

The French Rolling Census: a decade of experience

Jean-Michel Durr¹, François Clanché²,

¹JMStat, Consulting and Analysis. jmdurr@jmstat.com,

²Institut National de la Statistique et des Etudes Economiques (Insee).
francois.clanche@insee.fr

Abstract

Since 2004, France has replaced the traditional form of the census by a rolling census, with the goal of producing annually updated information at local levels as well as to spread the cost of the census over the years. Almost ten years later, a review of new census can be made. The new French census reached most of its goals and publishes good quality results every year. However, some difficulties inherent to annual data collection are faced: costs are directly proportional to the size of the population and are rising accordingly, and data processing workload is not decreasing after years. Several improvements are planned, such as collecting data via the Internet and building a database of dwellings. The paper describes the methodology of the French rolling census, the information produced, and the assessment of the operation after one decade, in addition to the new developments planned.

Keywords: rolling census, alternative census methodology, French census

1. Introduction

Since 1801, France has been conducting regularly censuses for complete enumeration of the population. Censuses were carried out every five years until the Second World War. From 1946, the interval between censuses increased, up to nine years between 1990 and 1999, mainly due to the high cost of the census. Factors associated with political and administrative decentralization in progress in the last 20 years, which led to increased responsibilities of municipalities, as well as the changing profile of demand, more focused on detailed and timely statistical information in support to local public policies, contributed to emerge the need to recast the model of censuses conducted so far.

To make possible to smooth the cost of the census operation over years and to produce more timely information, a model based on census periodic non-overlapping samples was adopted from 2004 in replacing the model of complete enumeration of the population (Durr, 2005). The project started in 1995, with first thoughts based on Leslie Kish ideas. In 2001, a project team was set up to implement the project. Extensive consultations of users were conducted prior to the vote in 2002 of the law enacting the new census methodology.

Since 2008, the fifth year of data collection in the new system, population estimates for all geographic levels, including the most disaggregated such as statistical areas with around 2,000 inhabitants in cities, are obtained annually from the accumulation of five latest annual samples. With the redesign of the census, the estimated cost of the total operation held over five years, is just 70% of the estimated cost to complete enumeration of the population conducted in a single year, thus comparable with the cost of a traditional census would be conducted every seven years.

2. Methodology

An important feature of the territorial organization of France is the existence of a very large number of municipalities, 36,680, most of them very small, but for which an official population figure is needed for the application of many legal texts. Only 980 municipalities have more than 10,000 inhabitants, while 27,000 have less than 1,000 inhabitants. The 36,680 French municipalities were divided into two groups: municipalities of fewer than 10,000 inhabitants and 10 000 inhabitants or more. Each group includes about half of the population, which is about 65 million people. Municipalities with fewer than 10,000 inhabitants, in each region of France, were divided into five groups of rotation, balanced according to the statistical characteristics, based on the 1999 census. Each year, all the municipalities of the rotation group of the year are fully enumerated, as in a traditional census. After five years all municipalities of less than 10,000 people have been enumerated. Groups are enumerated in sequence, and every year the group enumerated five years ago is enumerated again, and so on.

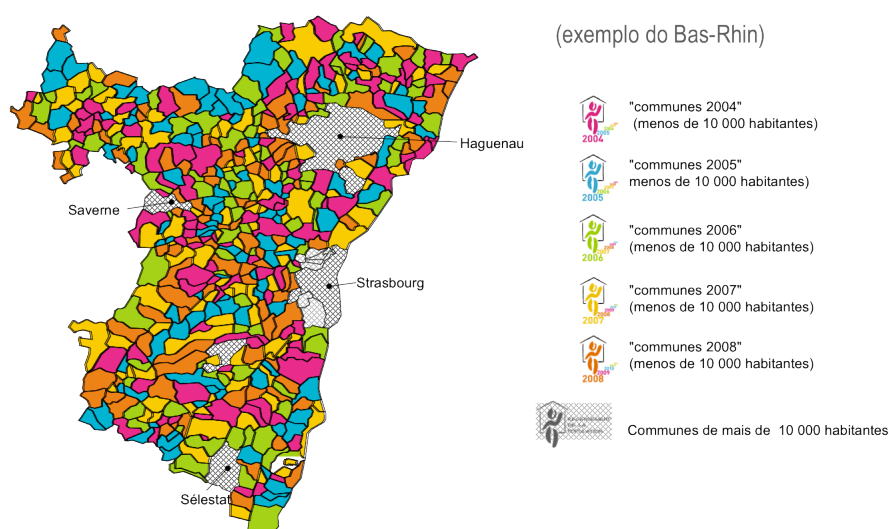


Figure 1 - Five groups of municipalities with less than 10,000 inh.

All municipalities with 10,000 or more inhabitants are visited annually, but only households sampled are listed and characterized. Five rotation groups are formed from a register of residential addresses "Répertoire d'immeubles localisés - RIL" balanced according to the same criteria used for the municipalities of less than 10,000 inhabitants. Each year a sample of addresses corresponding to 8% of households is enumerated. Groups are also considered in sequence and every year, a new sample is taken from the group visited five years ago, taking into consideration new or demolished buildings.

After five years, 70% of households in the country are enumerated, about 50% in municipalities with less than 10,000 inhabitants, which have been made in the municipal complete enumeration of the population, and 20% in larger cities, where you have been investigated a sample of households.

Each year, data collected during the last five years is used to produce updated population estimate of each municipality of France and detailed characteristics of population and housing for all geographic levels. The method for producing estimations varies with the size of the municipality (Durr, Grosbras, 2003), and use additional information as trends observed in administrative sources, such as the housing tax file, that keeps record of all residential dwellings, occupied or not.

3 The surveys are being conducted in highly satisfactory conditions

The tenth census survey was completed in February 2013. Like its predecessors, it went well. Over the years, the municipal personnel acquire an experience that contributes to the operation's efficiency. This professionalization is particularly intensive in large cities—which conduct annual surveys—but is also significant in smaller towns.

At INSEE as well, we have drawn lessons from the initial collections. We have gradually refined the protocols and instructions, made marginal changes in printed collection and management forms, and taken other steps to increase data-collection efficiency and speed. For example, since 2010, the survey in institutions (retirement homes, residential schools,..) takes place at the same time that "classical" household to limit double-counting. From this standpoint, we have clearly reached our goal of improving survey quality control (see Cézard and Lefebvre, 2009).

From year to year, we do not observe any deterioration in collection quality. The non-response rate remains very low and is not increasing, and the number of collections that INSEE needs to "adjust" via additional surveys remains minimal.

4 The results are published on time.

Since December 2008, Insee disseminates each year population official figures of France's 36,680 municipalities. The first year, more than 1,200 mayors actually asked for explanations in the weeks that followed. After receiving the information, very few municipalities expressed dissatisfaction. The number of formal complaints (in the legal sense) was insignificant and did not concern key aspects of the method.

We already repeated the above operation four times. Technically speaking, this entailed no additional difficulty for INSEE, as the method for determining the figures was identical. By contrast, the communication program was complicated by the proximity with the previous year's figures: data freshness—one of the new method's key contributions—creates demanding requirements when annual variations for 36,680 municipalities need to be checked and justified.

The number of requests for clarification addressed to INSEE declining year on year to 290 in 2013. Yet we cannot take the acceptance of the method for granted everywhere. A small number of mid-sized municipalities have challenged the latest figures, sometimes vehemently. INSEE devotes considerable time to checking the data and then explaining the method to these elected officials. While these isolated complaints call for the greatest vigilance, the decision to prepare annual figures of the official population at all geographic levels has clearly proved to be a winning proposition.

Six months after publishing the first population figures based on the new census, INSEE released a broad set of statistical data on its website (Clanché, 2010). This release was designed to satisfy demand from varied segments of the public, from "general public" (user-friendly format and presentation, data directly accessible and retrievable on our website) to specialists and professionals (downloadable databases requiring subsequent handling).

Each summer, we "refresh" all these data and make them consistent with the official populations: Although only some of the data actually reflect updated information (only one-fifth of the information is truly fresh), INSEE has decided to make all the data available again each year.

Feedback from users is largely positive: the census statistics have found their target audiences, both in local communities and among analysts and researchers. Users appreciate the quantity and variety of the information, as well as the richness of the documentation. Like the other arrangements, the dissemination system is not frozen : Developments take place every year, on the basis of user requests, to increase the number of available data, but mostly to facilitate the navigation on the web-site.

Today, we can confidently state that the census data are widely disseminated and used by a varied and generally satisfied public.

5 The quality of the data produced is not challenged

The radical change in the census method created a moderate risk of decline in quality relative to “traditional” censuses, particularly because of the introduction of sample surveys and the longer time frame for the collection.

During the “ramp-up” period, i.e., between 2004 and 2008, we conducted studies to verify the plausibility of the provisional data in demographic terms and the credibility of statistical results, in consultation with INSEE specialists in the relevant fields such as employment, education, and housing. Once we were sure of the quality of the national data, we performed validation tests on local data. Since the initial publication, our validation tests are more abbreviated.

The response from national and local users of these data over the past four years has convinced us that the quality of the results of the new census is at least as high as that of the figures from older censuses.

The quality of the new census holds, first of all, with innovations brought by the new formula compared to a traditional census. The availability of an annual directory of buildings in towns with more than 10,000 people ensures that no address has been omitted in data collection operations. The distribution over five years of data collection in municipalities with fewer than 100,000 inhabitants decreases the burden on the statistical institute, and allows a streamlined monitoring of these communes, in favor of a higher quality.

More importantly, the annualisation of the operations of census allows a true control of the process, enabling the funding of improvements over successive years. Since the introduction of the rolling census in France, no statistical adjustment was necessary to understand the evolution of populations, contrary to past experience.

However, it would be inaccurate to state that the census results are flawless.

Some defects are trivial and can be explained by failure to collect data (isolated cases) or to edit data (for which we can take remedial action). These defects were already present in earlier censuses, but there was no hope of correcting them in later rounds.

Other defects are more “structural” and require deeper methodological scrutiny. The first is the apparent underestimation of young children aged 0-4, which becomes visible when we compare the census numbers with vital statistics or school attendance data. This is a known problem in censuses, both in France and elsewhere, but the new method does not solve it.

6 The data processing workload is not diminishing year after year

The repetitiveness of the construction of census results gave us reason to hope that, once the system was up and running and the first series of detailed results had been published, the workload involved in preparing and validating the data would diminish. Our experience shows that this is hardly the case. A ten-member team is responsible for the specifically “statistical” aspects of the census that range from sampling to determination of legal populations, processing of variables, weightings, validations, and production of microdata files. This workforce is as large as it was in the “project design” period.

Why are we not achieving “productivity gains” here? Basically for three reasons:

a) Although we have now stabilized the method, some parameters of the statistical environment are evolving

The census method is based on the five-year stability of the “municipality” entity whose legal population we seek to determine. But every year a small number of municipalities merge or, on the contrary, are created through separations, or adjust their borders through territorial exchanges. Thus, it may happen that the collection be carried out in a certain geography and statistical results are disseminated in another geography.

While municipalities are born and disappear, they also change size, notably by crossing the 10,000-inhabitant threshold upward (approximately twenty a year) or

downward (two or three a year). Here as well, therefore, we need to define specific calculation procedures for the municipalities' entire "transition" period between the old and new calculation methods that concern them.

It was also necessary to integrate the change of the classification of activities in 2008, and to modify the questionnaire to adapt it to the 2011 EU census regulations. The publication of detailed results based on data collected with different questionnaires is a costly statistical challenge.

b) Over a five-year period, some of the statistical ingredients of the census are proving unstable

The method set up both to estimate populations and to describe them presupposes explicitly or implicitly that the phenomena observed will follow a certain "trend" over the five-year period. But, with respect to these trends, some phenomena are "accidents" inadequately addressed by the basic method.

The first example is the temporary closing, for renovation, of an institution (retirement home, student residence) for the two months of the collection period—an event that potentially "deprives" the municipality of the institution's population for several years.

The second example is the demolition of a large building preceded by a period in which the dwellings are gradually vacated: if the building is surveyed at a time when it is almost empty just before it disappears from the address register, population may be underestimated.

Another situation is the change in legal status of a retirement home (institution) that makes it subject to the occupancy tax (*taxe d'habitation*) in a small municipality: the normal extrapolation may cause an increase in the number of "private" dwellings to show up in the figures, whereas the institution has already been included in the total.

In all those situations, specific adjustments are necessary.

c) The methodological teams are responsible for correcting the accidental defects in the basic census data

Lastly, an inevitable number of quality accidents occur in the census—as with all very large-scale statistical operations: errors in the sampling and editing frames, collection errors (such as forgotten units, and surveys carried out by mistake in large municipalities), deficiencies in the occupancy-tax database (of which INSEE not control the compilation process), and so on. All these errors, when spotted in time and statistically significant, are subject to corrective calculation.

A total of some one hundred and fifty municipalities a year are subjected to an *ad hoc* calculation, and at least fifty are reviewed in an in-depth analysis that does not result in an adjustment. These various adjustments and corrections make up approximately one-third of the "data processing" workload for the census.

7 The cost of the census remains high

As regards census costs, the goal of the redesign was not to reduce the cost of the operation but to smooth it over time. This objective has been achieved, and expenditures related to census surveys are now listed annually in the budget of INSEE. Census is no longer exposed to the same risk as the former censuses: it should be recalled that France's last general census was postponed from 1997 to 1999 for budget reasons.

However, the cost of operation is high: more than 54 million Euros per year, or € 0.83 per capita. In total, it is not lower than that of the traditional census exhaustive, but it allows to have fresh data each year, that did not permit the old system.

Beyond the financial aspects, there are also human-resource issues involved. At INSEE, 450 staffers are assigned full-time to census work. During the collection period, another 460 staffers are responsible for training, accompanying, and supervising municipal personnel conducting census operations. Our survey experience shows only a small decrease in this human-resource requirement. However, like all

government agencies, the Institute is experiencing a steady, significant decline of its workforce.

In a period of pressure on public finances, while the increase in population automatically increases the cost of collection, the issue of medium-term sustainability of funding and carrying out the census is on the table.

8 Introduction of online collection and modernization of address-register management

The 2004 census redesign was undertaken without changing the collection protocol: the census forms are dropped off by the enumerator, filled out by households, retrieved by the enumerator, then scanned for data acquisition.

The first notable change in this protocol will be to introduce online response for households starting with the 2014 survey. This modernization will offer INSEE an opportunity to launch an overall modernization of the collection-monitoring system with the aim of saving resources and controlling quality.

At the same time, INSEE will upgrade the management of its “register of localized buildings” (Répertoire des Immeubles Localisés: RIL) that serves both as a sampling frame for annual surveys and a reference for calculating legal populations in large communes. It is extremely costly to maintain, as it occupies 140 staffers in our central and regional offices. The updating process and the underlying IT infrastructure will be redesigned to ensure better control of operations and their statistical impact and reduce management costs by taking advantage of the technical possibilities and administrative databases, that did not exist in the early 2000s

9 Conclusion: Developments considered for the 2020's Census

The establishment of a rolling census, which was an audacious bet when it was launched 15 years ago, is clearly a success. Ongoing projects will reduce its cost. However, Given the economic situation, INSEE thinks, like many other national statistical institutes to new developments of its mode census.

The institute has already initiated the building of a statistical register of dwellings, compiled from tax data and be enhanced by systematic mapping through the cadastral survey reference, henceforth linked to a geographic information system (Système d'Information Géographique: SIG). This register would allow greater use of administrative sources in the future French census.

No decision has been taken so far, as the administrative role played by the Census, beyond its statistical functions, requires broad consultations before any significant change.

References

Cézard, M., and Lefebvre, O., 2008, Census quality, INSEE (available at www.insee.fr, “INSEE and official statistics - Quality at the Insee - 4- quality commitment ” section)

Clanché, F. , 2010, The dissemination of French census results since 2009, Joint UNECE/Eurostat Group of Experts on Population and Housing Censuses, Working paper 18 (available at <http://www.unece.org/stats/documents/2010.07.census.htm>)

Durr, J.-M, 2005. The French new rolling census. Statistical Journal of the UNECE. Sept. 2005. Geneva.

Durr, J.-M., Grosbras, J.-M., 2003. La rénovation du recensement français: principes et méthodes. Insee, programme de rénovation du recensement, Paris.