportion to the time elapsed between the first census and January 1 of each year of the intercensal period. This approximation does not affect my results when the flows are weak or if they follow the approximation used. One can consider that this was true between 1901 and 1911: During the rural exodus, the migrations took place progressively. But this approximation is not verified during war periods. The May–June 1940 exodus is an emblematic example. To escape the advance of German troops, the populations of the Northeast migrated in mass toward the South and the West. I cannot take into account this exodus with the methodology used: For example, Ardennes' population in 1941 is largely overestimated. This issue is presented on several occasions in the *Statistique Annuelle du Mouvement de la Population* between 1939 and 1942.[10] It has been suggested that this publication estimated the population based on the distribution of ration tickets. Howewer, Alary et al. (2006) showed that these tickets were circumvented during the war, questioning their reliability for counting the present population. Bonnet

---

[9] See, for example *Statistique Annuelle du Mouvement de la Population,* 1939–1942, page 4.
[10] See *Statistique Annuelle du Mouvement de la Population,* 1939–1942, pages 3–4, 47, and 55.

(2019) tried to estimate these departmental populations, but only for females and for the total population.

Another issue relating to interdepartmental migrations concerns migration after the birth of children. These migrations have potential consequences for infant mortality estimates. Since cities concentrated health facilities, mothers living in the countryside came to give birth to their children. In official publications, these births are reported in the mother's home department, but it is possible that some of those who died just after birth were registered in the cities (Fariñas and Oris 2016). In addition, mothers living in cities sent their children to nurseries in the countryside shortly after birth. These births were recorded in the cities, but some of those who subsequently died were recorded in the countryside. Depending on the weight of these two effects, infant mortality in urban departments is overestimated or underestimated in this new database.

4.2.3 Domiciliation of deaths during the two world wars

The sources I use to estimate life expectancies during the two world wars are incomplete: Military and deportee deaths were recorded by birth department and not by home department. I have build matrices linking birth department and home department before the deportation, but they rely on strong assumptions about the representativity of pre- and post-war situations concerning the phenomena that took place during the war. The few statistics kept for this period limit the possibilities to go further. Regarding military deaths, I assume that the home department was similar to the birth department concerning the *Morts Pour La France*. If this hypothesis seems weaker than those assumed for deportees, it is not entirely satisfactory. Again, I lack reliable and available data to overcome this issue.

4.2.4 Small department figures

Estimating fertility or mortality rates is difficult when figures are small, namely around zero. Papers tackle this issue by using Bayesian estimation processes (Asunção et al. 2005 and Schmertmann et al. 2014 for fertility rates, Alexander et al. 2017 for mortality rates). I considered using these methods to supplement the HMD protocol. However, the French departments are not as small as the geographical units used in these studies. For example, the minimum according to population is 50,000 women (Territoire de Belfort, 1901), compared to 2,000 for some counties. However, these estimation models may be applied in the future, particularly to compute confidence intervals around departmental life expectancies.

5. Conclusion

In this paper, I have presented the sources and methods used to estimate lifetables by sex for all French metropolitan departments from 1901 to 2014. To do so, I collected vital records and census statistics at the departmental level since the beginning of the 20th century. Since the two world wars afflicted France in 1914–1918 and 1939–1945, military deaths and deaths in deportation were of great importance in the lifetables estimates; these statistics were collected at the departmental level in original sources, namely the *Mémoire des Hommes* and MemorialGenWeb databases.

To estimate departmental lifetables, I have referred to the methods used in a large number of countries by the researchers of the Human Mortality Database. These methods transform the collected raw data into homogeneous data. They include the use of cubic splines to estimate deaths by age groups, the Kannisto model to extrapolate deaths at older ages, and a panel of methods to estimate populations at January 1 of each year. The HMD protocol has been amended to take into account the French data specificities. This concerns false stillbirths reintroduced in the statistics of births and infant deaths, and territorial breaks such as those that affected the Paris region in 1968.

This work provides a new database on departmental mortality for the entire 20th century, which will be updated with the most recent data in a continuous way. Coupled with Bonneuil's (1997) estimations for the 20th century, it provides an overview of the local trends in mortality since the French Revolution. As they have been calculated for each sex, this database shed new insights on the reasons explaining the differences in life expectancy between men and women. Moreover, beyond mortality statistics, this new database can be used to analyze all demographic fields at the local level: birth rates since it includes annual births, the spatial distribution of population since it provides yearly populations by age, and finally internal migrations. These fields of research are on my future agenda.

6. Acknowledgments

References

Alary, E., Vergez-Chaignon, B., and Gauvin, G. (2006). *Les Français au quotidien, 1939–1949*. Paris: Perrin.

Alexander, M., Zagheni, E., and Barbieri, M. (2017). A flexible Bayesian model for estimating subnational mortality. *Demography* 54(6): 2025–2041. doi:10.1007/s13524-017-0618-7.

Assunção, R.M., Schmertmann, C.P., Potter, J.E., and Cavenaghi, S.M. (2005). Empirical Bayes estimation of demographic schedules for small areas. *Demography* 42(3): 537–558. doi:10.1353/dem.2005.0022.

Barbieri, M. (2013). La mortalité départementale en France. *Population* 68(3): 433–479. doi:10.3917/popu.1303.0433.

Bonnet, F. (2019). Beyond the exodus of May–June 1940: Internal migrations in France during the Second World War. Paris: Paris School of Economics (Working Paper 2019-23).

Bonneuil, N. (1997). *Transformation of the French demographic landscape 1806–1906*. Oxford: Clarendon Press.

Cheung, S.L.K., Robine, J.-M., Tu, E.J.C., and Caselli, G. (2005). Three dimensions of the survival curve: Horizontalization, verticalization, and longevity extension. *Demography* 42(2): 243–258. doi:10.1353/dem.2005.0012.

Coale, A.J., Demeny, P., and Vaughan, B. (2013). *Regional model life tables and stable populations: Studies in population*. Amsterdam: Elsevier.

Daguet, F. (2006). *Données de démographie régionale de 1954 à 1999*. Paris: INSEE.

Dupâquier, J. and Bardet, J.P. (1988). *Histoire de la population française*. Paris: Presses Universitaires de France. doi:10.3917/puf.dupaq.1988.01.

Fariñas, D.R. and Oris, M. (2016). *New approaches to death in cities during the health transition*. Berlin: Springer. doi:10.1007/978-3-319-43002-7.

Fries, J.F. (2002). Aging, natural death, and the compression of morbidity. *Bulletin of the World Health Organization* 80(3): 245–250.

Huber, M. (1931). *La population de la France pendant la guerre*. Paris: Presses Universitaires de France.

Lagrou, P., Audouin-Rouzeau, S., Becker, A., Ingrao, C., and Rousso, H. (2002). Les guerres, les morts et le deuil. Bilan chiffré de la Seconde Guerre Mondiale. In: Asséo, H. (ed.). *La violence de guerre, 1914–1945. Approches comparées des deux conflits*

mondiaux. Brussels: Éditions Complexe: 313–327.

Pedroncini, G., Corvisier, A., and Blanchard, A. (1992). *Histoire militaire de la France. 3. De 1871 à 1940*. Paris: Presses Universitaires de France.

Preston, S., Heuveline, P., and Guillot, M. (2000). *Demography: Measuring and modeling population processes*. Hoboken: Wiley-Blackwell.

Prost, A. (2008). Compter les vivants et les morts: L'évaluation des pertes françaises de 1914–1918. *Le Mouvement Social* (1): 41–60. doi:10.3917/lms.222.0041.

Schmertmann, C., Zagheni, E., Goldstein, J.R., and Myrskylä, M. (2014). Bayesian forecasting of cohort fertility. *Journal of the American Statistical Association* 109(506): 500–513. doi:10.1080/01621459.2014.881738.

Vallin, J. and Meslé, F. (2001). *Tables de mortalité françaises pour les XIXe et XXe siècles et projections pour le XXIe siècle*. Paris: Éditions de l'Institut National d'Etudes Démographiques.

Vallin, J. and Meslé, F. (2005). Convergences and divergences: An analytical framework of national and sub-national trends in life expectancy. *Genus* 61(1): 83–124.

Van de Walle, É. (1974). *The female population of France in the 19th century: A reconstruction of 82 Départements*. Princeton: Princeton University Press.

Wilmoth, J.R., Andreev, K., Jdanov, D., Glei, D.A., Boe, C., Bubenheim, M., Philipov, D., Shkolnikov, V., and Vachon, P. (2007). Methods protocol for the Human Mortality Database (version 31/05/2007). Berkeley: University of California, Rostock: Max Planck Institute for Demographic Research.

Wilmoth, J.R. and Horiuchi, S. (1999). Rectangularization revisited: Variability of age at death within human populations. *Demography* 36(4): 475–495. doi:10.2307/2648085.