The French Human Mortality Database

Florian Bonnet[*], Hippolyte d'Albis[†]

April 4, 2022

Abstract

Political and societal debates about territorial disparities are numerous in France. However, historical data to inform these debates are often lacking. This study aims to fill some of these gaps by making available on a dedicated website the statistics of mortality by age and sex at the regional level in France from 1901 to 2020. It is based on the exhaustive collection and digitization of population movement (births and deaths by age) and census data (populations by age) in the French departments, as well as on the data about both military and deportation mortality during the two World Wars. The methodology used is mainly the one used to build the national lifetables available in the Human Mortality Database. This new database will be updated annually to incorporate the most recent mortality data. It completes a limited statistical offer on the dynamics of mortality at the local level in developed countries.

[*]Institut National d'Etudes Démographiques, florian.bonnet@ined.fr
[†]Paris School of Economics & CNRS, hdalbis@psemail.eu

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 men 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. As such, significant differences exist between 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. In order to inform public decision-makers in their choices, it is important to know the history of these departmental differences. Accordingly, I have calculated regional life tables since 1901. This document presents the sources and methods used.

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 at each census (population by age and sex). I exploit a French unique characteristic: since 1789, this country is 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, in order to take into account the two World Wars that affected France between 1914–1918 and 1939–1945, 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 1st.

These lifetables at the regional level complete a still incomplete literature. Bonneuil (1997) works on departmental mortality in the 19th century: he computes women lifetables 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 in the same way men's mortality, because of strong fluctuations due to the wars which afflicted France at this time. From 1954 to 1999, Daguet (2006) groups lifetables established at the departmental level, but only for the census years. Barbieri (2013) uses in her study departmental lifetables calculated by INSEE for the period 1975–2008 . However, the data was provided exceptionally. Vallin and Meslé (2005) uses departmental life expectancies for the period 1906–1954. However, both reconstruction methods and data have never been published. Lastly, various 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, which is not the case of 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 regional level in few developed countries. The results according to Canadian provinces from 1921 onwards are available in the Canadian Human Mortality Database[1]; those according to the Japanese provinces since 1975 are available in the Japanese Mortality Database[2]; those according to the federal states in Germany since 1901 in van Raalte et al. (2020); those according to states of the USA since 1959 in the United States Mortality Database[3]; those according to the territories and states of Australia since 1971 in the Australian Human Mortality Database[4]. This paper therefore complements a still limited supply of regional mortality data freely available by adopting an internationally recognized protocol; this allows international comparisons without methodological bias.

Lifetables are available on a dedicated website, the French Regional Database.[5] They can be freely used by any person interested in mortality at the local level in France. They are available for 3 different geographical levels corresponding to the 3 levels of the Nomenclature of Territorial Units for Statistics: the departments (NUTS 3 level, 97 units), the administrative regions that existed between 1970 and 2015 (NUTS 2 level, 22 units) and the current administrative regions (NUTS 1 level, 13 units). This database will be updated annually to get the most recent lifetables. To date, regional lifetables are available from 1901 to 2020.

2 Sources

Computations of departmental lifetables requires two types of raw statistics: vital statistics (deaths and births domiciled) and population censuses. The deaths collected do not only concern civilian deaths: both military deaths during the two World Wars and deportation deaths between 1939 and 1945 have been included.

[1]Canadian Human Mortality Database. Department of Demography, Université de Montréal (Canada). Available at www.demo.umontreal.ca/chmd/

[2]"Japanese Mortality Database", National Institute of Population and Social Security Research, Available at http://www.ipss.go.jp/p-toukei/JMD/index-en.asp

[3]United States Mortality DataBase. University of California, Berkeley (USA). Available at usa.mortality.org

[4]https://demography.cass.anu.edu.au/research/australian-human-mortality-database

[5]https://frdata.org/en/french-human-mortality-database/

2.1 Deaths

Civilian deaths of each department, sex and year over the period 1901–2020 have been retrieved from the statistics published by *Statistique Générale de la France* (SGF) and then by *Institut National de la Statistique et des Etudes Economiques* (INSEE). I have retrieved deaths by age group recorded in home department. 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 Defense Ministry's website.[6] They are available by birth year at the departmental level, and by birth year and death year at the national level. Table 7, Table 8 and Table 9 present the sources from which deaths were retrieved.

Individuals who died during deportation in the Second World War are not included in the civilian population movement. However, they were nearly 100,000. I have decided to include them in my lifetables, using data from "*MemorialGenWeb*" website.[7] This database records deportees who left France and died in deportation published in the *Journal Officiel*, by birth department if they were born in France, birth country otherwise. 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 1901–2019 and the period 1853–1900. Table 10, 11 and 12 present the sources from which births were retrieved. I have also recovered stillbirths by mother's home department and year without distinction of sex.

2.3 Censuses

Population census in France were conducted at regular intervals from 1901 to 1999. There have been 15 censuses: 1901, 1906, 1911, 1921, 1926, 1931, 1936, 1946, 1954, 1962, 1968, 1975, 1982, 1990, 1999. Since 2004, the census has been based on an annual collection of information, covering successively all territories over a five-year period. The first census collected using this method (rolling census) was in 2008. Since then, each year, census results are produced from the five most recent annual surveys: information from the oldest survey is dropped and the new survey is taken into account.

Consequently, I have collected populations by birth year, home department, and sex for each global

[6] http://www.memoiredeshommes.sga.defense.gouv.fr/, downloaded on February, 2016.
[7] http://www.memorialgenweb.org/memorial3/deportes/index.php, forwarded on March, 2016.

census of the period 1901–1962 from hard-copy publications of SGF and INSEE. For the period 1968–1999, these statistics for global censuses have been found in on-line sources. Finally, I get populations from rolling censuses in 2008, 2013, 2014 and 2015 in on-line sources. Table 13 presents the sources from which populations at census were retrieved.

From 2016 to 2021, I use the annual populations by five-year age groups estimated by INSEE. These estimated populations will be gradually replaced by census populations as they become available.

3 Methods

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

3.1 Births

The number of total births by department and sex is available by mother's place of residence for each year during the period 1853 - 2020. 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 deaths 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. The number of stillbirths by sex is available at the national level for the period 1899-1974. The proportion of false stillbirths among stillbirths remained close to 15% from 1899 to 1950, and 20% from 1950 to 1974. These false stillbirths represented 0.6% of births recorded in the archives in 1899; this figure decreased continuously until 1974 to reach 0.3%.

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 national level by Vallin and Meslé (2001) is correct. Moreover, these departmental lifetables do not allow to study the issue of infant mortality including stillbirths.

3.2 Deaths

3.2.1 Civil Deaths

Civil deaths are available by department, sex, and age according to the place of residence of the deceased, for each year during the period 1901-2020. These civil deaths were reprocessed in 4 steps to provide single age statistics.

Distribution of deaths with unknown age The raw statistics collected include an "unknown age" category during the periods 1914-1924 and 1931-1955. In most cases, the share of deaths of unknown age is very small (less than 0.5% of total deaths). However, this share is higher than this threshold in some departments for the years 1919-1920 and 1940, 1943-1945. For men, there were 57 departments with a share of deaths of unknown age between 1% and 5% in 1944 (3% in 1919, 11% in 1940, 2% in 1943, 4% in 1945 and 1% in 1946), and 7 departments with a share of deaths of unknown age between 5 and 10% for the same year. For women, there were 7 departments with a share of deaths of unknown age between 1 and 5% in 1940, and 9 departments in 1944.

The deaths of unknown age in each department and for each sex were distributed pro rata to the number of deaths in each age group.

Estimation of deaths by single age: cubic splines and hermitan splines To get a 1×1 format (single age, year of death) for the deaths between 1901 and 1967, I adjust the curve of cumulative deaths by means of cubic splines and hermitan splines.

Cubic Spline is a semi-parametric estimation method which joins the points of a cumulative distribution by third degree polynomials. Let $Y(x) = \sum_{u=0}^{x-1} D_u$ be the cumulative number of deaths up to age x. $Y(x)$ is known for a limited collection of ages including 1, 5, 10... etc from the raw data, the highest age in the distribution and the age above which no further deaths are observed, set at 105. Equation (1) fits a cubic spline by using these values ($I(.)$ equals one if the logical statement within parentheses is true and zero otherwise):

$$Y(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 + \beta_1 (x - k_1) I(x > k_1) + \ldots + \beta_n (x - k_n) I(x > k_n). \tag{1}$$

I have to estimate the vector $(\alpha_0; \alpha_1; \alpha_2; \alpha_3; \beta_1; \ldots; \beta_n)$ which contains $n+4$ coefficients, but I only know $n+2$ values of $Y(x)$, and therefore $n+2$ constraints. Two further constraints must be introduced to identify the model. First I assume that there is no death at the upper bound, namely 105. Second I assume that deaths observed between 1 and 5-year-old occured between 1 and 2 year-old. $\hat{Y}(x)$ are calculated for all ages, for each department, sex, and year using function cubicspline of the pracma package of R software. Deaths at age x are found as follows:

$$\hat{D}(x) = \hat{Y}(x+1) - \hat{Y}(x).$$

This method can not be used to estimate deaths by single age at young ages. Indeed, the cumulative death curve is fitted by polynomials of degree 3: it is therefore not constrained to be increasing at any point. Consequently, one could obtain negative deaths at young ages, which is indicated in the HMD protocol. To overcome this issue, I use cubic hermite splines (function pchip of the pracma package of R software), which constrains the fitted curve of cumulative deaths to be increasing at any point. Nevertheless, this method does not allow a good fit of the deaths for ages beyond the open interval, when the open interval starts at an age where the number of deaths is still important. For intermediate ages, the two adjustments coincide.

In order to use the advantages of both adjustments, I use the deaths estimated by the cubic hermite splines for ages less than 40, and I use the deaths estimated by the cubic splines for ages greater than or equal to 40.

Death adjustment at old ages Deaths estimated by cubic splines and hermitan splines are too imprecise to be used at advanced ages since open-age interval of deaths is too low. For years 1901 to 1967, these deaths are adjusted by means of the Kannisto model, which assumes a survival curve of logistic shape, with a zero-asymptote for very old ages. I use this method for deaths beyond the open-age interval – different according to the periods, I keep a maximum of 95 so that estimates are not hindered by too small figures – and rely on the deaths observed for ages 10 years below this limit. Thus, if the open-age interval begins at age 90, I use the ages 80–89. Formally, I compute a fictitious survival curve $S(80+x)$:

$$S(80+x) = \frac{\sum_{u=80+x}^{105} D_u}{\sum_{u=80}^{105} D_u} \quad \text{for } x = 0,1,...,9. \tag{2}$$

This survival function conditional on reaching age 80 may be seen as tracking a "synthetic extinct cohort", since it is based on annual deaths and not on deaths in the cohort itself. Assuming that this fictitious cohort displays survival probabilities that can be fitted by the Kannisto model, the survival function $s(x)$ is:

$$s(x) = \left(\frac{1+a}{1+ae^{b(x-80)}} \right)^{1/b}. \tag{3}$$

I compute $\hat{s}(x)$ and $d(x) = \hat{s}(x) - \hat{s}(x+1)$ with estimated values for a and b. Then, I obtain deaths at each age:

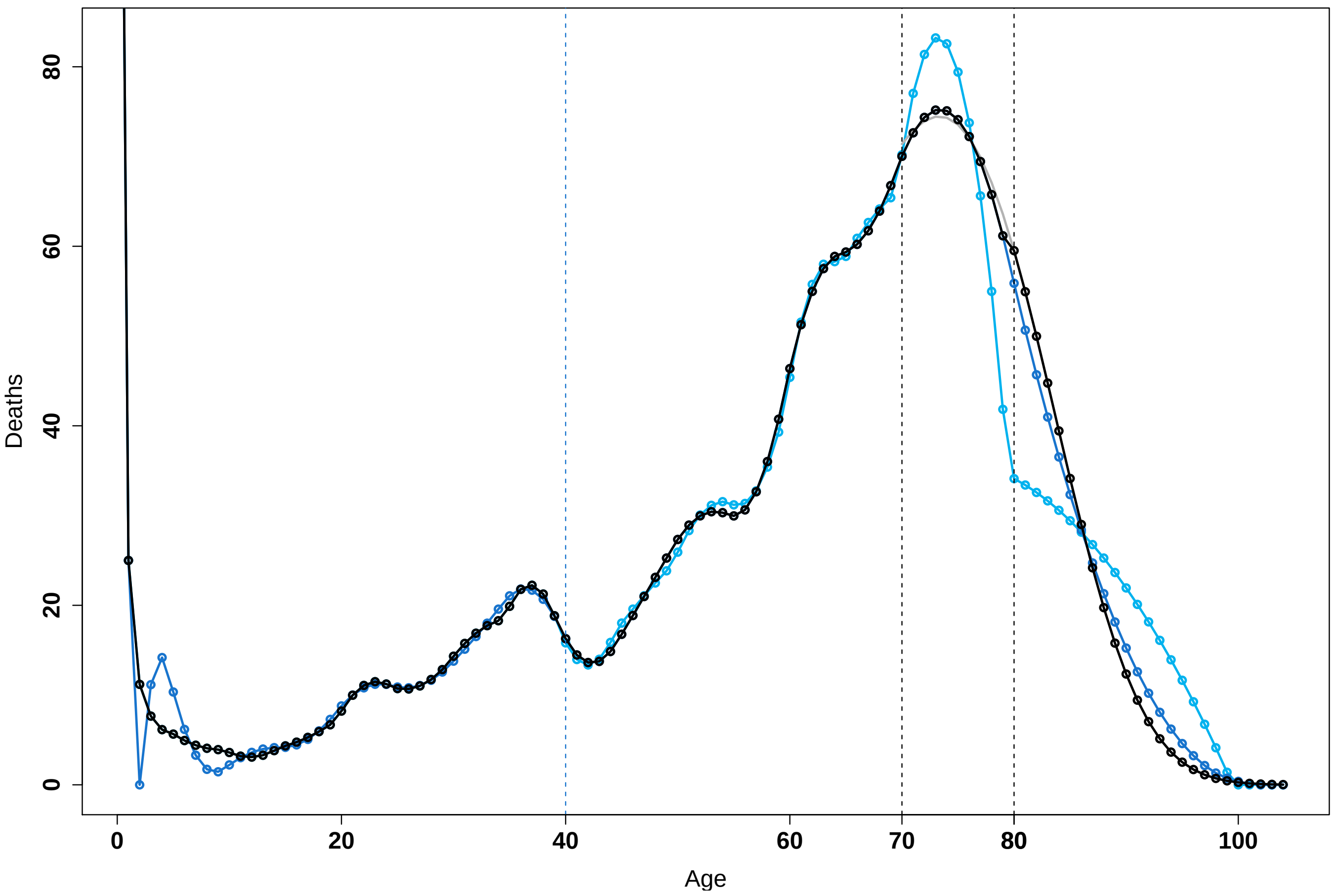
$$D(x) = \sum_{u=90}^{105} D_u \times \frac{d(x)}{\hat{s}(90)}. \tag{4}$$

Figure 1 summarizes the process of estimating deaths by age using the example of male deaths in the department of Ain in 1936. The dark blue curve represents the deaths by age estimated by the cubic splines method, the light blue curve the deaths by age estimated by the hermite splines method. The grey curve represents the age-specific deaths estimated by the Kannisto model over the support age interval (in this specific case, between ages 70 and 80). The black curve reveals the deaths by single age used: it is the junction of the deaths estimated by the hermite splines method up to age 40, the deaths estimated by cubic splines from age 40 to 80, and the deaths estimated by the Kannisto model beyond age 80. One can see that deaths estimated by the cubic splines method are highly volatile up to age 10, and negative at age 3. Between ages 20 and 70, the light blue and dark blue curves merge with the black curve, the estimates being roughly the same for both methods. Beyond age 70, the deaths estimated by the hermite splines method are no longer realistic: the open age interval observed in the raw statistics is too low, and the number of deaths at this age is still very high. Finally, the Kannisto model reveals that deaths are underestimated by the cubic spline method below age 90, and overestimated above that age.

False stillbirths To get unbiased infant mortality rates, false stillbirths included in births are also included into deaths of age 0.

Consistency with national data To get estimations consistent with the lifetables computed for France as a whole, the sum of departmental deaths by age is compared with the deaths by single age calculated by Vallin et Meslé (2001). The national lifetables were estimated using more accurate national statistics, which provide more reliable deaths by single age. I calculate the difference between the sum of departmental deaths and national deaths for each age, sex, and year in the period 1901-1967, during which deaths are

Figure 1: ESTIMATION OF DEATHS BY SINGLE AGE: OVERVIEW

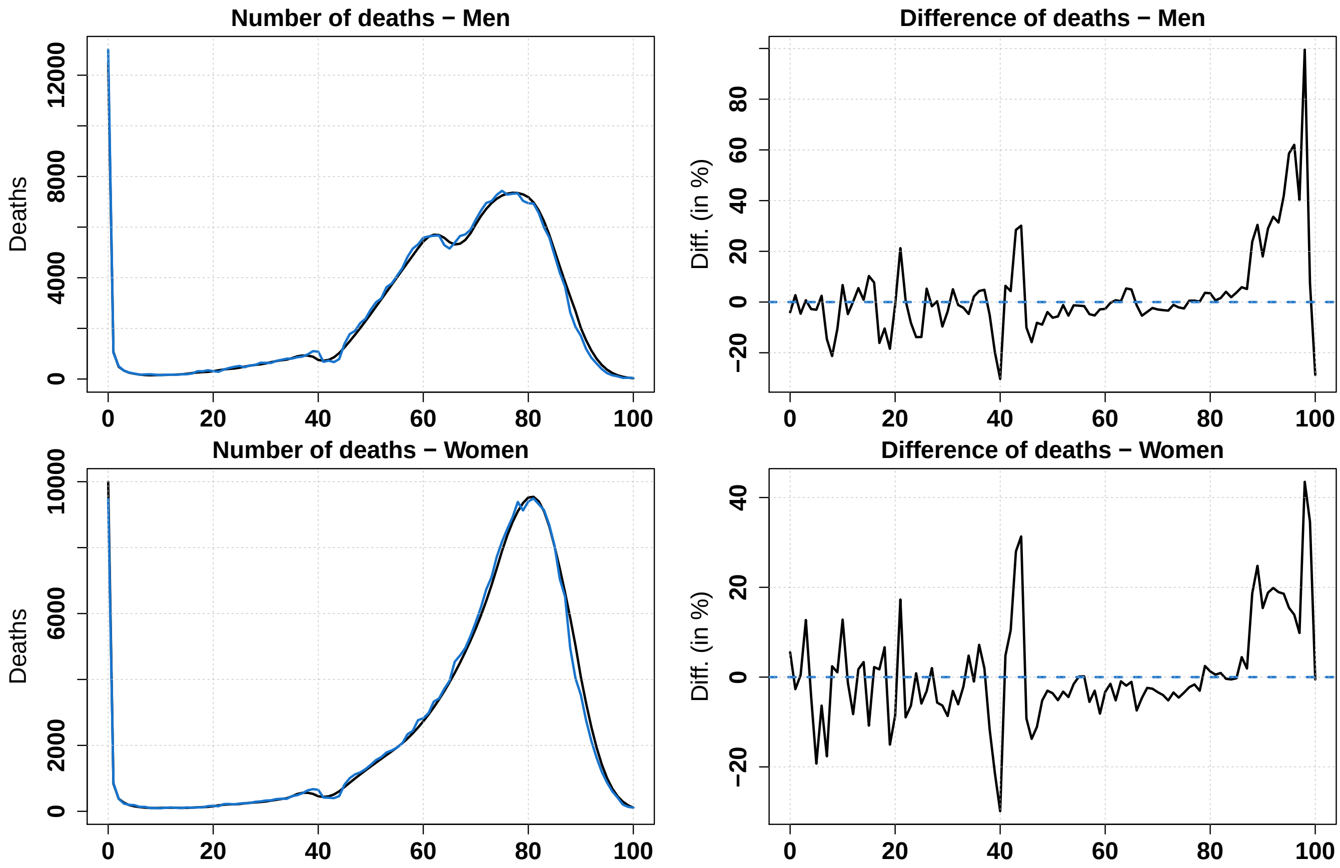


Note: Number of men deaths by single age in the department of Ain in 1936. Dark blue curve for deaths by age estimated by the cubic splines method; light blue curve for deaths by age estimated by the hermite splines method. Grey curve for deaths by age estimated by the Kannisto model over the support age interval (in this specific case, between ages 70 and 80). Black curve for deaths by single age used.

reported by age groups. I uniformly adjust the departmental deaths so that the two sources match. This adjustment is crucial in 1967 since deaths are available by very large age groups.

Figure 2 presents the year 1960. The estimated distribution of deaths by single age is depicted in black on the quadrants on the left, for males (top) and for females (bottom). The estimated distribution of deaths by single age estimated by Vallin is depicted in blue on the same figures. The difference between the two curves is plotted in black on the quadrants on the right. One can see that the curves are different for the cohorts born during the First World War, which are around age 40 at that date, because of a strong variability of fertility from one year to another. They are also different for male cohorts whose mortality was very high during the First World War, which are at age 65 to 70 at that date. Finally, the large relative differences observed at older ages can be explained by a low number of deaths for ages 90 and more.

Figure 2: CONSISTENCY OF DEATHS BY SINGLE AGE IN 1960



Note: Sum of departmental deaths by single age in black on left quadrants, for males (top) and for females (bottom). Deaths by single age for France estimated by Vallin et Meslé (2001) in blue. Difference between the two curves plotted in black on right quadrants.

3.2.2 Military Deaths

Overview of the method The two World Wars had significant demographic effects both at national and departmental level. The first is due to internal migrations caused by the conflict and the France's division into occupied and unoccupied zones in 1940. Raw statistics give no direct indication for this question. The second 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 subnational level.

According to military deaths during the two World Wars, I couple 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 his 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 national level (Pedroncini (1992), Prost (2008), Hubert (1931), Lagrou et al. (2002)), so as to verify the overall consistency of the various sources.

Military deaths by département, year of birth and year of death Ideally, the statistics of military deaths should be available according to the age and the year of the soldier's death, as well as his home department before the war. Since the sources used are incomplete, I couple two different matrices.

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 classification of departments from the *"Mémoire des Hommes"* website is modified to fit the classification for civilian deaths. Problems concern *Corse* (two departments counting as one) and the old departments of *Seine* and *Seine-et-Oise*. For these last two, deaths are given according to the new departments. To allocate deaths between *Seine* and *Seine-et-Oise* I first sum all deaths in *Ile-de-France* (without *Seine-et-Marne*), then I allocate these military deaths pro rata of population in the cohorts born from 1880 to 1896. These cohorts account for 83% of total military deaths in the First World War. Concerning the distribution of deaths in the Parisian departments between *Seine* and *Seine-et-Oise* for the Second World War, I allocate them pro rata of populations born between 1905 and 1921 (70% of total deaths during the Second World War). *Seine*'s deaths are equal to 78.6% of the total.

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 his personal research on a specific soldier, can inform both his birth year and death year. This work has been done for just over 20% of total deaths. I wonder if this sample is representative of the distribution by death year. To do so, I use Pedroncini (1992)'s work: it gives total military deaths by death year. Table 1 shows these distributions according to both sources. Even if discrepancies exist, I consider that I can use the sample coming from *Mémoire des Hommes*. Data by birth year and death year are therefore extracted from the Defense Ministry's database.

TABLE 1: DISTRIBUTION OF SOLDIERS DEATHS, BY YEAR OF DEATH

Source		1914	1915	1916	1917	1918	Total
Mémoire des hommes	Deaths	75,403	82,878	50,933	34,436	52,459	296,109
	% of the total	25.46%	27.99%	17.20%	11.63%	17.72%	100%
Pedroncini (1992)	Deaths	301,000	349,000	252,000	164,000	235,000	1,301,000
	% of the total	23.14%	26.83%	19.37%	12.61%	18.06%	100%

Note: Distribution of military deaths by year of death according to "*Mémoire des hommes*" website and Pedroncini (1992).

By cross-referencing the two matrices, I get a matrix giving total deaths by department, birth year and death year. I assume that there is little variation between departments in the death year according to the cohort.

To ease the collection of data from the website, military deaths have been retrieved by birth year for the youngest (born after 1889), then by five-year group for those born in 1889 and earlier. These deaths must be split by birth year, which is done by hermite cubic splines for each departement and year of death.

General adjustment by the total of military deaths This distribution of deaths is then adjusted by the total of deaths as estimated by researchers at national level, so as to verify the overall consistency of the various sources. Prost (2008) makes an inventory of the statistical estimates of deaths during the First World War. He uses the Marin's report, followed by Hubert (1931) and Dupaquier (1988). Roure's report cited by Prost (2008) reveals 1,357,800 military casualties, taking into account deaths of foreigners. Hubert (1931) added 40,000 soldiers dead during the 6 months after the armistice, as well as sailors. Table 2 summarizes these numbers. Regarding the 28,600 deaths that occurred 6 months after the armistice, I assume that they had been included in the 1919 deaths of the population movement and do not take them into account. With regard to the 75,700 deaths of soldiers coming from settlements and abroad, I do not keep them in the total since these populations were not registered in 1911 in the French *départements* and were surely recorded in the civilian deaths of their home country. Finally, I obtain 1,304,400 deaths.

TABLE 2: TOTAL OF MILITARTY DEATHS DURING THE FIRST WORLD WAR

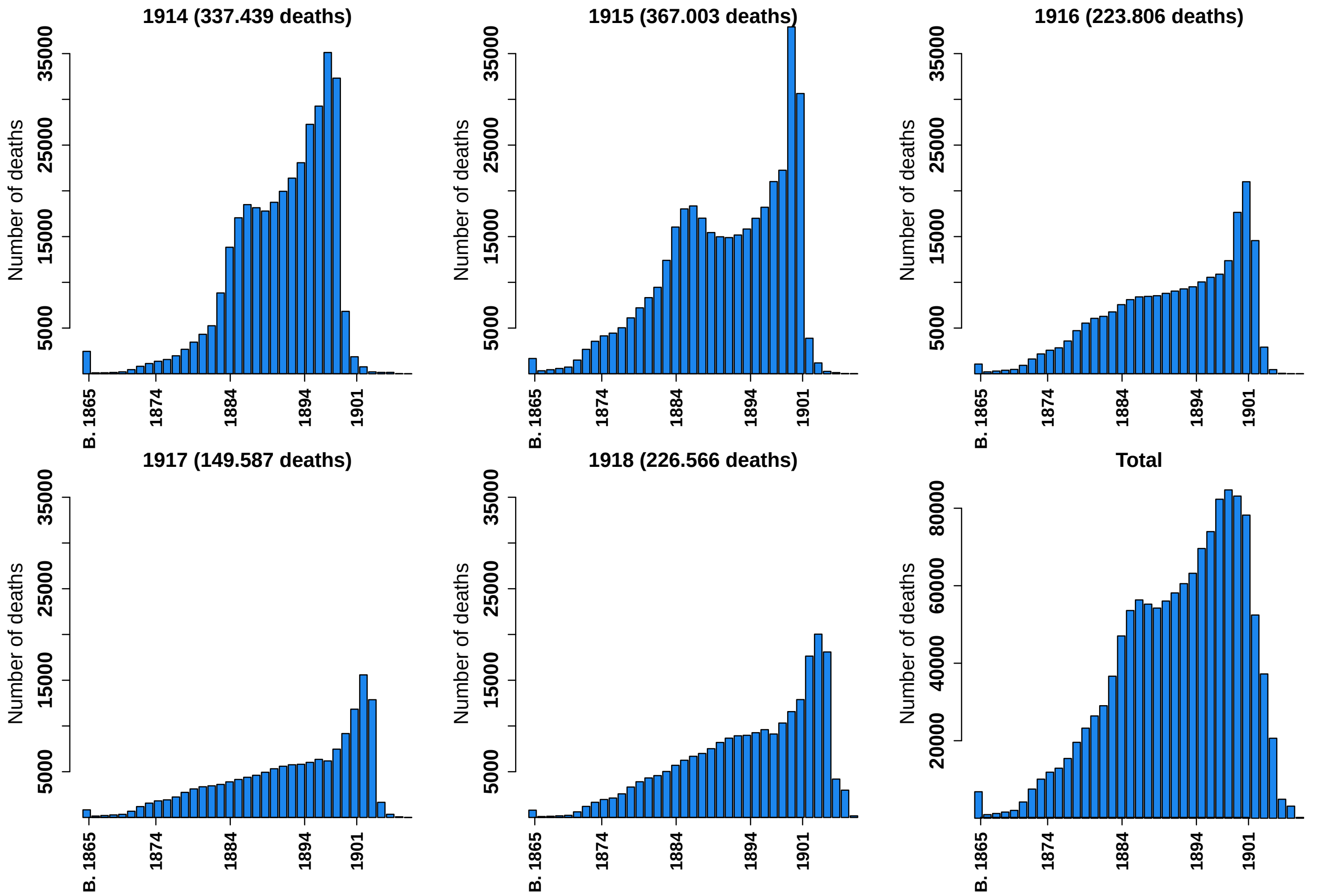
Source	Variable	Deaths
Roure	Total of French military deaths	1,282,100
	Total foreign-born and settlements	75,700
	Total Roure	1,357,800
Hubert	Deaths 6 months after armistice	28,600
	Sailors	11,400
	Final total	1,397,800

Note: Military deaths during the First World War in France according to Roure's report and Hbert (1931) cited by Prost (2008).

To illustrate these computations, Figure 3 shows the number of deaths by year of death and cohort between 1914 and 1918, as well as the number of deaths by cohort for the whole First World War. The total number of deaths per year is also indicated. First World War was particularly deadly for the cohorts born between 1895 and 1901, with nearly 400,000 deaths.

Figure 4 maps military death rates at department level during the First World War. It is computed by

Figure 3: NUMBER OF DEATHS BY COHORT AND YEAR OF DEATH IN FRANCE DURING THE FIRST WORLD WAR



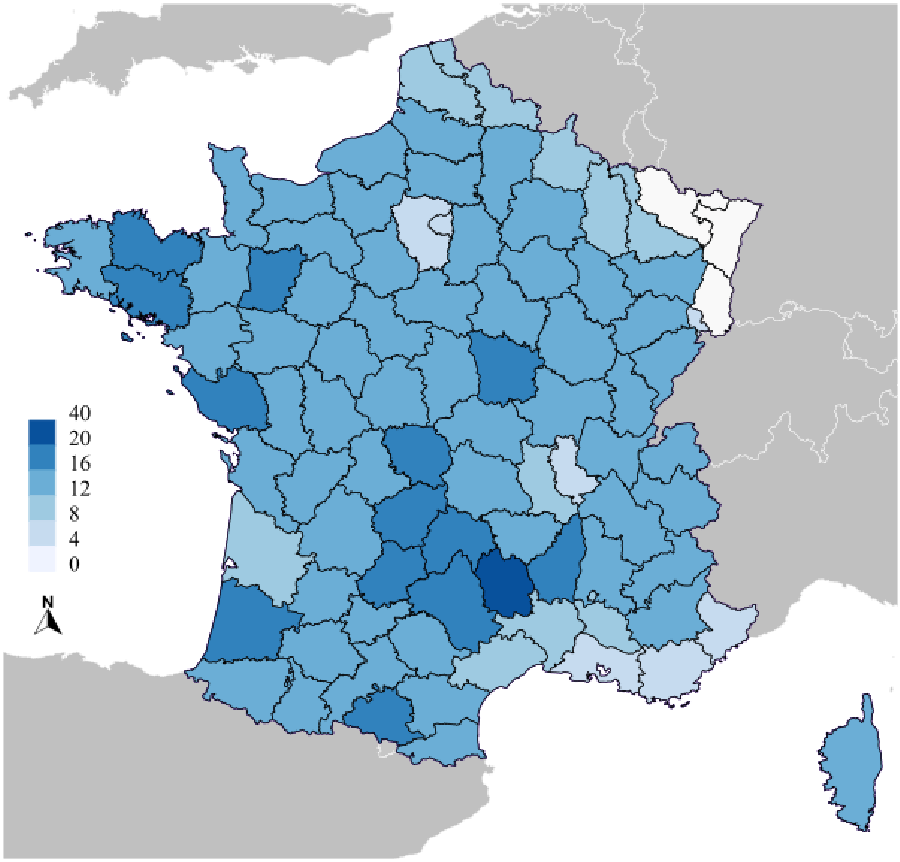
Note: Number of deaths by cohort and year of death in France during the First World War. "B. 1865" means "cohorts born in 1865 and before".

dividing the total number of deaths in each department by the population of men born between 1865 and 1900 in 1911 census. Values are expressed per 100 men. A light color means that mortality was high, a dark color means that mortality was low. At the national level, this value is 12.2%. One can see that military mortality was lower near the Mediterranean coast, in some departments with large cities (Paris, Lyon, Bordeaux), and near the northern and eastern borders. Conversely, it was high in the south of the Massif Central, in the south of the Alps and in the departments of Brittany or close to this region (Mayenne, Vendée). This map echoes Gilles et al (2014), which shows that mortality was lower than elsewhere in Provence and Ile-de-France, and higher in Limousin and Brittany. This result could be partly explained by the economic characteristics of the territories such as the share of the agricultural population. Nevertheless, their model fails to explain the under-mortality in Provence.

The principle is the same for the Second World War. The two matrices combined come from the Defense Ministry's database. The total of deaths I use is 200,000, in line with Lagrou et al. (2002).

Figure 5 shows the number of deaths by year of death and cohort between 1939 and 1945, as well as the

Figure 4: MILITARY DEATH RATE (IN %) DURING THE FIRST WORLD WAR



Note: Military death rate is the ratio of the total number of military deaths divided by the population of men born between 1865 and 1900 in the 1911 census. Death rates are non available in Moselle, Bas-Rhin and Haut-Rhin. Sample includes 90 *départements*.

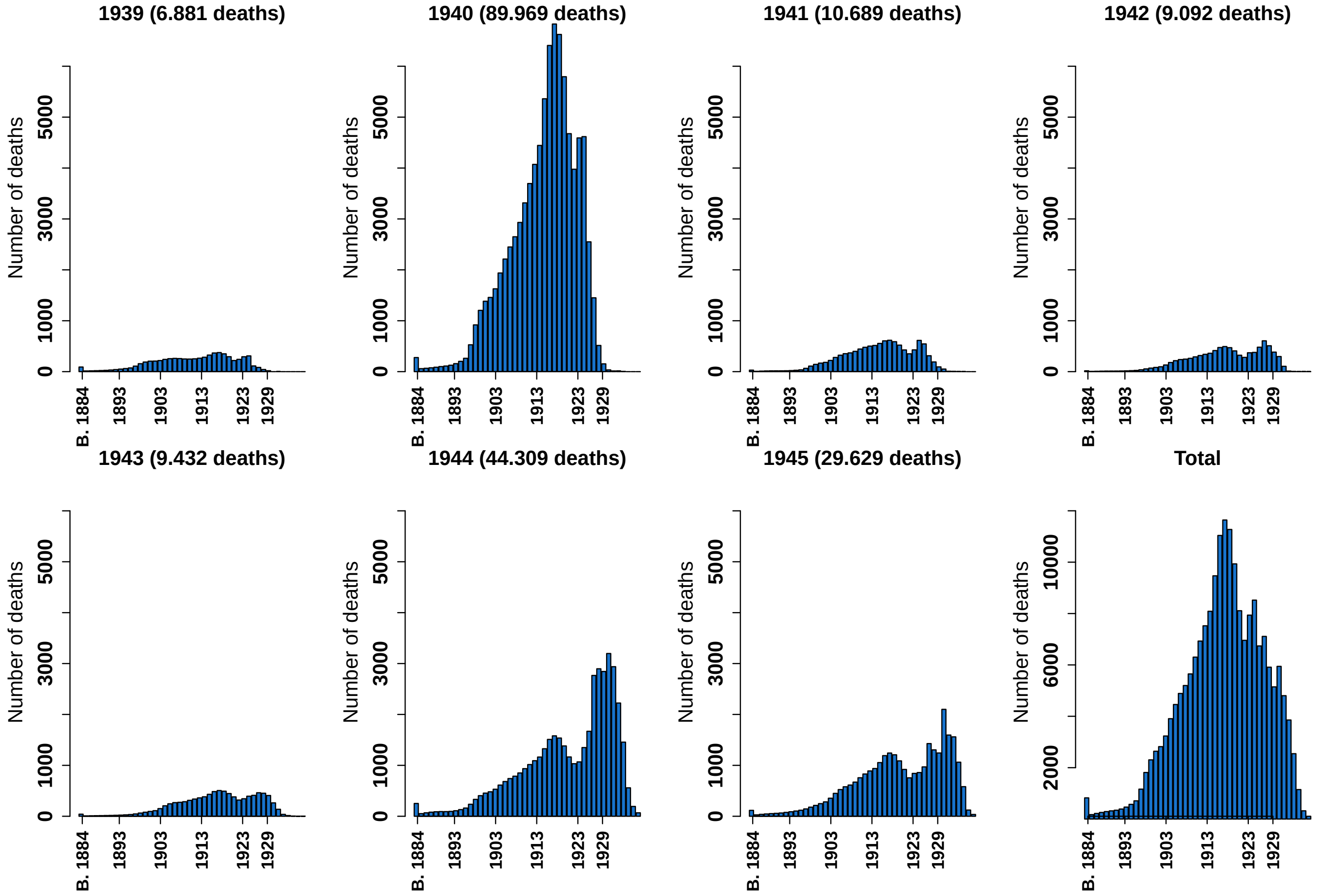
number of deaths by cohort for the whole Second World War. The total number of deaths per year is also indicated.

Figure 4 maps military death rate during the Second World War. It is computed by dividing the total number of deaths in each department by the population of men born between 1896 and 1925 and in 1936 census. Values are expressed per 100 men. At the national level, this value is 2.14%. One can see that military mortality was lower in Ile-de-France, in occupied départements of Bas-Rhin, Haut-Rhin and Moselle, in the Southeast and along the Garonne. The Mediterranean coast and Corsica have a particularly low death rate, less than 1.5%. Conversely, it was higher than 3.5% in Vosges and Haute-Saône, as well as Finistère, Morbihan and Côtes-d'Armor in Britanny.

3.2.3 Deaths in Deportation

Overview of the method According to deportation during the Second World War, deportees are classified by birth place in the database, 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

Figure 5: NUMBER OF DEATHS BY YEAR OF BIRTH AND YEAR OF DEATH IN FRANCE DURING THE SECOND WORLD WAR

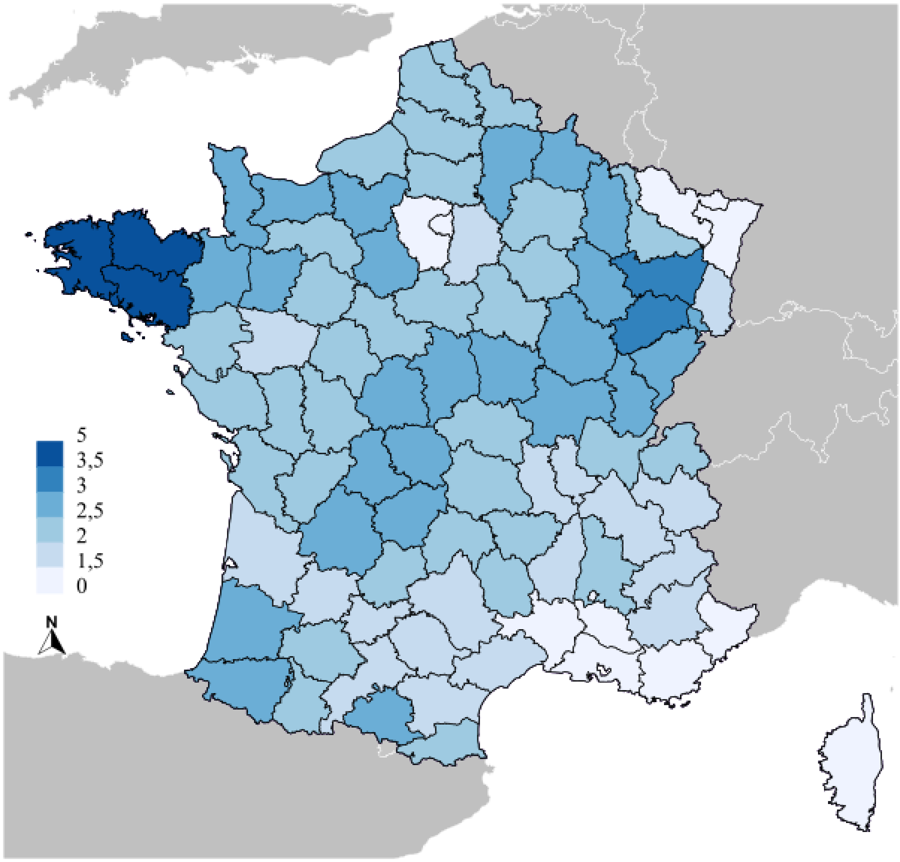


Note: Number of deaths by birth cohort and year of death in France during the Second World War. "B. 1884" means "cohorts born in 1884 and before".

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).

Deportees by sex, year of birth, year of death and birth place The deportee database is nominative (1 line for each deportee). Sex, birth department (or country of birth if born abroad), day-month-year of birth, day-month-year of death were extracted. The age of death in days-months-years follows. For dates of birth and death, data are kept since the year was available. Thus, if only the year was available, the date chosen was January 1st. Likewise, if only the month and year of birth were available, the full date of birth was set to the first day of the month. If the date was considered irrelevant (namely if the date of birth follows date of death), the date is erased. For individuals whose year of death was after 1946 (about forty individuals), I consider that those are unknown. 93% of the deceased have well-informed data for the four variables (sex,

Figure 6: MILITARY DEATH RATE (IN %) DURING THE SECOND WORLD WAR



Note: Military death rate is the ratio of the total number of military deaths divided by the population of men born between 1896 and 1925 in the 1936 census. Sample includes 90 *départements*.

date of death, age, place of birth). For those with one or more variable missing, data were not used. From these nominative data, I thus extract matrices crossing the age of death, the year of death (1940–1946), the place of birth and the sex.

One of the variables available in the deportee database is the place of birth. One has to differentiate this variable from the home place before deportation, that is where the deceased would have to be located in the lifetables. Since a 40-year-old have a non-zero probability to migrate in a different department from where he is born, I may infer the home-department before deportation. Similarly, deportees born abroad must be located in a French department.

From birth place to home place before deportation : deportees born abroad There are 33,609 deaths of born-abroad deportees, some 44% of the database. Those born outside France need to be allocated across France on the assumption that they immigrated before they were arrested and deported. One may suppose that these deportees born outside France fled Nazi persecution and settled in France before the start of the war. I make the assumption that the probability of being in each department can be infered by the spatial distribution of foreigners in 1936. Moreover, I assume that this distribution does not vary by age, and also that the 1936 distribution is representative of the war-time one. I can construct the following matrices:

1. N: Birth country $\times$ Age (90×105), retrieved from the "*MemorialGenWeb*" database, available for each sex and year between 1939 and 1946,

2. P: Department of residence $\times$ Birth country (91×48), available for census year 1936 and for the sum of men and women,

3. R: Department of residence $\times$ Age (90×105).

The first modification concerns *Seine*. Matrix P comprises 91 departments and not 90 because of the distinction between the city of Paris and the inner suburbs. These two lines are summed to get the same administrative boundaries in the two matrices. Next P is transformed so that the matrix gives the probability that an individual born in country i lives in department j.

Third, the names of countries of birth for Matrices P and N must be linked: there are 48 countries or regions in Matrix P and 90 countries in Matrix N. The level of detail in the "*MemorialGenWeb*" database is quite high, whereas the one in the census is lower : many Asian, South American and African countries are not directly specified, and colonies are often included in the generic term "French possessions in Africa". I reclassify them to calculate the product of Matrices N and P. Thus I get a Matrix $P^* (90 \times 90)$ and calculate the R Matrix for each sex and year between 1939 and 1946:

$$R' = N'P^*$$

Ultimately, each element R_{sj} in Matrix R corresponds to the sum of individuals aged s born in each of the countries i who emigrated to département j before being arrested and deported.

From birth place to home place before deportation : native deportees There are 43,055 deaths of French-born deportees in the database. I cannot assume that any deportee born in a department stayed in that department. A transfer matrix must therefore be constructed linking department of birth and department of residence before deportation. I use the matrix cross-referencing department of residence and department of birth in the 1946 census. This matrix distinguishes males and females. I assume both this matrix is representative of the pre-war situation and of deportee migrations, and that the probability of migration is equal for all ages.

I make a few preliminary modifications. The main is to allocate the deportees according to the post-1968 departments between *Seine* and *Seine-et-Oise*. The allocation key is the same as the one used for military deaths in the Second World War. I construct the following matrices:

1. N: Birth department × Age (90 × 105), retrieved from "*Mémoire des Hommes" database*, available for each sex and year between 1939 and 1946,

2. P: Department of residence × Department of birth (90 × 90), available for census year 1946 and for the sum of men and women,

3. R: Department of residence × Age (90 × 105).

P is transformed in P^* so that the matrix gives the probability that an individual born in *département i* lives in department j. Thus I calculate the R Matrix for each sex and year between 1939 and 1946:

$$R' = N'P^*.$$

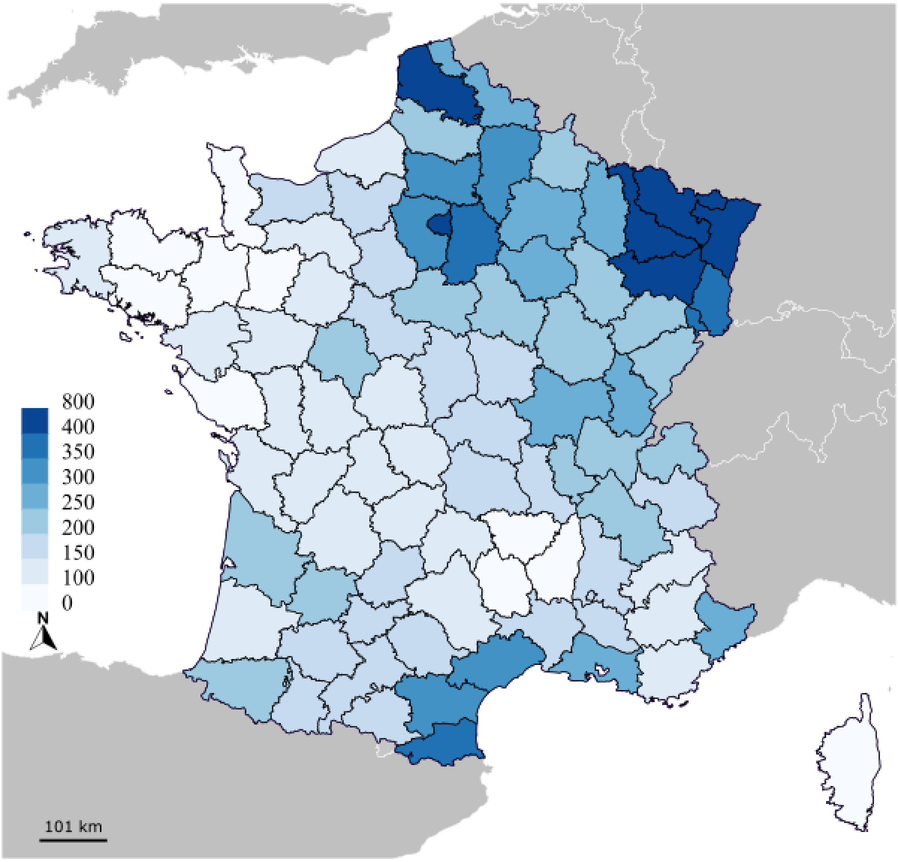
General adjustment The matrices of French and foreign-born deportees are finally summed. This final matrix is the sum for each department, each age, each sex and each year, of the departees born in a French department and deportees born outside France but living in France when they were arrested. For the total number of deportees, I based my computations on Dupaquier (1988). He reports 27,000 resisters dead in deportation and 83,000 jewish and other deportees. Consequently, I consider that 110,000 individuals died in the camps, compared to 76,664 included in the database (around 70%).

Figure 7 maps mortality due to deportation in French departments during the Second World War. The number of deportees is divided by the population of the department in the 1936 census. The indicator is expressed per 100,000 inhabitants. A dark color means that mortality by deportation was low, while a light color means that mortality by deportation was high. One can see that mortality was low in the west of the country, particularly in Brittany, and in the southeast of the Massif Central. The Paris region, the northern borders, and to a lesser extent the Mediterranean coast, were much more affected. Maximum values were reached in Moselle (735), Pas-de-Calais (520) and Seine (514). Note that 74% of the deportees were men.

3.2.4 Lexis Triangles

Civilian, military, and deportation deaths are added to get total of deaths by year, sex, single age, and department. These deaths 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 could 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 sufficiant. This is not, 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

Figure 7: DEATH RATE IN DEPORTATION (PER 100,000 INHABITANTS IN 1936)



Note: Death rate is the ratio of the total number of deportees divided by the population in the 1936 census. Death rates are expressed per 100,000 inhabitants. Sample includes 90 *départements*.

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. This equation takes into account the relative size of two successive cohorts, age, some historical events (e.g Spanish influenzia), and the infant mortality rate. If we call x the age and t the year, these sex-specific equations are as follows (Equation (5) for women, Equation (6) for men):

$$\begin{aligned}\hat{\pi}_d(x,t) =& 0.4710 + \hat{\alpha}_F + 0.7372\,[\pi_b(x,t) - 0.5] \\ & + 0.1025\, I_{t=1918} - 0.0237\, I_{t=1919} \\ & - 0.0112\, logIMR(t) - 0.0688\, logIMR(t)\, I_{x=0} + 0.0268\, logIMR(t)\, I_{x=1} \\ & + 0.1526\,[logIMR(t) - log(0.01)]\, I_{x=0} I_{IMR(t)<0.01}\end{aligned} \quad ; \qquad (5)$$

$$\begin{aligned}\hat{\pi}_d(x,t) =& 0.4836 + \hat{\alpha}_H + 0.6992\,[\pi_b(x,t) - 0.5] \\ & + 0.0728\,I_{t=1918} - 0.0352\,I_{t=1919} \\ & - 0.0088\,logIMR(t) - 0.0745\,logIMR(t)\,I_{x=0} + 0.0259\,logIMR(t)\,I_{x=1} \\ & + 0.1673\,[logIMR(t) - log(0.01)]\,I_{x=0} I_{IMR(t)<0.01}\end{aligned} \quad . \tag{6}$$

$\hat{\pi}_d(x,t)$ is defined as the proportion of death of a given year and age allocated in the lower triangle. α_F and α_H are age-specific values coming from the HMD protocol. $\pi_b(x,t)$ is defined as the ratio of births between two successive cohorts and calculated only once for both sexes:

$$\pi_b(x,t) = \frac{B(t-x)}{B(t-x) + B(t-x-1)}. \tag{7}$$

Long historical series are required to calculate this ratio for all the cohorts tracked between since 1901. One can take individuals aged 80 in 1901 as an example. To calculate this ratio one needs birth in 1820 and 1821. I was unable to do so since my birth records only go back to 1853. For earlier years I assume a birth ratio of 0.5.

$IMR(t)$, the same for both sexes, is calculated as follows:

$$IMR(t) = \frac{D(0,t)}{\frac{1}{3}B(t-1) + \frac{2}{3}B(t)}. \tag{8}$$

If births are not available for one of the two years, $IMR(t)$ is calculated as follows[8]:

$$IMR(t) = \frac{D(0,t)}{B(t^*)}, \tag{9}$$

with t^* the year for which births are available.

Note that I obtain proportions of deaths in the lower triangle greater than 1 for 28 female orbservations and 30 for 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 of 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 0 death in the upper triangle for these observations.

[8] When $IMR(t)$ is equal to zero because of no infant deaths, I assume a $0,00000001$ value so that $logIMR(t)$ can be calculated.

3.3 Population at Census

3.3.1 Distribution of Population of Unknown Age

Population of unknown year of birth in 1901 For the 1901 census, individuals whose birth year is unknown are put together in the open-age interval. To allocate them I use the 1911 census, which has a useful degree of detail. The process follows three steps.

First, I compute the quotient of individuals aged 95 and over by individuals aged 80 and over for each department i and each sex j in 1911:

$$R_{95ij}^{1911} = \frac{\sum_{s=95}^{105} P_{sij}^{1911}}{\sum_{s=80}^{105} P_{sij}^{1911}}. \tag{10}$$

Second I apply these quotients to the 1901 census to compute the proportion of individuals aged 95 and over among individuals aged 80 and over:

$$\sum_{s=95}^{105} P_{sij}^{1901} = R_{95ij}^{1911} \times \sum_{s=80}^{105} P_{sij}^{1901}. \tag{11}$$

Third I deduce population of unknown year of birth for each department and sex by substraction.

Distribution of population of unknown year of birth The raw statistics collected include an "unknown age" category in 1901, 1906, 1911, 1921, 1926, 1931, 1936 and 1946. In most cases, the share of population of unknown year of birth is very small (less than 1% of total population). This share exceeds 1% only in Corsica from 1901 to 1921 (between 1 and 2%). Population of unknown year of birth in each department and for each sex was distributed pro rata to the population in each group of year of birth.

3.3.2 Consistent Groups for Pre-1946 Censuses

Censuses from 1901 to 1946 did not use the same methodology for populations in the first three age groups. Some groups have to be combined or split, as shown in Table 3, in italics. For that purpose I assume that births were spread uniformly over time.

Finally, the 1911 census is rather different because it provides data for each birth year and not per five-year groups. However, these numbers fluctuate considerably. There were two possible methods: either use the numbers given, or combine the numbers in five-year groups as for the other censuses and apply cubic splines. Although the first method provides more information, it includes inconsistent fluctuations at adult ages. Since I need to maintain consistency, I choose the second method. Raw data in 1911 have to

TABLE 3: CLASSIFICATION AND AVAILABILITY OF POPULATIONS BORN TWO YEARS BEFORE THE CENSUS

Census	1^{st} class	2^{nd} class	3^{rd} class
1901	Born from 01/01/01 to 04/03/01	Born in 1900	Born in 1899
1906	Born from 01/01/06 to 03/06/06	Born in 1905	Born in 1904
1911	Born from 01/01/11 to 03/05/11	Born in 1910	Born in 1909
1921	Born from 01/01/21 to 03/05/21	*Born from 03/06/20 to 12/31/20*	*Born from 01/01/20 to 03/05/20*
1926	Born from 01/01/26 to 03/07/26	*Born from 03/08/25 to 12/31/25*	*Born from 01/01/25 to 03/07/25*
1931	Born from 01/01/31 to 03/07/31	*Born from 03/08/30 to 12/31/30*	*Born from 01/01/30 to 03/07/30*
1936	Born from 01/01/36 to 03/07/36	*Born from 08/03/35 to 31/12/35*	*Born from 01/01/35 to 7/03/35*
1946	*Born from 03/10/45 to 03/09/46*	*Born from 01/01/44 to 03/09/45*	Born in 1943

Note: Periods in italics in the table have to be combined or split to get populations by year of birth. 01/01/01 means 01/01/1901.

be thoroughly reprocessed: I keep the first fifteen birth year groups, and then combine them by five-year groups (1891–1895, 1886–1890, etc.), plus the open-age interval “1820 and earlier”.

3.3.3 Estimation of Population by Single Year of Birth

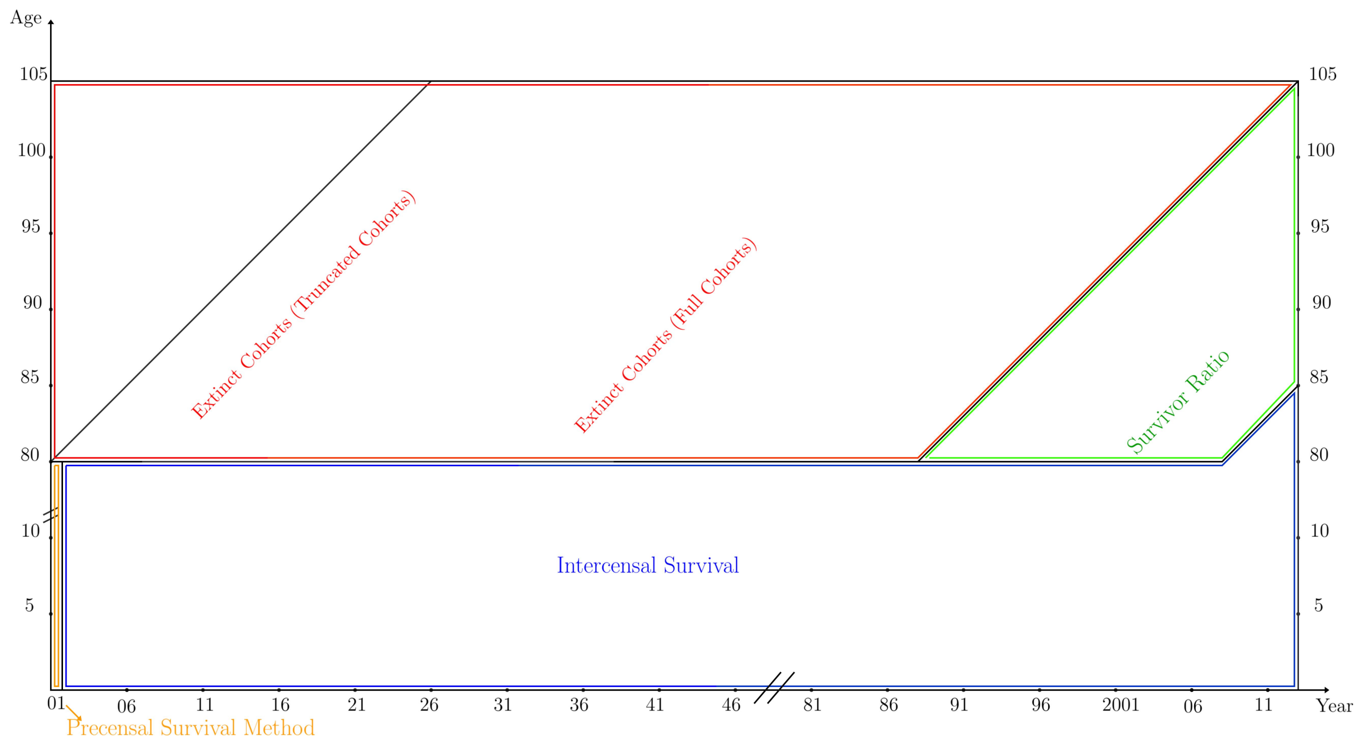
For my purposes it is simpler to compute population figures by year of birth. Nevertheless, populations available in censuses prior to 1968 are not available by single year of birth. Moreover, estimated populations from 2016 to 2021 are available by five-year age groups.

To get populations by single year of birth, I use the cubic spline method. The cubic splines are fitted to the cumulative curve of population born before 1^{st} January of the census year. For example, according to the 1901 census, I consider the population born before 1^{st} January 1901. The population born between 1^{st} January 1901 and the day of the census provides no further information and would involve fractional knots. The cubic splines adjustment takes into account that populations were given by age and not by birth year from 1968 onwards.

3.4 Population at 1^{st} January

I need populations by age at 1^{st} January for each year from 1901 onwards to calculate the mortality rates. I get populations by age at each 1^{st} January between 2013 and 2020 from official statistics so I may calculate populations by age at each 1^{st} January between 1901 and 2012. Figure 8 reveals the four methods used for various periods and ages.

Figure 8: METHODS FOR COMPUTATIONS OF POPULATION AT 1ST JANUARY



Note: Methods used to compute populations by age at each 1st January.

3.4.1 Intercensal Survival

The first method used to compute yearly populations on 1st January is called "Intercensal Survival". With this method I can estimate population by age (for those aged under 80) for each intercensal period. Populations at the second census (e.g. 1911 for 1906–1911) are not estimated in the same way for all cohorts. Figure 9 presents the three types of cohorts which exist in this method. There are "Pre-existing cohorts" (born before the census year), "Infant cohort" (born during the census year) and "Birth cohorts" (born after the census year). The gaps between the census date and 1st January of the census year are crucial. This gap is called f_1 for the first census and f_2 for the second.

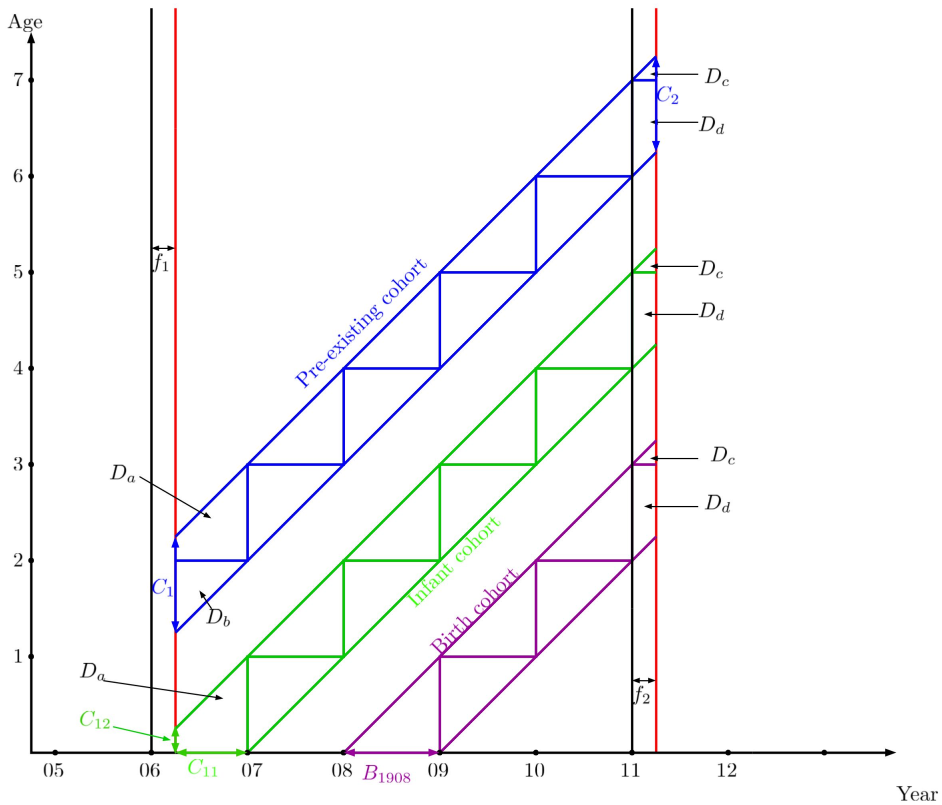
I begin with "Pre-existing cohorts". I estimate age-population at date of the second census. Let t and $t+N$ be the first and last 1st January in the intercensal period. N is the number of full calendar years between censuses. The dates of the two censuses are:

$$t_1 = t - 1 + f_1,$$

$$t_2 = t + N + f_2.$$

The elapsed time between the censuses is thus:

Figure 9: CLASSIFICATION OF DIFFERENT COHORTS FOR INTERCENSAL SURVIVAL METHOD



$$t_2 - t_1 = N + 1 - f_1 + f_2.$$

The cohort tracked (Figure 9, in blue) was 1- or 2-years-old at the time of the 1906 census and was born in 1904. The data is by birth year and not by age, which simplifies computations. I assume a uniform distribution of deaths in each Lexis triangle, so that for the cohort aged x on 1st January of the year of the first census,

$$D_a = (1 - f_1^2) \times D_L(x, t - 1),$$

$$D_b = (1 - f_1)^2 \times D_U(x - 1, t - 1),$$

$$D_c = f_2^2 \times D_L(x + N + 1, t + N),$$

$$D_d = (2f_2 - f_2^2) \times D_U(x + N, t + N).$$

This cohort's estimated population at the second census may be called $\hat{C}_2$ and is calculated as follows:

$$\hat{C}_2 = C_1 - (D_a + D_b) - \sum_{i=0}^{N-1} [D_U(x+i, t+i) + D_L(x+i+1, t+i)] - (D_c + D_d). \quad (12)$$

$\Delta_x = C_2 - \hat{C}_2$ is the difference between the estimated population and that recorded at the date of the second census. It comprises estimation errors and intercensal migrations within the cohort. In order to compute age-population at 1st January of each intercensal year, Δ_x must be split between the age-populations in each intercensal year. I asume that these rough migrations are uniformly distributed over time. Population by age is calculated as follows:

$$P(x+n, t+n) = C_1 - (D_a + D_b) - \sum_{i=0}^{n-1} [D_U(x+i, t+i) + D_L(x+i+1, t+i)] + \frac{1 - f_1 + n}{N + 1 - f_1 + f_2} \Delta_x. \quad (13)$$

There is only one "Infant cohort" to track for each intercensal period: in Figure 9, it is the cohort born in 1906. Thus, $C_1 = C_{11} + C_{12}$, with $C_{11} = (1 - f_1) \times B_{t-1}$ and C_{12} the population recorded as born during the year of the census. Thus,

$$\hat{C}_2 = C_1 - D_a - \sum_{i=0}^{N-1} [D_U(i, t+i) + D_L(i+1, t+i)] - (D_c + D_d), \quad (14)$$

and

$$P(n, t+n) = C_1 - (D_a + D_b) - \sum_{i=0}^{n-1} [D_U(i, t+i) + D_L(i+1, t+i)] + \frac{\frac{1}{2}(1 - f_1^2) + n}{N + \frac{1}{2}(1 - f_1^2) + f_2} \Delta_0. \quad (15)$$

Finally, since N is the number of full calendar years during the intercensal interval, I track N birth cohorts. A cohort born in year $t + j$ is aged $K = N - j - 1$ on $01/01/t + N$. The estimated population of this cohort may be expressed as:

$$\hat{C}_2 = B_{t+j} - D_L(0, t+j) - \sum_{i=1}^{N-1} [D_U(i-1, t+j+i) + D_L(i, t+j+i)] - (D_c + D_d). \quad (16)$$

Note that the number of intermediate populations produced by the various cohorts depends on K. For $k = 0, ..., K$, the intermediate populations of each cohort are computed as follows:

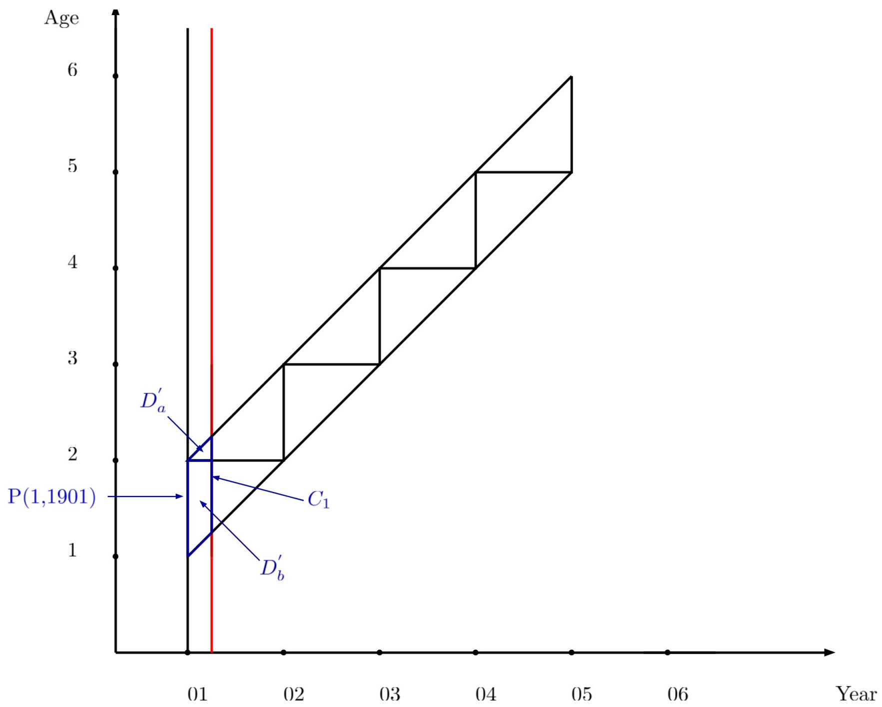
$$P(k,t+j+k+1) = B_{t+j} - D_L(0,t+j) - \sum_{i=1}^{k} [D_U(i-1,t+j+i) + D_L(i,t+j+i)] + \frac{2k+1}{2K+1+2f_2}\Delta_{t+j}. \tag{17}$$

3.4.2 Precensal Survival Method

The second method I use is "Precensal Survival", to compute populations for the first 1^{st} January of the whole period. Figure 10 presents the computations for population of age 1 in 1901. To do so, I must add D'_a et D'_b to the population born in 1901 and recorded on 6 March 1901. If t_1 is the first 1^{st} January of the intercensal period, then:

$$P(x-1,t_1-1) = C_1 + D'_a + D'_b. \tag{18}$$

Figure 10: PRECENSAL SURVIVAL METHOD



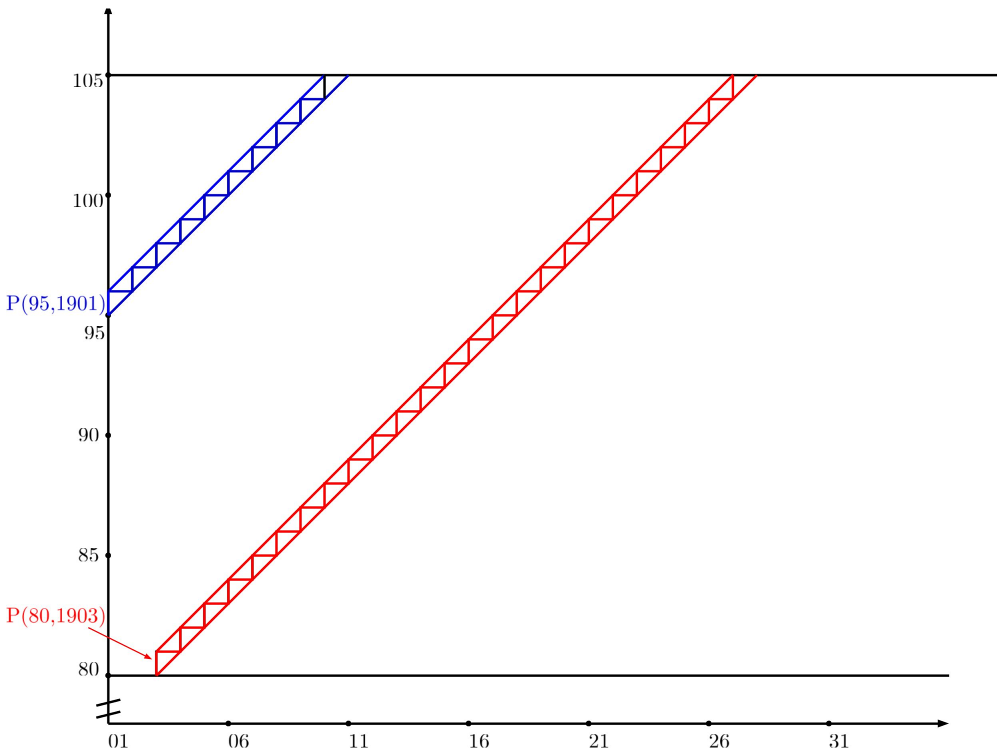
3.4.3 Extinct Cohorts Method

The third method I use is "Extinct Cohorts", to calculate age-population for all the cohorts extincted in 2013. Since the maximum age in my database is 105, a cohort is considered to be extinct if it reached 105 or over in 2013.

Figure 11 reveals that my data comprises two kinds of extinct cohorts. The first are "Full cohorts" (Figure 11, in red), which can be tracked from ages 80 to 105 in 1901–2013. Thus, the 80-year-old population in 1903 equals the sum of the cohort's Lexis triangles between ages 80 and 105. The others are "Truncated cohorts" (Figure 11, in blue), aged 80 and over in 1901. Thus, the 95-year-old population in 1901 equals the sum of the cohort's Lexis triangles between 95 and 105. More generally, the population of age x in year t can be calculated as follows:

$$P(x,t) = \sum_{i=0}^{\infty} \left[D_U(x+i, t+i) + D_L(x+i, t+i) \right].$$

Figure 11: EXTINCT COHORTS METHOD



3.4.4 Survivor Ratio Method

The last method I use is "Survivor ratio", to calculate non-extinct cohorts of age 85 and over in 2013. Figure 12 presents the computations for the cohort aged 104 in 2013. The survivor ratio R may be defined as the number of individuals alive at age x on 1st January t, divided by the number of individuals in the same cohort alive k years previously. Formally:

$$R = \frac{P(x,t)}{P(x-k, t-k)}.$$

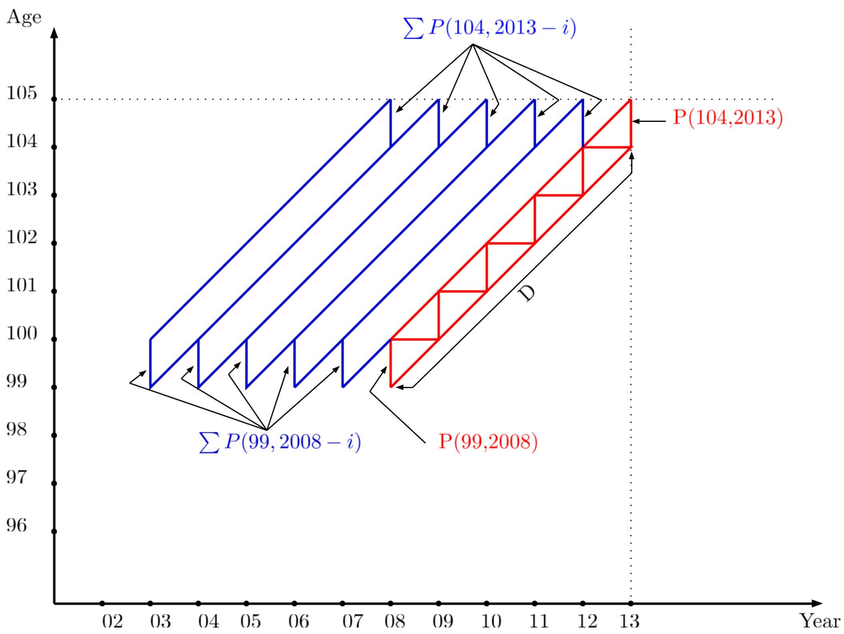
I assume that there is no migration at these ages. R may also be expressed:

$$R = \frac{P(x,t)}{P(x,t) + \dot{D}}.$$

where $\dot{D} = \sum_{i=1}^{k} [D_U(x-i, t-i) + D_L(x-i+1, t-i)]$. Finally, $P(x,t)$ may be expressed as a function of R:

$$P(x,t) = \frac{R}{1-R}\dot{D}. \tag{19}$$

Figure 12: SURVIVOR RATIO METHOD



Since the survivor ratio cannot be directly observed for a cohort, I use preceding cohorts whose age-populations have been calculated by the "Extinct Cohorts" method. I asume that the survival ratio has roughly the same value in the studied cohort and in the preceding ones. As such, the mean ratio R^* of the preceding m cohorts may be calculated as follows:

$$R^*(x, 2013, k, m) = \frac{\sum_{i=i}^{m} P(x, 2013-i)}{\sum_{i=i}^{m} P(x-k, 2013-k-i)}.$$

I may then estimate $\tilde{P}(x, 2013)$:

$$\widetilde{P}(x, 2013) = \frac{R^*}{1-R^*}\dot{D}.$$

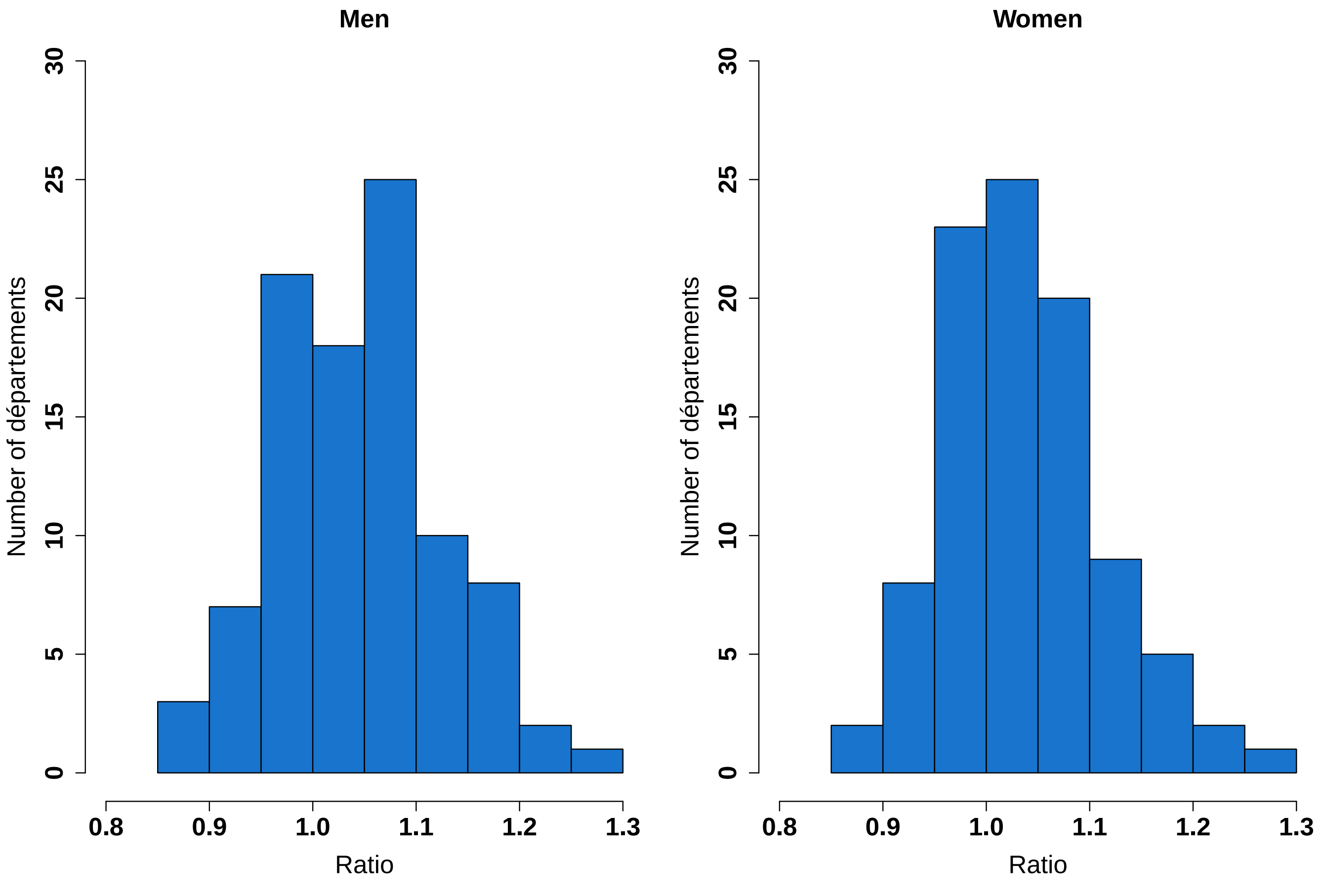
Subsequently, I may track the cohort back in time and estimate $\tilde{P}(x-1, 2012)$, $\tilde{P}(x-2, 2011)$, ... by adding step by step the cohort's deaths. I apply this method for any non-extinct cohort in 2013. For my estimations I follow the guidelines of the HMD Protocol, with $k = m = 5$.

The assumption of a constant survivor ratio over time is strong; I may control by the recorded population on 1st January 2013. I compare the 85-and-over population on 1st January 2013 – retrieved from the census of that year and called P^{Rec}_{85+} – with the 85-and-over population on 1st January 2013, as calculated by the Survivor Ratio method and called P^{SR}_{85+}. Thus, populations at each age in 2013 can be computed as follows:

$$\hat{P}(x,2013) = c\widetilde{P}(x,2013) = c\frac{R^*}{1-R^*}\dot{D},$$

where $c = \frac{P^{Rec}_{85+}}{P^{SR}_{85+}}$. Figure 13 reveals the departmental distributon of c in 2013.
As before, each cohort is then back-followed: I make estimates for $\hat{P}(x-1,2012), \hat{P}(x-2,2011), \ldots$

Figure 13: RATIO c OF POPULATION AGED 85 AND OVER ESTIMATED WITH SURVIVOR RATIO METHOD AND POPULATION AGED 85 AND OVER AVAILABLE IN CENSUS IN 2013



Note: Ratio c equals to the quotient of population aged 85 and over in 2013 census by population estimated with "Survivor Ratio" method in 2013.

3.5 Lifetables

3.5.1 Mortality Rates Adjustments

I can compute departmental mortality rates by age and sex with deaths in Lexis triangles and populations at each 1st January. Moreover, I sum deaths and populations to compute lifetables for both sexes combined.

Mortality rates are the ratio between the number of deaths and the number of individuals exposed to the risk:

$$M_{xt} = \frac{D_{xt}}{E_{xt}} = \frac{D_L(x,t) + D_U(x,t)}{\frac{1}{2}\left[P(x,t) + P(x,t+1)\right] + \frac{1}{6}\left[D_L(x,t) - D_U(x,t)\right]}. \tag{20}$$

For some cases, I have not $P(x,t+1)$. To estimate mortality rates for t, I assume that the population at each age in $t+1$ is equal to the one in t, and the formula becomes:

$$M_{xt} = \frac{D_{xt}}{E_{xt}} = \frac{D_L(x,t) + D_U(x,t)}{P(x,t) + \frac{1}{6}\left[D_L(x,t) - D_U(x,t)\right]}. \tag{21}$$

These rates are not used directly to calculate lifetables. I smooth mortality rates beyond age 90 in order 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 $b \geq 0$):

$$\mu_x(a,b) = \frac{ae^{b(x-80)}}{1 + ae^{b(x-80)}}. \tag{22}$$

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

$$M_x(a,b) = \mu_{x+0,5}(a,b). \tag{23}$$

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:

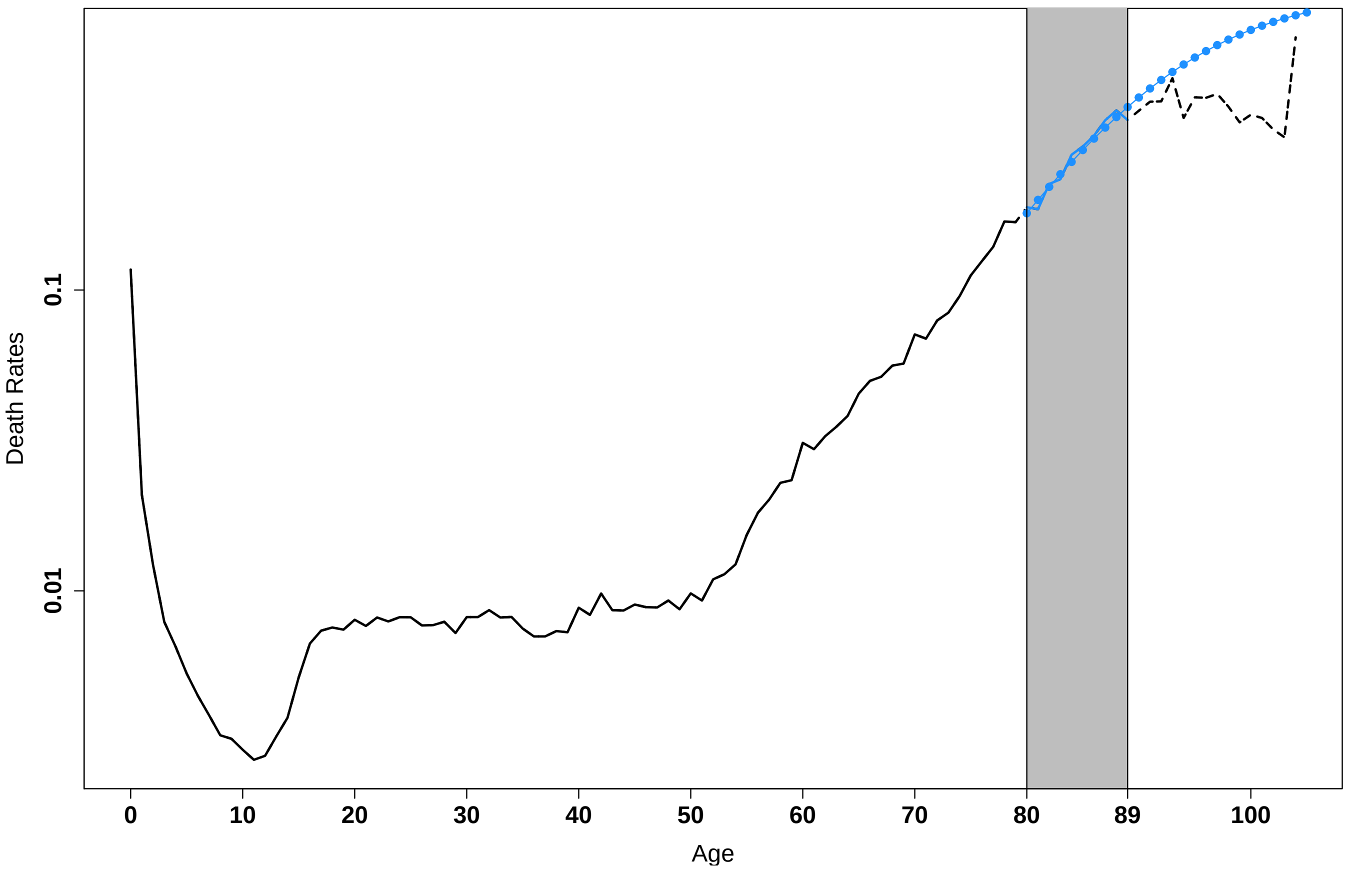
$$-logL(a,b) = \sum_{x=80}^{105} \left[D_x log\mu_{x+0,5}(a,b) - E_x\mu_{x+0,5}(a,b)\right]. \tag{24}$$

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):

$$\begin{cases} m_x = M_x & x \in [0,89] \\ m_x = \hat{M}_x & x \in [90,105] \end{cases}. \tag{25}$$

Figure 14 presents this adjustment for male mortality in the department of Ain in 1901. Mortality rates computed using raw data are in solid or dashed lines, while mortality rates estimated using Kannisto model for ages 80 and more are in blue dotted line. Blue solid line represents mortality rates between age 80 and 89 used in the Kannisto model to estimate mortality rates beyond age 90 and replace raw mortality rates in dashed black line, which are highly volatile.

Figure 14: EXAMPLE OF COMPUTATIONS OF DEATH RATES ABOVE AGE 90



Note: Mortality rates for males in the department of Ain in 1901. Raw mortality rates in solid or dashed lines. Mortality rates estimated using raw mortality rates between ages 80 and 89 and Kannisto model in blue dotted line.

3.5.2 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:

$$\begin{cases} a_x = 1/2 & x \in [1, 104] \\ a_x = \frac{1}{m_{105}^{\infty}} & x = 105+ \end{cases} . \tag{26}$$

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

$$
\begin{cases}
m_0 \geq 0.107 & \begin{cases} a_0 = 0,350 & \text{for women,} \\ a_0 = 0,330 & \text{for men,} \end{cases} \\
m_0 < 0.107 & \begin{cases} a_0 = 0,053 + 2.800 m_0 & \text{for women,} \\ a_0 = 0,045 + 2.684 m_0 & \text{for men.} \end{cases}
\end{cases}
\tag{27}
$$

The probabilities of death may be calculated as follows:

$$
\begin{cases}
q_x = \frac{m_x}{1+(1+a_x)m_x} & x \in [0,104] \\
q_x = 1 & x = 105+
\end{cases}
.
\tag{28}
$$

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) i. e. for each age and each group of 5 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[, ... etc until ages 105 and over. Values in abridged lifetables are computed from previous variables.

3.6 Territorial Changes and Missing Data

3.6.1 General Overview

The main advantage of the French departments is their stability since the beginning of the 19$^{\text{th}}$ 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 (*Corse* counting as one), as well as the *Seine* and *Seine-et-Oise* departments which existed before 1968. Territorial breakdowns are twofold in this study: either departmental boundaries changed because of a territorial reorganization, or the data is missing within the unified departmental classification that I use.

Territorial changes The departmental boundary changes are of two types. The first concerns the pre-1901 period. *Savoie* and Nice's *Comté* were attached to France following the 22 and 23 April 1860 *plébiscite*. *Savoie* and *Haute-Savoie* were created ex nihilo on 14 June 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*.[9] In addition, the department boundaries of *Haut-Rhin*[10], *Bas-Rhin* and *Vosges*[11] changed. For this period, I distributed births of the old-classification departments between the unified classification of departments.

The second change concerns the post 1901 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 *départements* were divided between *Paris, Yvelines, Essonne, Hauts-de-Seine, Seine-Saint-Denis, Val-de-Marne* and *Val d'Oise*.

Missing data The missing data in the unified departmental classification is also of two types. The first concerns 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, and *Moselle*, *Bas-Rhin* and *Haut-Rhin* for the period 1939–1945. *Corse* is also concerned in 1943 and 1944. The second category is *départements* temporarily under German control, namely *Bas-Rhin*, *Haut-Rhin* and *Moselle* before 1919.

Computation periods Finally, computation periods vary by department. I distinguish them according to four classes, as one can see in Figure 15.

C_1 All départements except *Moselle, Bas-Rhin, Haut-Rhin, Seine-et-Oise* and *Ile-de-France* (except *Seine-et-Marne*). These 85 departments are tracked from 1901 onwards. Computations of population at each 1^{st} January are done as shown in Figure 8.

C_2 *Seine* (75) and *Seine-et-Oise* (78). The lifetables for these departments are estimated for the period 1901–1968. Figure 16 presents the methods used to compute populations at each 1^{st} January. One can see that the "Survivor Ratio" method is applied to the 1968 census and not the 2013 census.

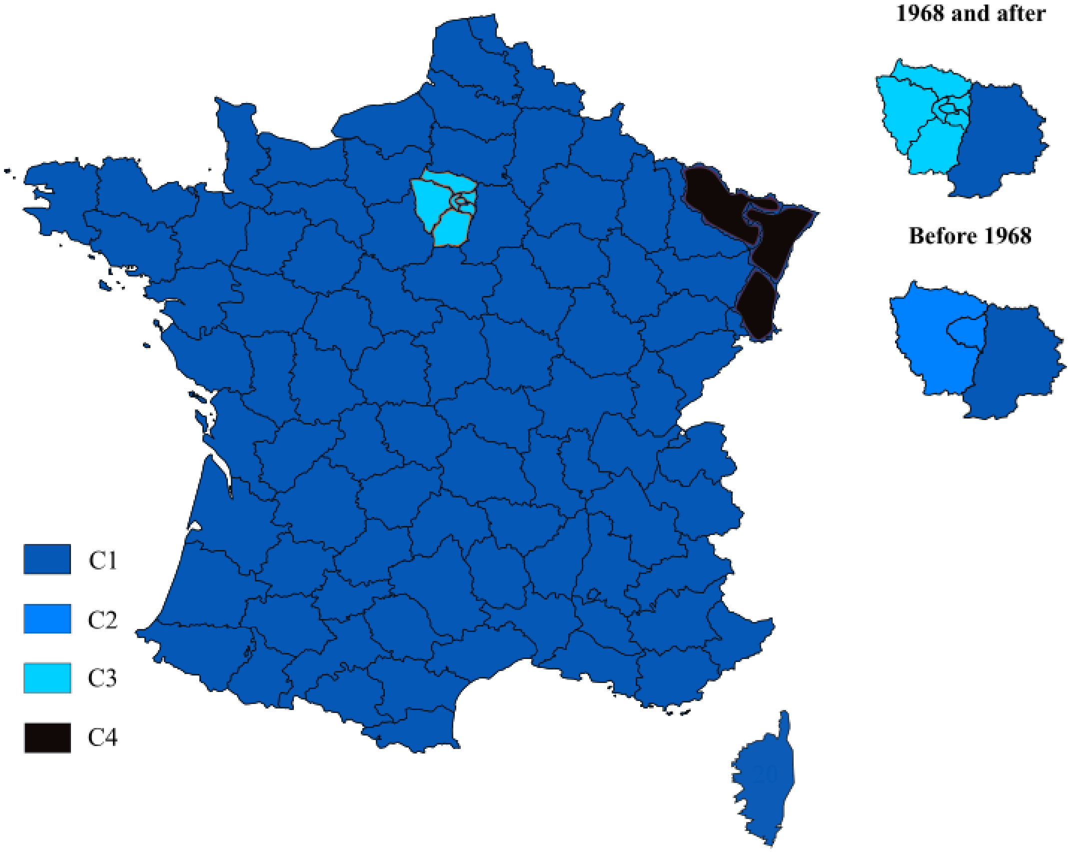
C_3 The new departments in *Ile-de-France: Essonne* (91), *Hauts-de-Seine* (92), *Seine-Saint-Denis* (93), *Val-de-Marne* (94), *Val d'Oise* (95), *Paris* (96), *Yvelines* (97). These lifetables are estimated from 1968 onwards. Figure 17 presents the methods used to compute populations at each 1^{st} January.

[9]Until 1870, two departments existed, namely *Meurthe* and *Moselle*. Their gathering 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 T*hionville, 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, i.e. the districts of *Luneville*, *Nancy* and *Toul* for the old *Meurthe* and the canton of *Briey* for the old *Moselle.*

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

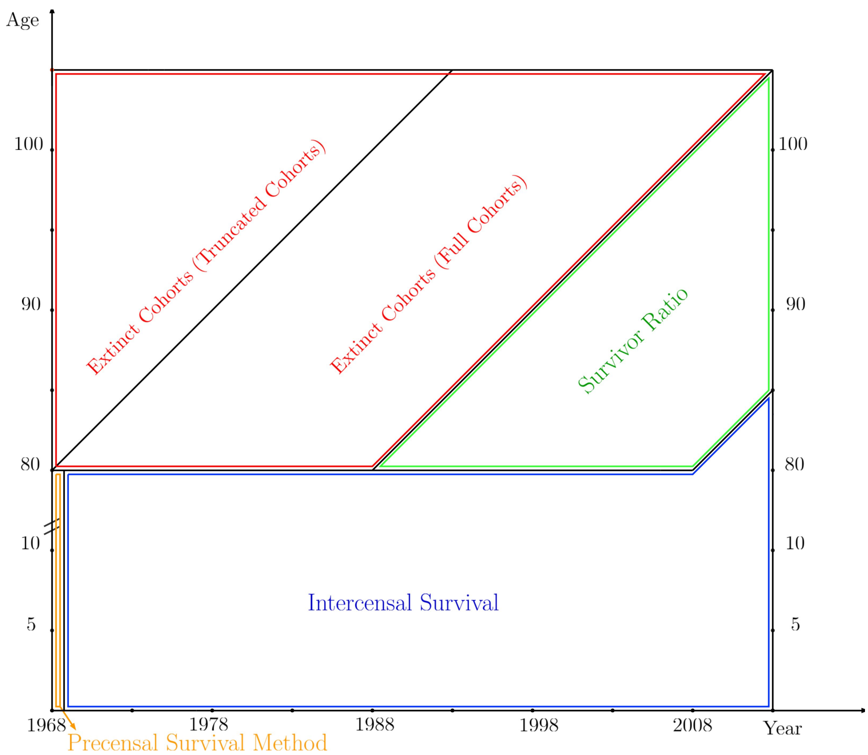
[11]In 1870, the former cantons of *Schirmeck* and *Saales* (in *Vosges*) were attached to *Bas-Rhin*, under German control. The new boundaries of these two *départements* are those that we know nowadays.

Figure 15: CLASSIFICATION OF DEPARTMENTS ACCORDING TO THE YEARS AVAILABLE IN THE FRENCH HUMAN MORTALITY DATABASE



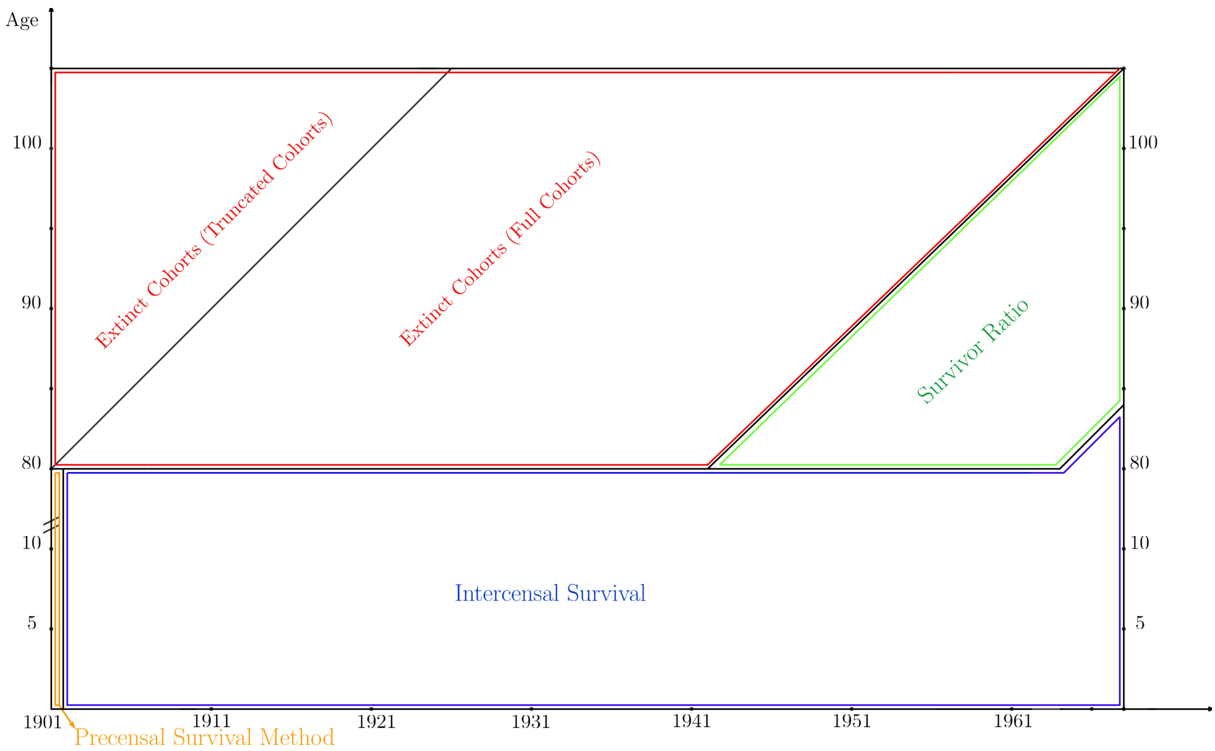
Note: Lifetables for dark blue departments available from 1901 onwards (Class C_1). Lifetables for medium blue departments available between 1901 and 1968 (Class C_2). Lifetables for light blue departments available from 1968 onwards (Class C_3). Lifetables for black departments available from 1921 onwards (Class C_4).

Figure 17: ESTIMATIONS OF POPULATIONS FOR DEPARTMENTS OF CLASS C_3



Note: Methods used to compute populations by age at each 1st January for departments of class C_3, namely *Essonne* (91), *Hauts-de-Seine* (92), *Seine-Saint-Denis* (93), *Val-de-Marne* (94), *Val d'Oise* (95), *Paris* (96), *Yvelines* (97).

Figure 16: ESTIMATIONS OF POPULATIONS FOR DEPARTMENTS OF CLASS C_2



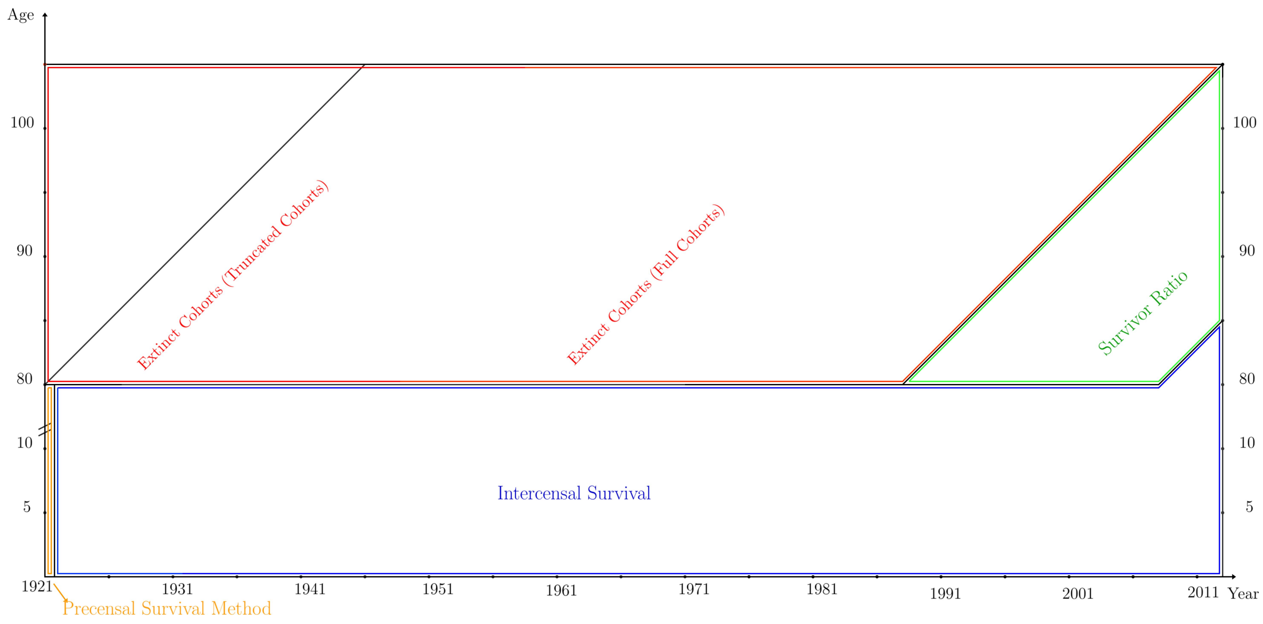
Note: Methods used to compute populations by age at each 1st January for departments of class C_2, namely *Seine* (75) and *Seine-et-Oise* (78).

C_4 *Moselle, Bas-Rhin* and *Haut-Rhin*. From 1870 to 1918 these three departments were under German administration. Consequently, the public records were not kept by the French authorities. I have not been able to do research in Germany to find data for this territory, so my estimates begin at the first available census, namely 1921, as shown in Figure 18.

3.6.2 Estimation of Births during the 19th Century

Births of the missing departments during the period 1853–1900 are first estimated. Recall that 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 along 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 *départements*), 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.

Figure 18: COMPUTATIONS OF POPULATIONS FOR DEPARTMENTS OF CLASS C_4



Note: Methods used to compute populations by age at each 1st January for departments of class C_3, namely *Moselle* (57), *Bas-Rhin* (67) and *Haut-Rhin* (68).

3.6.3 Vital Statistics during the Two World Wars

Data from the population movement for missing departments during the two World Wars is also estimated. There are 10 departments (*Aisne, Ardennes, Marne, Meurthe-et-Moselle, Meuse, Nord, Oise, Pas-de-Calais, Somme, Vosges*) with missing data during the First World War, and 4 during the Second World War (*Corse* between 1943 and 1944, *Moselle, Bas-Rhin, Haut-Rhin* between 1939 and 1945). These missing data are of 2 types: births and stillbirths, as well as deaths. 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 during the 19th Century, and endogenize the choice of the reference department. The general assumption used for the estimations of these missing data is that the neighbours departments have similar evolutions concerning their demographic variables, because of their culture and their shared living conditions. For each couple of department × missing period, I choose a panel of geographically close departments whose data is available. Table 4 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 *département* is defined for each department with missing data and used to estimate these values. This method is used to total births, stillbirths, deaths by age of men and deaths by age of women (sum of civilian, military and in deportation deaths). Table 5 gives the reference *département* for each missing *département* and each variable.

Table 4: 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*

Note: Departments with missing data, by period, and panel of potential reference departments for each department with missing data.

Table 5: DEPARTMENTS WITH MISSING VALUES AND REFERENCE DEPARTMENTS USED

Period	Missing departments	Reference departments used			
		Stillbirths	Births	Deaths of men	Deaths of women
1870-1871	*Seine*	*Seine et Oise*	*Seine et Oise*		
1914-1919	*Aisne*	*Aube*	*Haute-Marne*	*Eure*	*Haute-Marne*
1914-1919	*Ardennes*	*Aube*	*Haute-Marne*	*Eure*	*Haute-Marne*
1914-1919	*Marne*	*Haute-Saône*	*Haute-Saône*	*Haute-Marne*	*Haute-Marne*
1914-1919	*Meurthe-et-Moselle*	*Haute-Saône*	*Haute-Saône*	*Seine Inférieure*	*Seine Inférieure*
1914-1919	*Meuse*	*Seine et Marne*	*Haute-Saône*	*Eure*	*Haute-Marne*
1914-1919	*Nord*	*Haute-Saône*	*Seine et Marne*	*Eure*	*Seine Inférieure*
1914-1919	*Oise*	*Aube*	*Haute-Marne*	*Haute-Marne*	*Seine Inférieure*
1914-1919	*Pas-de-Calais*	*Aube*	*Haute-Saône*	*Eure*	*Seine Inférieure*
1914-1919	*Somme*	*Aube*	*Haute-Marne*	*Eure*	*Seine Inférieure*
1914-1919	*Vosges*	*Haute-Saône*	*Haute-Saône*	*Haute-Saône*	*Seine Inférieure*
1939-1945	*Moselle*	*Meurthe-et-Moselle*	*Meurthe-et-Moselle*	*Meurthe-et-Moselle*	*Meurthe-et-Moselle*
1939-1945	*Bas-Rhin*	*Meurthe-et-Moselle*	*Meurthe-et-Moselle*	*Meurthe-et-Moselle*	*Meurthe-et-Moselle*
1939-1945	*Haut-Rhin*	*Meurthe-et-Moselle*	*Meurthe-et-Moselle*	*Meurthe-et-Moselle*	*Vosges*
1943-1944	*Corse*	*Alpes Maritimes*	*Var*	*Hérault*	*Bouches du Rhone*

Note: Departments with missing data by period and variable, and reference departments used to estimate missing data.

Births and stillbirths The choice of the reference department for each missing department and each sub-period must consider how their demographic variables were synchronized. For that purpose, I may define a support interval and then track changes in the ratio between the variable in the missing department and the variable in the reference department during that interval. Let t_1 and t_2 be the first and last years of the subperiod for which there are missing data, $\Omega_\Delta = [t_1, t_2]$ the subperiod for which there are missing data, $\Omega_t = [t_1 - h, t_1] \cup [t_2, t_2 + h]$ the support interval with $h = 4$, i the missing department, j the potential reference department. The ratio R_{ij}^t is calculated for a demographic variable V:

$$R_{ij}^t = \frac{V_j^t}{V_i^t}, t \in \Omega_t \quad .$$

The mean $\bar{x}_{ij}$ and the standard deviation σ_{ij} of R_{ij}^t are calculated over the interval Ω_t. The stability of the ratio is measured as the coefficient of variation of R_{ij}^t over the interval Ω_t:

$$CV_{ij} = \frac{\sigma_{ij}}{\bar{x}_{ij}}.$$

The reference department j^* chosen is the one with the lowest coefficient of variation among all the possible reference departments. This criterion is used for both stillbirths and births. After choosing the reference department for each missing department, the missing data for department i and variable V is estimated as follows:

$$V_i^t = V_{j^*}^t \times \bar{x}_{ij*}.$$

Deaths The method used to estimate missing deaths is similar to the one used for stillbirths and births. Note that computations are made for total deaths (including military deaths and deportees). Let t_1 and t_2 be the first and last years of the subperiod for which there are missing data, $\Omega_\Delta = [t_1, t_2]$ the subperiod for which there are missing data, $\Omega_t = [t_1 - h, t_1] \cup [t_2, t_2 + h]$ the support interval with $h = 4$, i the missing department, j the potential reference department. The ratio R_{xij}^t is calculated for deaths D at age x:

$$R_{xij}^t = \frac{D_{xj}^t}{D_{xi}^t}, t \in \Omega_t.$$

The mean $\bar{x}_{xij}$ and the standard deviation σ_{xij} of R_{xij}^t are calculated over the interval Ω_t. The stability of the ratio is measured as the coefficient of variation of R_{xij}^t over the intervalΩ_t:

$$CV_{xij} = \frac{\sigma_{xij}}{\bar{x}_{xij}}.$$

The fit between missing department and reference department needs to take the lowest value of the coefficient of variation over a number of ages Ω_x and not a single point. I calculate a score S_{ij}:

$$S_{ij} = \frac{1}{\Omega_x} \sum_{x \in \Omega_x} CV_{xij},$$

where Ω_x is defined as ages 0–4 and 50–89 in order to avoid erratic results due to small number of deaths.

The reference department j^* chosen is the one with the lowest score among all the possible reference departments. After choosing the reference department for each missing department and subperiod, deaths at age x for the department i are estimated as follows:

$$D^t_{xi} = D^t_{xj^*} \times \bar{x}_{xij^*}.$$

3.6.4 Changes in Ile-de-France region

By changing the three departments of *Ile-de-France* (*Seine, Seine-et-Marne, Seine-et-Oise*) in eight new ones (*Paris, Seine-et-Marne, Yvelines, Essonne, Hauts-de-Seine, Seine-Saint-Denis, Val-de-Marne, Val-d'Oise*), the reorganization of this region in 1968 creates a discontinuity in the data. I change my methodology so as to track each of these departments over the most appropriate period. Note that *Seine-et-Marne* was not affected by these changes. When I talk about *Ile-de-France* hereafter, I mean the *Ile-de-France* region less *Seine-et-Marne*.

For the intercensal period 1901–1962, I can track the old departments since I have all the censuses between these years and vital statistics. From 1968 onwards, I can track the new departments since I have all the censuses and population flows. For the intercensal period 1962–1968, I have 1962 and 1968 censuses for the new departments, but no population flows. For the same intercensal period, I have population flows and the 1962 census for the old departments, but no data according to the 1968 census. I choose to track the old departments until 1968, and the new ones from 1968 onwards. To do so, I make two adjustments. The first is about pre-1968 births for the new departments, useful to split deaths in Lexis triangles. The second is about populations of the old departments in 1968, to estimate the 1st January population of these departments between 1962 and 1968.

To estimate births of the new departments before 1968, I use the 1968 distribution. I assume that the weight of each department remains constant. Although this is a strong assumption if one want to know the accurate number of births, it is less strong for the relative size of two successive cohorts.

I am not able to calculate 1st January populations of the 1962–1968 intercensal period for *Seine* and *Seine-et-Oise*. Indeed, the turning census available for both old and new departments is the 1962 one. In order to estimate pre-1968 population, one need population aged 85 and over to implement the "Survivor Ratio" method, and populations aged 0 to 84 to implement the "Intercensal Survival" one. To estimate the population aged 85-and-over for *Seine* and *Seine-et-Oise*, I assume that the weight of the two departments in the population of age 85 and over in Ile-de-France did not vary between 1962 and 1968. It is more difficult concerning the population aged 0 to 84. To do so, I draw on the Intercensal Survival method. First, I calculate the estimated population in 1968 for *Seine* and *Seine-et-Oise* and the sum of these two departments $\hat{P}^{68}_{IdF}(x)$, by subtracting from each cohort counted in 1962 deaths occurring during the intercensal period. I also know the population estimated for these two *départements* in 1968 (called $P^{68}_{IdF}(x)$) by summing the new departments. I can therefore deduce the migratory profile for *Ile-de-France*:

$$R_{IdF}^{68}(x) = \frac{\hat{P}_{IdF}^{68}(x)}{P_{IdF}^{68}(x)}.$$

I assume this profile was similar for each of the old departments j and use this migratory profile to compute 1968 census populations:

$$P_j^{68}(x) = R_{IdF}^{68}(x) \times \hat{P}_j^{68}(x).$$

Figure 19 reveals the population of the Seine and Seine-et-Oise departments in 1962 (in black) and 1968 (in blue) for men (top left) and women (top right). It also represents the estimated age migration profile (in values) between 1962 and 1968 for men (bottom left) and women (bottom right). It is interesting to recall that the 1962 population statistics are available by five-year birth groups for those born before 1948 (aged 15 and over), whereas the 1968 population statistics are available by single age. Therefore, estimates of the migration pattern by single age for cohorts born during First World War are less reliable.

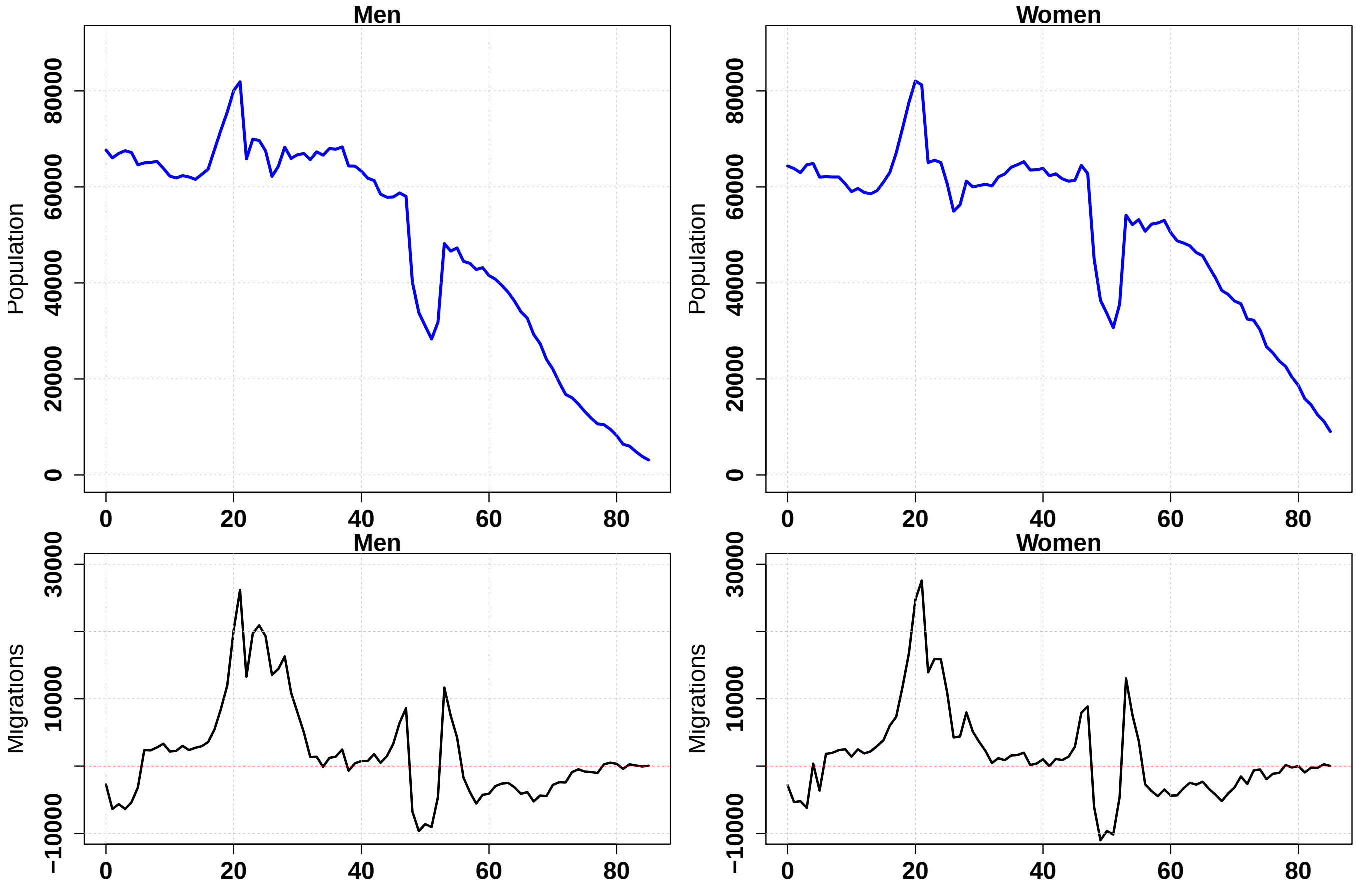
4 Available Results and Discussion

4.1 Available Results

Lifetables are available for 3 different geographical levels corresponding to the 3 levels of the Nomenclature of Territorial Units for Statistics: the departments of the current classification (*Corse* counting as one) as well as the old *Seine* and *Seine-et-Oise* (NUTS 3 level, 97 units), the administrative regions that existed between 1970 and 2015 (NUTS 2 level, 22 units) and the current administrative regions (NUTS 1 level, 13 units). Figures 20 and 21 present these 3 geographical levels in an extensive manner. Note that lifetables are available from 1921 onwards in the regions *Alsace* and *Lorraine* (NUTS 2 level) and in the region *Grand-Est* (NUTS 1 level) since lifetables are available only during this period for the departments of Moselle, Bas-Rhin and Haut-Rhin. Lifetables are available for men, women, and both sexes combined.

To illustrate these results, I present in Figure 22 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 South-West to the North-East, 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

Figure 19: POPULATIONS BY AGE AND SEX IN 1962 AND 1968 AND PROFILE OF MIGRATIONS FOR THE DEPARTMENTS OF SEINE AND SEINE-ET-OISE

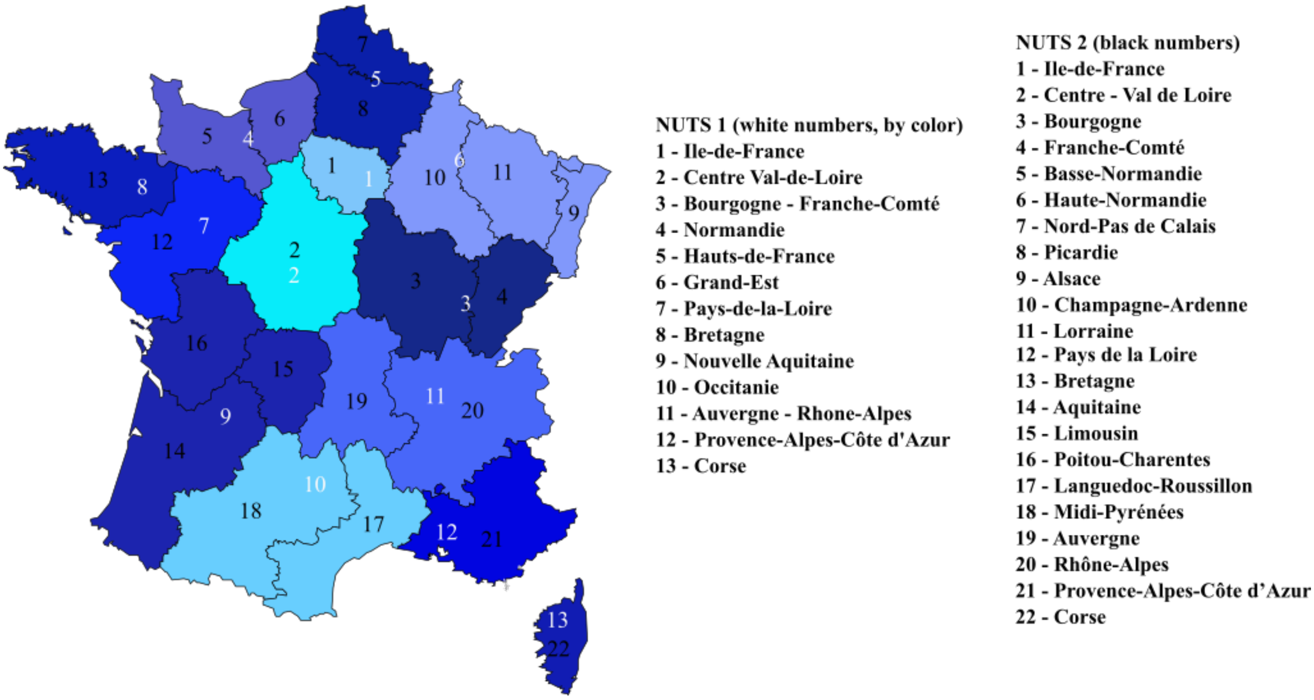


Note: Populations by age and sex in the sum of departments Seine and Seine-et-Oise in 1962 (in black) and in 1968 (in blue) in the upper quadrants. Age profile of migrations by sex in the lower quadrants. Age profile of migrations is the difference between populations by age at the census date and populations by age estimated at the census date.

20% higher than the metropolitan average. In contrast, life expectancies at birth in the South-East, *Seine* and *Bretagne* are significantly lower than the metropolitan average (between 5% and 20% according to the *département*). 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.

Rather than analyzing synthetic indicators such as life expectancy, one can look at age-specific indicators. Since they impacted strongly life expectancies at birth, Figure 23 presents mortality rates between 0 and 5 for women. One more time I have chosen to present the results for women, but these results are available for men too. 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 in 1901 dramatically high. One can see that in extreme case (*Bouches-du-Rhône*), for a thousand children under 5 years, more than 270 (271) died before their fifth birthday. Rates were generally higher in the

Figure 20: CLASSIFICATION OF GEOGRAPHICAL UNITS USED IN THE FRENCH HMD (NUTS 1 & 2)

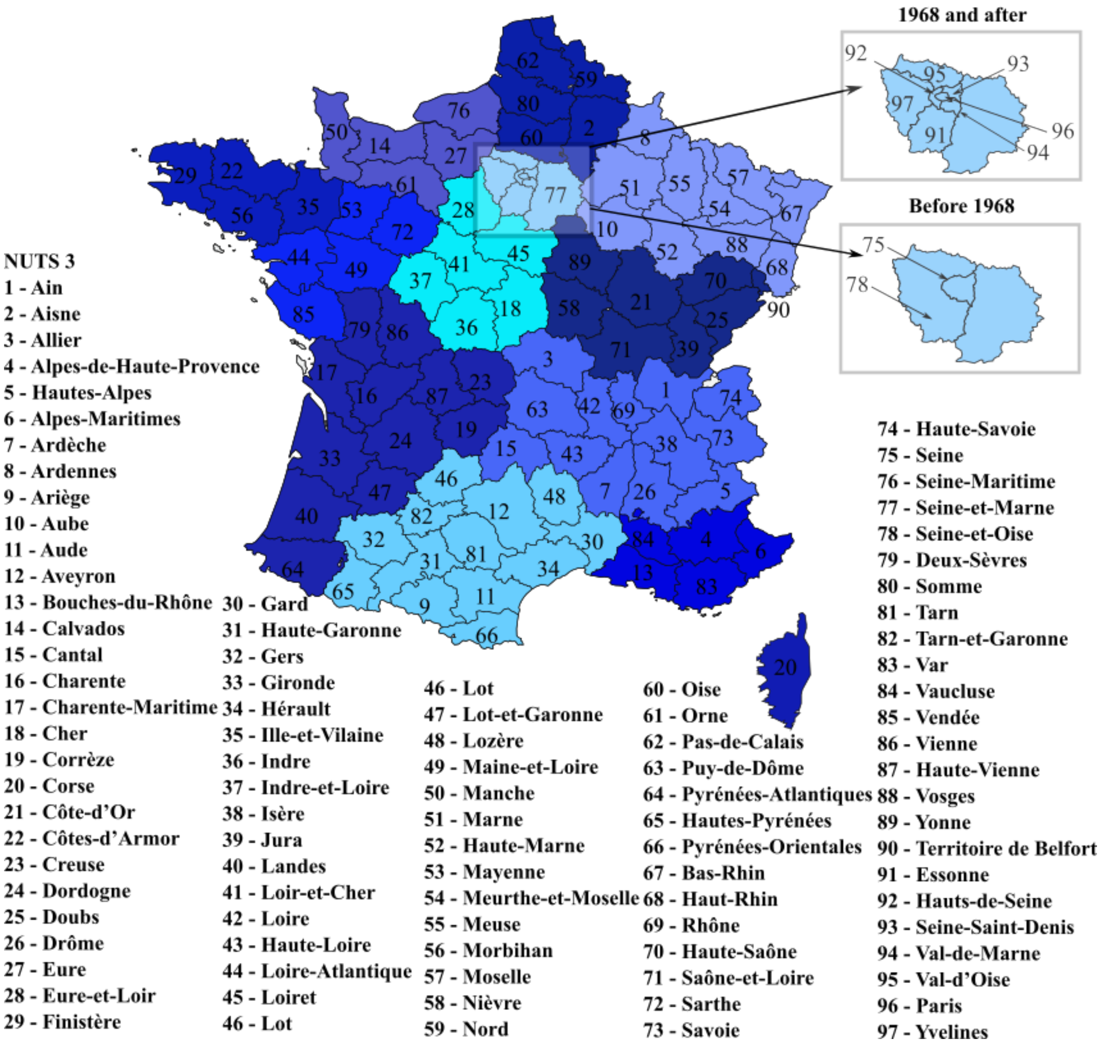


Note: New regions used in the French HMD (NUTS 1) presented by color with associated numbers in white. Regions which existed between 1970 and 2015 (NUTS 2) delimited by the black lines with associated numbers in black.

North and the South-East, while they were lower in a broad central area 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 1926. Child mortality rates decreased between the two years since they were globally around 110 per thousand in 1926.

Finally, one can analyze evolvements of a single department from 1901 onwards. Figure 24 shows women survivors at each age for different dates in *Morbihan*. I have chosen this department since it was a place or high mortality in 1901. Indeed, there was high infant mortality at that time: there were only 850 survivors at age 1 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 ages 60 and 80, then beyond age 80 from 1999 onwards. 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 2018 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 21: CLASSIFICATION OF GEOGRAPHICAL UNITS USED IN THE FRENCH HMD (NUTS 3)



Note: Departments used in the French HMD (NUTS 3). Change of nomenclature used in Ile-de-France presented on the top right.

4.2 Discussion

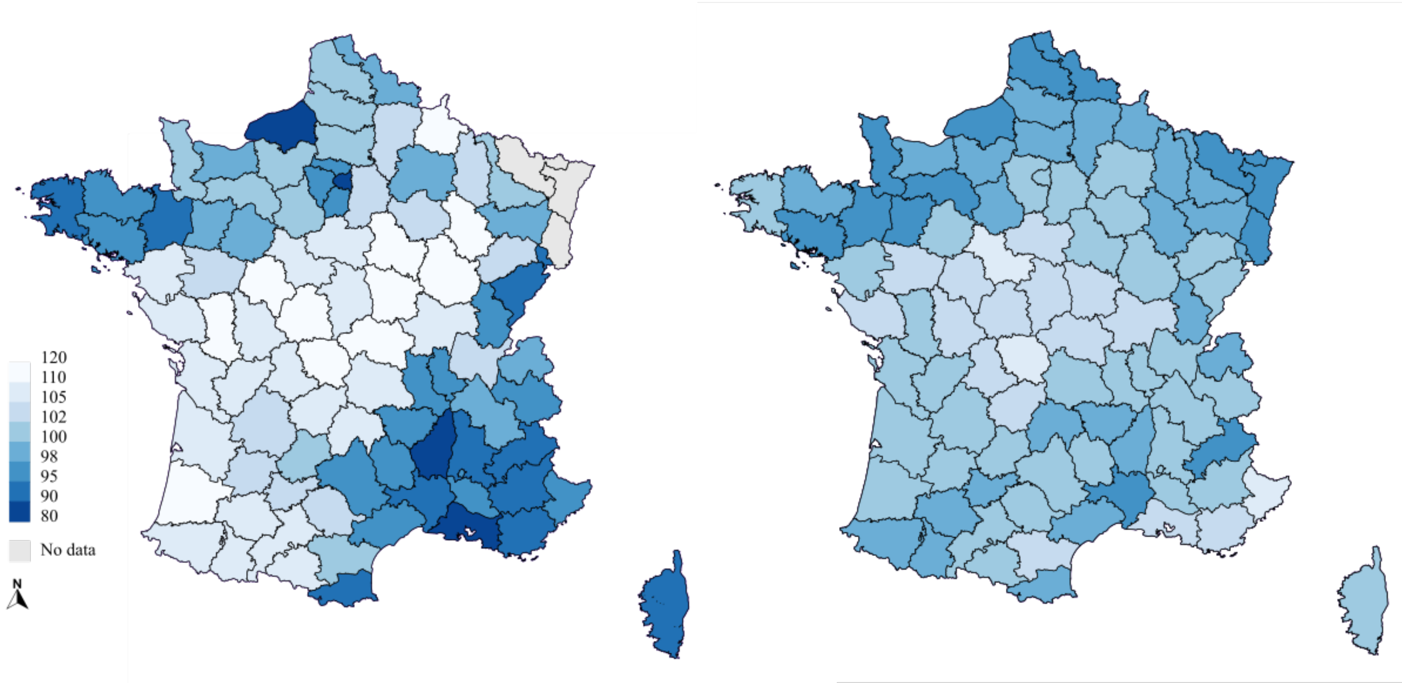
4.2.1 Lifetables Consistency with Other Studies

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

First, I test how departmental and national data are 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 are equal to national values. These expectations were true, which testify to the quality of the raw data. My results are 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

Figure 22: LIFE EXPECTANCY AT BIRTH FOR WOMEN (IN % OF THE METROPOLITAN MEAN), 1901 AND 1946



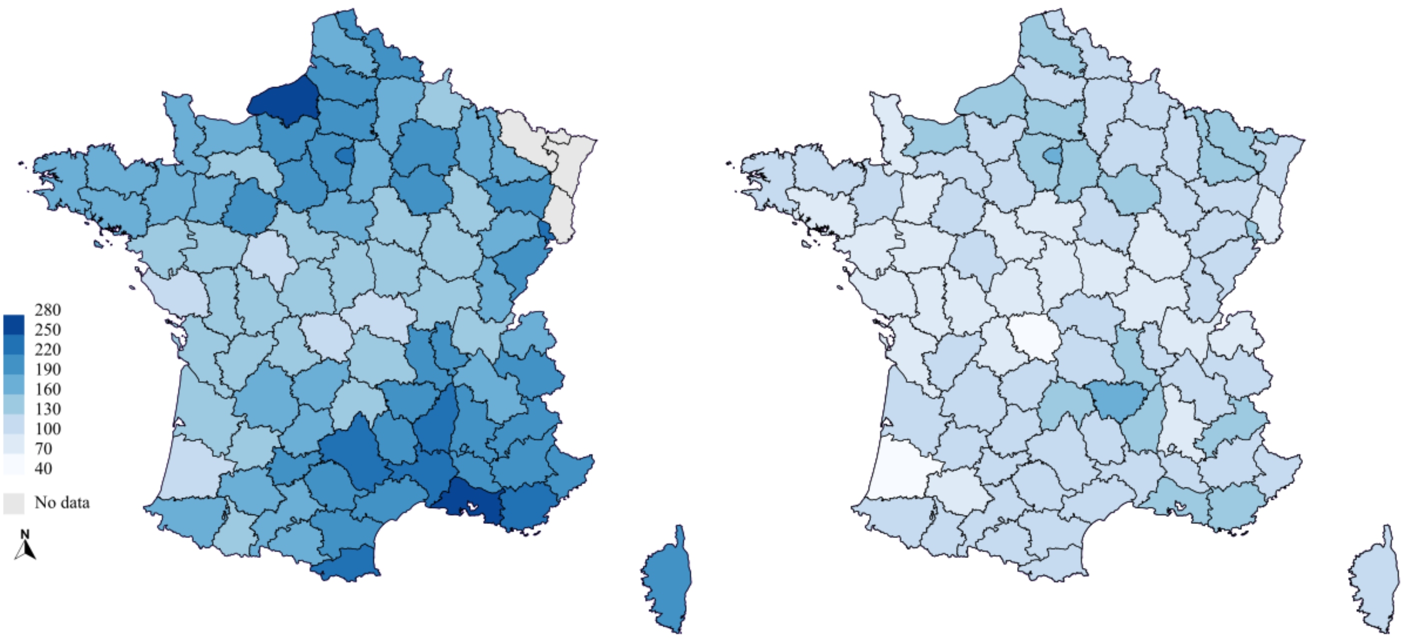
Note: *Moselle, Bas-Rhin and Haut-Rhin* values are non available in 1901 (*départements* under German administration). Sample includes 90 *départements*.

do so, I calculated the differences between my departmental life expectancies and those of Bonneuil (1997) and Daguet (2006). Results are presented in Table 6.

Bonneuil (1997) calculates the life expectancies of women in 1901–1905. I have calculated life expectancies for the same period as well. 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 comprised 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.

Figure 23: MORTALITY RATES BETWEEN 0 AND 5 FOR WOMEN (PER THOUSAND), 1901 AND 1926



Note: Sample includes 90 *départements*. *Moselle, Bas-Rhin and Haut-Rhin* values are non available in 1901 (*départements* under German administration).

Table 6: DIFFERENCES OF DEPARTMENTAL LIFE EXPECTANCIES AT BIRTH WITH OTHER STUDIES

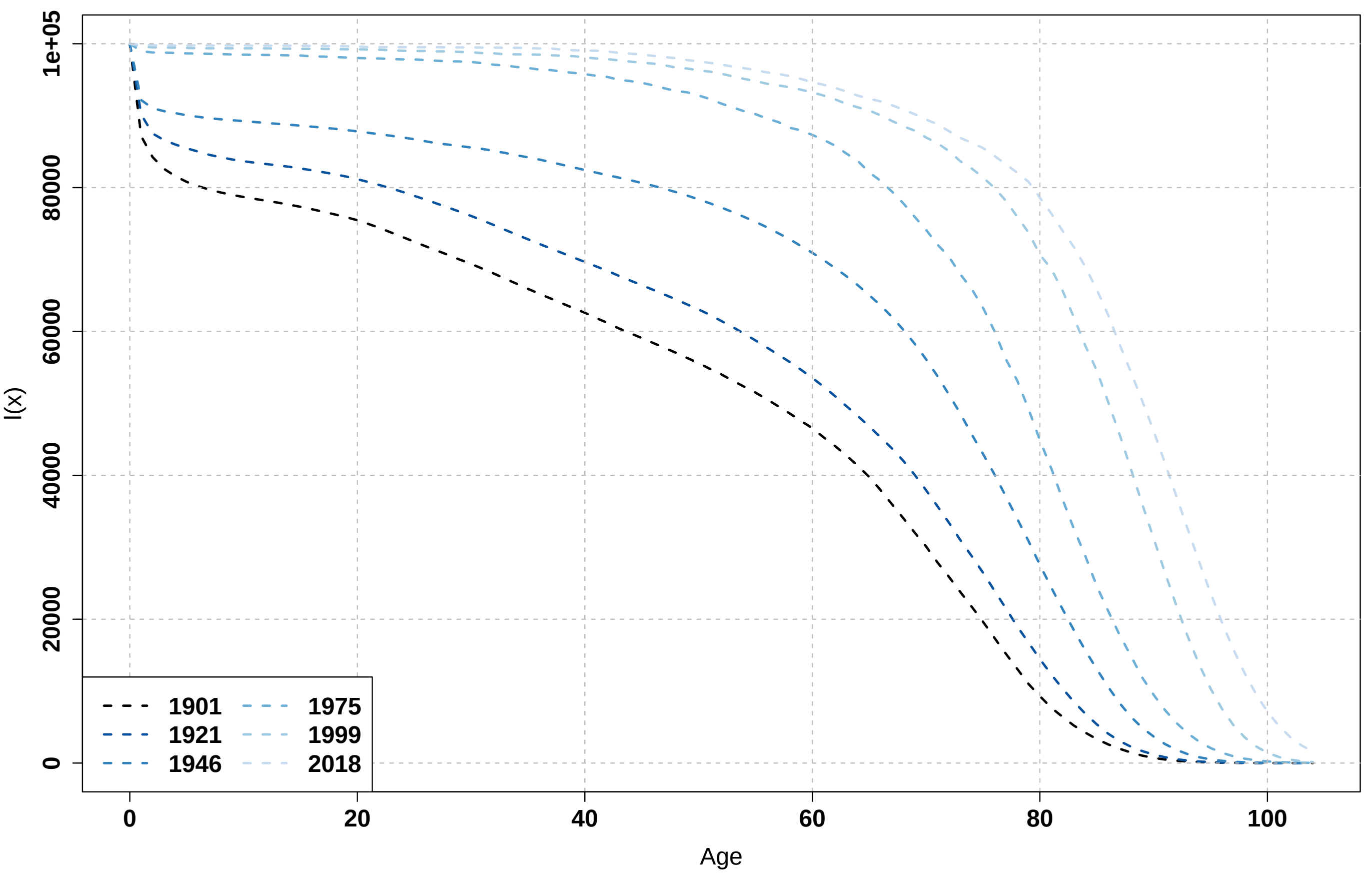
	Men			Women		
	1st Quartile	Median	3rd Quartile	1st Quartile	Median	3rd Quartile
1901–1905				0.49	3.34	6.05
1954	0.18	0.65	1.00	0.54	0.84	1.34
1962	0.00	0.40	0.72	–0.01	0.37	0.68
1968	0.17	0.38	0.73	–0.02	0.33	0.78
1975	–0.17	0.15	0.50	–0.11	0.19	0.47
1982	0.01	0.27	0.59	0.04	0.21	0.50
1990	0.09	0.31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

Note: Differences in % of my computations. Distribution of 90 or 95 departmental differences, according to the classification of the year.

4.2.2 General discussion

Censuses Reliability With population censuses, one know the spatial distribution of the population by age and sex between the French departments along 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: in order to obtain a fair representation in local or national assemblies, electoral divisions are divided so that each of them

Figure 24: EVOLUTION OF SURVIVORS AT EACH AGE IN MORBIHAN



Note: Survival curves for women in *Morbihan.*

represents roughly the same population percentage. Censuses therefore had a very strong political impact. As a result, some regions have sought to inflate their census populations in order to get greater financial or electoral weight. Historians and statisticians have shown that *Marseille*'s population was overestimated in the 1930s.[12] 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.

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 1st 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. At the opposite, this approximation is not verified during war periods. The May-June

[12]See for example *Statistique Annuelle du Mouvement de la Population,* 1939–1942, page 4.

1940 Exodus is an emblematic example. To escape the advance of German troops, the populations of the North-East migrated in mass towards 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.[13] This publication has suggested to estimate the population with ration tickets dispensed to the population. However, Alary et al. (2006) showed that these tickets were circumvented during the war, questioning their reliability in counting the present population. Bonnet (2019) try to estimate these annual populations at department level, 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 a nursery 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.

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 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 miss reliable and available data to overcome this issue.

Small Department Figures Estimating fertility or mortality rates is difficult when figures are small, namely around 0. Papers tackle this issue by using bayesian estimation process (Asunção et al. (2005) and Schmertmann et al. (2014) for fertility rates, Alexander et al. (2017) for mortality rates). The question

[13] See *Statistique Annuelle du Mouvement de la Population,* 1939–1942, pages 3-4, 47 and 55.

arose of using these methods to supplement the HMD Protocol. However, the French departments are not as small as 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.

References

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

Alexander, M., E. Zagheni, and M. Barbieri (2017). A Flexible Bayesian Model for Estimating Subnational Mortality. *Demography 54*(6), 2025–2041.

Assunção, R. M., C. P. Schmertmann, J. E. Potter, and S. M. Cavenaghi (2005). Empirical Bayes Estimation of Demographic Schedules for Small Areas. *Demography 42*(3), 537–558.

Barbieri, M. (2013). La mortalité départementale en France. *Population 68*(3), 433–479.

Bonneuil, N. (1997). *Transformation of the French Demographic Landscape 1806-1906.* Oxford England Clarendon Press.

Cheung, S. L. K., J.-M. Robine, E. J.-C. Tu, and G. Caselli (2005). Three Dimensions of the Survival Curve: Horizontalization, Verticalization, and Longevity Extension. *Demography 42*(2), 243–258.

Coale, A. J., P. Demeny, and B. Vaughan (2013). *Regional Model Life Tables and Stable Populations: Studies in Population*. Elsevier.

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

Dupâquier, J. and J.-P. Bardet (1988). *Histoire de la population française*, Volume 4. Presses universitaires de France.

Fariñas, D. R. and M. Oris (2016). *New Approaches to Death in Cities during the Health Transition*. Springer.

Fries, J. F. (2002). Aging, Natural Death, and the Compression of Morbidity. *Bulletin of the World Health Organization 80*(3), 245–250.

Gilles, H., J.-P. Guironnet, and A. Parent (2014). Géographie économique des morts de 14-18 en france. *Revue économique 65*(3), 519–532.

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

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

Preston, S., P. Heuveline, and M. Guillot (2000). *Demography: Measuring and Modeling Population Processes*. 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.

Schmertmann, C., E. Zagheni, J. R. Goldstein, and M. Myrskylä (2014). Bayesian Forecasting of Cohort Fertility. *Journal of the American Statistical Association 109*(506), 500–513.

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

Vallin, J. and F. Meslé (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*. University Press.

van Raalte, A. A., S. Klüsener, A. Oksuzyan, and P. Grigoriev (2020). Declining regional disparities in mortality in the context of persisting large inequalities in economic conditions: the case of germany. *International journal of epidemiology 49*(2), 486–496.

Wilmoth, J. R., K. Andreev, D. Jdanov, D. A. Glei, C. Boe, M. Bubenheim, D. Philipov, V. Shkolnikov, and P. Vachon (2007). Methods Protocol for the Human Mortality Database. *University of California, Berkeley, and Max Planck Institute for Demographic Research, Rostock. URL: http://mortality. org [version 31/05/2007] 9*, 10–11.

Wilmoth, J. R. and S. Horiuchi (1999). Rectangularization Revisited: Variability of Age at Death Within Human Populations. *Demography 36*(4), 475–495.

Sources of Raw Data

Table 7: SOURCES FOR CIVILIAN DEATHS, 1901–1919

Year	Departments		Classes		Publication	Book	Page
	Total	Missing	Step	Ceiling age			
1901	87	(2)	5	100	SAMP (Year 1901)	31	62–73
1902	87	(2)	5	100	SAMP (Year 1902)	32	62–73
1903	87	(2)	5	100	SAMP (Year 1903)	33	72–83
1904	87	(2)	5	100	SAMP (Year 1904)	34	62–73
1905	87	(2)	5	100	SAMP (Years 1905 et 1906)	35-36	62–73
1906	87	(2)	5	100	SAMP (Years 1905 et 1906)	35-36	140–144
1907	87	(2)	5	100	SAMP (Years 1907-1908-1909-1910)	1	190–193
1908	87	(2)	5	100	SAMP (Years 1907-1908-1909-1910)	1	194–197
1909	87	(2)	5	100	SAMP (Years 1907-1908-1909-1910)	1	198–201
1910	87	(2)	5	100	SAMP (Years 1907-1908-1909-1910)	1	202–205
1911	87	(2)	5	100	SAMP (Years 1911-1912-1913)	2	152–155
1912	87	(2)	5	100	SAMP (Years 1911-1912-1913)	2	156–159
1913	87	(2)	5	100	SAMP (Years 1911-1912-1913)	2	160–163
1914	77	(1)	5	100	SAMP (Years 1914–1919)	3	100–103
1915	77	(1)	5	100	SAMP (Years 1914–1919)	3	104–107
1916	77	(1)	5	100	SAMP (Years 1914–1919)	3	108–111
1917	77	(1)	5	100	SAMP (Years 1914–1919)	3	112–115
1918	77	(1)	5	100	SAMP (Years 1914–1919)	3	116–119
1919	77	(1)	5	100	SAMP (Years 1914–1919)	3	120–123

Note: "SAMP" refers to *Statistique Annuelle du Mouvement de la Population.*

(1) *Aisne - Ardennes - Marne - Meurthe et Moselle - Meuse - Moselle - Nord-Oise - Pas de Calais - Bas Rhin - Haut Rhin - Somme - Vosges.*
(2) *Moselle - Bas Rhin - Haut Rhin.*

Table 8: SOURCES FOR CIVILIAN DEATHS, 1920–1945

Year	Departments		Classes		Publication	Book	Page
	Total	Missing	Step	Ceiling age			
1920	90		5	100	SAMP (Years 1920–1924)	4	82–85
1921	90		5	100	SAMP (Years 1920–1924)	4	86–89
1922	90		5	100	SAMP (Years 1920–1924)	4	90–93
1923	90		5	100	SAMP (Years 1920–1924)	4	94–97
1924	90		5	100	SAMP (Years 1920–1924)	4	98–101
1925	85		5	100	SAMP (Year 1925) - CD	5	2–183
1926	90		5	100	SAMP (Year 1926) - CD	6	2–183
1927	90		5	100	SAMP (Year 1927) - CD	7	2–183
1928	90		5	100	SAMP (Year 1928) - CD	8	2–183
1929	90		5	100	SAMP (Year 1929) - CD	9	2–183
1930	90		5	80	SAMP (Year 1930) - CD	10	16–195
1931	90		5	80	SAMP (Year 1931) - CD	11	16–195
1932	90		5	80	SAMP (Year 1932) - CD	12	16–195
1933	90		5	80	SAMP (Year 1933) - CD	13	16–195
1934	90		5	80	SAMP (Year 1934) - CD	14	16–195
1935	90		5	80	SAMP (Year 1935) - CD	15	16–195
1936	90		5	80	SAMP (Year 1936) - CD	16	16–195
1937	90		5	100	SAMP (Year 1937)	17	54–57
1938	90		5	100	SAMP (Year 1938)	18	154–157
1939	87	(2)	5	100	SAMP (Years 1939–1942)	19	118–125
1940	87	(2)	5	100	SAMP (Years 1939–1942)	19	178–185
1941	87	(2)	5	100	SAMP (Years 1939–1942)	19	238–245
1942	87	(2)	5	100	SAMP (Years 1939–1942)	19	298–245
1943	86	*Corse* (2)	5	100	SAMP (Year 1943)	20	58–65
1944	86	*Corse* (2)	5	100	SAMP (Year 1944)	21	58–65
1945	87	(2)	5	100	SAMP (Year 1945)	22	60–67

Note: “CD” refers to *Causes de Décès*.

(2) *Moselle - Bas Rhin - Haut Rhin*

Table 9: SOURCES FOR CIVILIAN DEATHS, 1946–2020

Year	Departments		Classes		Publication	Book	Page
	Total	Missing	Step	Ceiling			
1946	90		5	100	SAMP (Years 1946–1947)	23	110–117
1947	90		5	100	SAMP (Years 1946–1947)	23	170–177
1948	90		5	100	SAMP (Years 1948–1949)	24	242–249
1949	90		5	100	SAMP (Years 1948–1949)	24	308–315
1950	90		5	100	SAMP (Years 1950–1951)	25	240–247
1951	90		5	100	SAMP (Years 1950–1951)	25	314–321
1952	90		5	85	SAMP (Year 1952)	26	196–203
1953	90		5	90	SAMP (Years 1953–1955)		291–294
1954	90		5	90	SAMP (Years 1953–1955)		360–363
1955	90		5	90	SAMP (Years 1953–1955)		434–437
1956	90		5	90	SAMP (Years 1956–1959)	II	104–115
1957	90		5	90	SAMP (Years 1956–1959)	II	272–283
1958	90		5	90	SAMP (Years 1956–1959)	II	438–449
1959	90		5	90	SAMP (Years 1956–1959)	II	608–619
1960	90		5	90	SAMP (Years 1960–1962)	II	134–145
1961	90		5	90	SAMP (Years 1960–1962)	II	364–375
1962	90		5	90	SAMP (Years 1960–1962)	II	594–605
1963	90		5	90	SAMP (Years 1963–1964)	II	140–145
1964	90		5	90	SAMP (Years 1963–1964)	II	312–317
1965	90		5	90	SAMP (Years 1965–1966)	II	156–165
1966	90		5	90	SAMP (Years 1965–1966)	II	360–369
1967	90		10	75	SCD (Years 1966–1967)		210–211
1968–2019	95		1	125	Detailled Files, INSEE (*)		
2020	95		1	125	Detailled Files, www.insee.fr (**)		

Note: "SAMP" refers to *Statistique Annuelle du Mouvement de la Population*; "SCD" refers to *Statistiques des Causes de Décès*.
INSEE (*): Detailled Files obtained with ADISP.
www.insee.fr (**) : https://www.insee.fr/fr/statistiques/4487988.

Table 10: SOURCES FOR BIRTHS, 1901–1935

Year	Departments		Publication	Book	Page
	Total	Missing			
1901	87	(2)	SAMP (Year 1901)	31	32
1902	87	(2)	SAMP (Year 1902)	32	31
1903	87	(2)	SAMP (Year 1903)	33	32
1904	87	(2)	SAMP (Year 1904)	34	32
1905	87	(2)	SAMP (Years 1905 et 1906)	35-36	32
1906	87	(2)	SAMP (Years 1905 et 1906)	35-36	113
1907	87	(2)	SAMP (Years 1907-1908-1909-1910)	1	128–131
1908	87	(2)	SAMP (Years 1907-1908-1909-1910)	1	132–135
1909	87	(2)	SAMP (Years 1907-1908-1909-1910)	1	136–139
1910	87	(2)	SAMP (Years 1907-1908-1909-1910)	1	140–143
1911	87	(2)	SAMP (Years 1911-1912-1913)	2	104–107
1912	87	(2)	SAMP (Years 1911-1912-1913)	2	108–111
1913	87	(2)	SAMP (Years 1911-1912-1913)	2	112–115
1914	77	(1)	SAMP (Years 1914–1919)	3	44–47
1915	77	(1)	SAMP (Years 1914–1919)	3	48–51
1916	77	(1)	SAMP (Years 1914–1919)	3	52–55
1917	77	(1)	SAMP (Years 1914–1919)	3	56–59
1918	77	(1)	SAMP (Years 1914–1919)	3	60–63
1919	90		SAMP (Years 1914–1919)	3	64–67
1920	90		SAMP (Years 1920–1924)	4	34–37
1921	90		SAMP (Years 1920–1924)	4	38–41
1922	90		SAMP (Years 1920–1924)	4	42–45
1923	90		SAMP (Years 1920–1924)	4	46–49
1924	90		SAMP (Years 1920–1924)	4	50–53
1925	90		SAMP (Year 1925) - CD	5	12–15
1926	90		SAMP (Year 1926) - CD	6	12–15
1927	90		SAMP (Year 1927) - CD	7	14–17
1928	90		SAMP (Year 1928) - CD	8	14–17
1929	90		SAMP (Year 1929) - CD	9	16–19
1930	90		SAMP (Year 1930) - CD	10	16–19
1931	90		SAMP (Year 1931) - CD	11	16–19
1932	90		SAMP (Year 1932) - CD	12	14–17
1933	90		SAMP (Year 1933) - CD	13	14–17
1934	90		SAMP (Year 1934) - CD	14	14–17
1935	90		SAMP (Year 1935) - CD	15	14–17

Note: "SAMP" refers to *Statistique Annuelle du Mouvement de la Population*; "CD" refers to *Causes de Décès*.

(1) *Aisne - Ardennes - Marne - Meurthe et Moselle - Meuse - Moselle - Nord-Oise - Pas de Calais - Bas Rhin - Haut Rhin - Somme - Vosges*
(2) *Moselle - Bas Rhin - Haut Rhin*

Table 11: SOURCES FOR BIRTHS, 1936–1971

Year	Departments		Publication	Book	Page
	Total	Missing			
1936	90		SAMP (Year 1936) - CD	16	14–17
1937	90		SAMP (Year 1937)	17	14–17
1938	90		SAMP (Year 1938)	18	114–117
1939	87	(2)	SAMP (Years 1939–1942)	19	78–81
1940	87	(2)	SAMP (Years 1939–1942)	19	138–141
1941	87	(2)	SAMP (Years 1939–1942)	19	200–203
1942	87	(2)	SAMP (Years 1939–1942)	19	260–263
1943	86	*Corse* (2)	SAMP (Year 1943)	20	18–21
1944	86	*Corse* (2)	SAMP (Year 1944)	21	18–21
1945	87	(2)	SAMP (Year 1945)	22	20–23
1946	90		SAMP (Years 1946–1947)	23	74–77
1947	90		SAMP (Years 1946–1947)	23	132–135
1948	90		SAMP (Years 1948–1949)	24	198–201
1949	90		SAMP (Years 1948–1949)	24	266–269
1950	90		SAMP (Years 1950–1951)	25	196–199
1951	90		SAMP (Years 1950–1951)	25	268–271
1952	90		SAMP (Year 1952)	26	152–155
1953	90		SAMP (Years 1953–1955)		274
1954	90		SAMP (Years 1953–1955)		334
1955	90		SAMP (Years 1953–1955)		408
1956	90		SAMP (Years 1956–1959)	II	53–54
1957	90		SAMP (Years 1956–1959)	II	203–204
1958	90		SAMP (Years 1956–1959)	II	371–372
1959	90		SAMP (Years 1956–1959)	II	541–542
1960	90		SAMP (Years 1960–1962)	II	56–57
1961	90		SAMP (Years 1960–1962)	II	252–253
1962	90		SAMP (Years 1960–1962)	II	494–495
1963	90		SAMP (Years 1963–1964)	II	70–72
1964	90		SAMP (Years 1963–1964)	II	240–243
1965	90		SAMP (Years 1965–1966)	II	69–71
1966	90		SAMP (Years 1965–1966)	II	267–269
1967	90		AS 1968 Table XVIII (*)		50
1968	95		SAMP (Year 1968)		136–137; 144–145
1969	95		SAMP (Year 1969)		136–137; 144–145
1970	95		SAMP (Year 1970)		138–139; 146–147
1971	95		SAMP (Year 1971)		140–141; 146–147

Note: “SAMP” refers to *Statistique Annuelle du Mouvement de la Population*; “CD” refers to *Causes de Décès*; “AS” refers to *Annuaire Statistique.*

(2) *Moselle - Bas Rhin - Haut Rhin*

(*) Since SAMP in 1967 does not exist, I collect the births for the two sexes and distribute them between boys and girls pro rata births in 1966.

Table 12: SOURCES FOR BIRTHS, 1972–2020

Year	Departments		Missing	Book	Page
	Total	Missing			
1972	95		SAMP (Year 1972)		138–139; 148–149
1973	95		SAMP (Year 1973)		138–139; 144–145
1974	95		SAMP (Year 1974)		136–137; 144–145
1975	95		SAMP (Year 1975)		148–151
1976	95		SAMP (Year 1976)		148–151
1977	95		SAMP(Year 1977)		148–151
1978	95		SCD (1978)	II	29–32
1979	95		SAMP (Year 1979)		146–149
1980	95		SAMP (Year 1980)		146–149
1981	95		SAMP (Year 1981)		146–149
1982	95		*Collec. de l'INSEE Série D - La Sit. Dem.* (Year 1982)		171–174
1983	95		*Collec. de l'INSEE Série D - La Sit. Dem.* (Year 1983)		171–174
1984	95		*Collec. de l'INSEE Série D - La Sit. Dem.* (Year 1984)		181–184
1985	95		*Collec. de l'INSEE Série D - La Sit. Dem.* (Year 1985)		172–175
1986	95		*Collec. de l'INSEE Série D - La Sit. Dem.* (Year 1986)		172–175
1987	95		*Collec. de l'INSEE Série D - La Sit. Dem.* (Year 1987)		150–153
1988	95		IR-DS n° 3–4		176–179
1989	95		IR-DS n° 10		174–177
1990	95		IR-DS n° 16–17		212–215
1991	95		IR-DS n° 26–27		186–189
1992	95		IR-DS n° 42–43		188–191
1993	95		IR-DS n° 49–50		188–191
1994	95		IR-DS n° 51–52		188–191
1995	95		IR-DS n° 65–66		188–191
1996	95		IR-DS n° 70–71		217–220
1997	95		IR-DS n° 75–76		194–197
1998–2013	95		www.insee.fr (*)		
2014–2020	95		www.insee.fr (**)		

Note: "SAMP" refers to *Statistique Annuelle du Mouvement de la Population*; "CD" refers to *Causes de Décès*; "AS" refers to *Annuaire Statistique*; "IR-DS" refers to *Insee Résultats-Demographie et Société.*

www.insee.fr (*) : https://www.insee.fr/fr/statistiques/2408051?sommaire=2117120.
www.insee.fr (**) :
2014: https://www.insee.fr/fr/statistiques/1406578.
2015: https://www.insee.fr/fr/statistiques/2106619.
2016: https://www.insee.fr/fr/statistiques/2898646.
2017: https://www.insee.fr/fr/statistiques/3576483.
2018: https://www.insee.fr/fr/statistiques/4190525.
2019: https://www.insee.fr/fr/statistiques/4647557.
2020: https://www.insee.fr/fr/statistiques/5414767.

Table 13: SOURCES FOR CENSUSES, 1901–2020

Date	Departments		Publication	Book	Table	Variable	Ceiling age
	Total	Missing					
March 4th, 1901	87	(1)	RGP Stat.	I to III	I et III	Y. of birth	95
March 6th, 1906	87	(1)	RGP Stat.	II and III	II	Y. of birth	80
March 5th, 1911	87	(1)	RGP Stat.	II	III	Y. of birth	105
March 6th, 1921	90		RGP Stat.	II and III	I	Y. of birth	80
March 7th, 1926	90		RGP Stat.	II and III	I	Y. of birth	80
March 8th, 1931	90		RGP Stat.	II and III	I	Y. of birth	80
March 8th, 1936	90		RGP Stat.	II and III	I	Y. of birth	80
March 10th, 1946	90		RGP - Results by dept	I to VI	I	Y. of birth	80
May 8th, 1954	90		RGP - Results by dept	I to VI	D1	Y. of birth	89
March 7th, 1962	94	(2)	DE - Results by dept	I to VI	D1	Y. of birth	84
March 1st, 1968	95		www.insee.fr (*)			Age	120
Feb. 20th, 1975	95		www.insee.fr (*)			Age	120
March 4th, 1982	95		www.insee.fr (*)			Age	120
March 5th, 1990	95		www.insee.fr (*)			Age	120
March 8th, 1999	95		www.insee.fr (*)			Age	120
January 1st, 2008	95		www.insee.fr (*)			Age	120
January 1st, 2013	95		www.insee.fr (*)			Age	120
January 1st, 2014	95		www.insee.fr (*)			Âge	120
January 1st, 2015	95		www.insee.fr (*)			Âge	120
2016-2021	95		www.insee.fr (**)			Âge	95

Note: "RGP" refers to *Recensement Général de la Population*; "DE" refers to *Dépoullement Exhaustif.* (*) refers to estimated population by quinquennial age groups.

(1) *Moselle - Bas Rhin - Haut Rhin*

(2) In 1962, the census made in *Corse* was irrelevant (cf p. 5 of the book). Only the 1/20th exploitation available in the regional *Provence-Alpes-Côte d'Azur* book was used. This one provided population by quinquennial group of birth years while the last class provided the 74 year-old and over, not the 84 year-old and over. To compute these age classes and get an homogeneous census, I use the distribution of the other *départements*. As an exemple, for ladies born between 1958 and 1962, 23.95% were born in 1961 elsewhere. So I apply this percentage on the sum of ladies born between 1958 and 1962 in *Corse* (4,860) and I deduct that 1,164 were born in 1961.

www.insee.fr (*): https://www.insee.fr/fr/statistiques/2414232.
www.insee.fr (**) : https://www.insee.fr/fr/statistiques/1893198.