United Nations ECE/CES/GE.41/2013/24

Economic and Social Council

Distr.: General
12 September 2013

Original: English

Economic Commission for Europe

Conference of European Statisticians

Group of Experts on Population and Housing Censuses

Fifteenth Meeting
Geneva, 30 September – 3 October 2013
Item 2 of the provisional agenda
Census methodology

The French rolling census, ten years after its launch[1]

Note by the National Institute of Statistics and Economic Studies, France

Summary

France was the first country to organize his population and housing census as a "rolling census": census surveys are spread over five years. The figures are established using extrapolation, interpolation, and weighting methods. Ten years after its establishment, what lessons can be learned from the implementation of this device?

The main objectives of the reform have been achieved: Census collections are held each year since 2004, the level of participation of the population is very high. The accumulation of experience from all relevant stakeholders ensures a better quality of field operations. Since 2008, each year, Insee establishes and disseminates official population figures at all geographic levels, and detailed statistics.

But the French method does not reduce the cost of the census. This comes from the number of surveys carried out, which is still important, but also from the methodological cost of the formula. The construction and updating of the sample frame, the establishment and maintenance of the data processing system, the specific treatments related to the allocation of 5 years of investigations (and frames) to establish the figures, require significant and qualified statistics and IT teams.

The device does not allow establishing annual series of results at all levels of geography: only national or large regional data are yearly reliable. For detailed geographic levels, only the dissemination of the 6th census, in early 2014, will allow comparisons with those established 5 years ago.

[1] Prepared by François Clanché.

I. Introduction

1. From 1801 to 1999, 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.

2. The decision, at the end of the 1990's, to move to a rolling census, was a response to two concerns:

(a) the demand of more local and timely statistical information to face de administrative decentralization in progress, without increasing of the census cost;

(b) the wish to smooth the cost and burden of census survey, to ensure sustainable funding and quality control.

3. The first thought took place inside Insee in 1995. After a first set of methodological works and field tests, strategic decisions on the new kind of census were taken in 1998. From 1999 to 2001, an extensive set of consultations was held. In February 2002, a law was voted enacting the new census methodology. In 2001, a project team was set up to implement the project. The first survey of the new census began in January 2004, and the first full results were released in July 2009, almost 15 years after the first reflexions inside the INSEE.

II. Methodology

4. The methodological details of French rolling census are linked to two characteristics of the country:

(a) The very large number of local administrative units (more than 36 000 municipalities, the *communes*). Half have less than 420 inhabitants, and a small minority (950 municipalities with more than 10,000 inhabitants) is home to more than half the country's population. This specificity excluded the establishment of a uniform statistical model throughout the country.

(b) The lack of a localized population register, used as survey or estimation frame, like for example the system put in place in Switzerland.

5. 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.

6. Municipalities with fewer than 10,000 inhabitants were divided into five groups of rotation, balanced according to the statistical characteristics in each region of France, 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.

7. In the interests of statistical consistency, but also regulatory equity, the official population figures of all the communes must refer to the same year, called "base year". To ensure good statistical robustness, it was decided that this base year is the middle year of the five-year cycle of collection. For example, population figures published at the end of the year 2013, based on surveys conducted between 2009 and 2013, the base year is 2011.

8. For the fifth of small communes surveyed in 2011, the result of the investigation is exactly taken into account. For the municipalities surveyed in 2012 or 2013, we assume a

steady growth between the last published figure (which had 2010 as reference) and the most recent survey (2012 or 2013) to project the year 2011 by linear interpolation. For the communes recorded in 2009 or 2010, we must "refresh" the data to determine a figure in 2011. An extrapolation rate is built, using evolution of the total numbers of dwellings in the occupancy-tax (*taxe d'habitation*) file and the trend of the average number of inhabitants per housing.

9. 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 Local building register ("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.

10. After five years, 40% of households are enumerated in large cities. The population figures are determined by multiplying the average number of inhabitants per housing (determined through the last 5 census surveys) by the total number of housing (included in the RIL of the base year).

11. All detailed statistical census results are established by weighting the survey data, in line with population figures. Thus in small towns, the weights are close to 1 (but deviate when the date of collection is remote from the base date), while in large towns they average 2.5.

III. The surveys are being conducted in highly satisfactory conditions

12. Although census surveys do not relate to the entire population, they remain major operations: more than 4.5 million homes are surveyed, and more than 9 million people answer to census each year. The surveys are conducted by 22,000 enumerators, supervised by 8 000 local officials. INSEE is present on the field with 450 “supervisors” who train, assist, advise and control the collection teams, in addition to the permanent staff in office (250 people in the regional or central team’s directions). As a collection of "traditional" census, we must prepare millions of documents and distribute them throughout the territory, design and deliver training, organize communication between stakeholders and for the general public, provide powerful computing, the whole in a completely constrained schedule. What is specific to the rolling census is that all this happens every year.

13. The tenth census survey was completed in February 2013. Like its predecessors, it went well.

14. Fears about the proper conduct of investigations focused on two points:

(a) The participation of residents in regular surveys on only a portion of the population, while traditional censuses focus their communication on the exceptional and universal character of the operation.

(b) Understanding and acceptance of the new system by the actors of the collection, in particular municipalities, to whom the law has entrusted the responsibility of the survey organization and supervision of enumerators.

15. "Traditional" periodic censuses can build on communication campaigns which presents the survey as a systematic, exceptional and time limited operation. The partial annual census cannot rely on these arguments. Nevertheless, a communication campaign focused on the usefulness of the census and, secondarily, on the innovative and money-

saving aspects of the new system has generated a fairly strong endorsement by the population.

16. Every year, posters are made available to municipalities and, now, banners for websites and posts on social networks. Also press conferences are held in Paris and in each region before the beginning of each survey, to remind the embodiment of investigations and illustrate with concrete examples the usefulness of census results. It has become customary for many journalists who are there. Press reports are numerous and very largely positive. Over time, the communication messages are less about the originality of the method, but are more on the practical uses of the census (it is possible to give the results of previous surveys) and the simplicity of the census operation of the citizen point of view.

17. Each year, we carefully observe the proportion of people, from which it is impossible to get an answer. Although this figure includes persons absent for the duration of the operation or unreachable, it is a good indicator of rejection of the operation among the population. This rate is very low (3% of households on average, with higher levels in large cities, lowest in the country), and it does not increase year after year. While enumerators reported an increasing proportion of people suspicious or aggressive. But probably due to their experience over time, field workers are able to maintain an excellent level of response.

18. The February 2002 Act establishing the new census called for a sharing of census-taking operations between INSEE and French municipalities: INSEE organizes and supervises the surveys, but the municipalities prepare and execute them, notably by hiring, managing, and paying the enumerators. In the context of decentralization and occasional tensions between the State and local government over the distribution of resources and responsibilities, it was by no means certain that this partnership would function properly.

19. In the event, the establishment of annual data-collection procedures went smoothly, thanks to major communication and support actions aimed at the municipalities. Even before the first surveys, INSEE launched a communication program to provide an overall explanation of the new system. Every year, INSEE now assists the municipalities conducting the survey in the quarter preceding the collection, then during the collection itself. INSEE staff train the municipal personnel in charge of the operation, as well as enumerators. 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. Each year, 70% of the municipal census staff had already at least one survey experience. Similarly, 40% of enumerators have already done this work at least once, and this rises to 60% in large towns. All surveys actors say they are more efficient and work better the second year than the first, and this is confirmed by their supervisors. Surveys are periodically organized among field stakeholders at the end of the collect, to know their problems and measure the adequacy of training and tools.

20. At INSEE as well, we have drawn lessons from the initial collections. We have gradually refined the protocols and instructions, made marginal changes in 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.

21. In addition to training and manual, have been implemented successively operations schedules for municipal census officials, "checklists" for INSEE supervisors and short "memory aid" for enumerators, to help them in their daily tasks. Often these documents were developed centrally because we realized that the local actors made their own, independently of each other, given their experiences. These documents are now adapted and updated each year.

22. From this standpoint, we have clearly reached our goal of improving survey quality control. Thanks to regular assessments of the process by all participants, we have achieved progress where needed (see Cézard and Lefebvre, 2009).

IV. The results are published on time

23. The second challenge of the new census was the determination of the official population figures of France's 36,680 municipalities (*communes*) at end-2008, using the 2004-2008 surveys, the register of localized buildings (Répertoire des Immeubles Localisés: RIL), the register of institutions, and the occupancy-tax (*taxe d'habitation*) data. Calculations must be completed (and the figures controlled) at the beginning of December, knowing that the investigation ends on the field at the end of February, and the data computer acquisition ends in August. A complex statistical and computing device has been designed to calculate and validate data as soon as possible, sometimes on provisional data collection (preliminary count). This task was completed on time.

24. Throughout 2008, INSEE conducted a specific communication campaign to inform mayors of the calculation method and announce the future annualization of the figures. In December, each mayor received the figures for his or her municipality a few weeks before the official publication of the list in a decree.[2] INSEE Regional Offices made arrangements to answer mayors' requests for information. 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.

25. 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. INSEE devotes considerable time to checking the data and then explaining the method to these elected officials.

26. 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. Most are municipalities whose 2006 population exceeded that of 1999 and has since been trending down. Even a mild decrease has a symbolic effect on the municipalities concerned, as well as an impact on local finances.

27. INSEE leads a permanent dialogue with representatives of local 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. It is now clear that even if all elected officials are not entirely convinced by the new method, the annualization of official population is considered by the mayors as great progress.

28. Six months after publishing the first official population figures based on the new census, INSEE disseminated a broad set of statistical data on its website[3]. Like the population figures, they were based on the 2004-2008 surveys and used 2006 as base year. The data were fully consistent with official population figures.

[2] See official population figures at *Insee - Populations légales.*

[3] See complete statistical results of census at *Insee - Les résultats du recensement*

29. The overall device was designed after numerous consultations with regular users of the census. Its implementation then mobilized members of the production team of census and dissemination specialists.

30. Obviously, the dissemination plan must meet statistical confidentiality. The nature and amount of data disseminated was discussed with the national authority responsible for ensuring compliance with civil liberties in connection with computerized data (CNIL). This release was designed to satisfy demand from varied segments of the public. We accordingly organized the data in several complementary forms. For the “general public,” the data are directly accessible and retrievable on our website, in a rather user-friendly format and presentation. For specialists and professionals, we provided downloadable databases requiring subsequent handling by the user.

31. In the first segment, we prepared easy-to-print pages of “key figures” on all municipalities—down to the smallest—with a few charts and data comparisons with earlier censuses. Also for the general public, but only for larger geographic units, we prepared detailed tables combining several variables, often comparable with those of the 1999 census. For specialists, we compiled databases containing the figures of the detailed tables and allowing all possible geographic aggregations. A few months later, we added databases at infra-municipal level and databases of individual anonymized data. For the first time, micro data from a census are available free of charge to a wide public and can be freely tabulated by users.

32. An educational effort was made and specific "precautions" were written for users because of the new method.

33. Each summer since 2010, we “refresh” all these data and made them consistent with the last official population figures. 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.

34. Feedbacks from users are 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. But, as always, this publication generated additional demands that was expressed by website visitors, during meetings with policy makers, in an ad hoc group at the National Council for Statistical Information (Conseil National de l’Information Statistique: CNIS), and in a July 2010 online satisfaction survey.

35. Like the other arrangements, the dissemination system is not frozen. We made marginal changes in summer 2010, and more important reorganisations in 2011, particularly to facilitate navigation on the website, and in 2013, to improve the integration of census results in the general dissemination of local data on the website of the INSEE. There are also evolutions each year to increase the number of disseminated data. But INSEE takes care not to complicate the offer too much, to continue to respect the limits imposed by the statistical confidentiality, and to ensure the quality and the precision of the disseminated data.

36. The number of visits of the subset “results of the census” of the website of INSEE varies from 70,000 to 110,000 per month. Thus, the results put online in July 2010 received in one year a million visit, and 11.8 million pages of results were consulted.

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

V. The quality of the data produced is not challenged

38. Publishing detailed census data every year, strictly on time, would not be a success if the quality of the data were not recognized. 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. We had weighed and estimated the risk in the design phases, notably through simulations.

39. 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 in cooperation with our Regional Office network. Lastly, the main dissemination products underwent multiple rereading's in the months before their online release. These progressive validation phases led to the publication of national and regional results as early as 2005. And they enabled us to verify the quality of the final results as thoroughly as possible.

40. Since the initial publication, our validation tests are more abbreviated: they concern (1) the official populations' figures of municipalities, with systematic analyses of major annual trends, and (2) selected statistical results for regions and departments.

41. 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.

42. 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.

43. More importantly, the annualisation of the operations of census allows a true control of the process, enabling the funding of improvements over successive years. These check operations and the continuity of the method from one year to another make it possible to reduce to the strict minimum the data-collection risks specific to the traditional censuses. Since the introduction of the rolling census in France, no statistical adjustment was necessary to understand the evolution of populations, contrary to past experience.

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

(a) 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.

(b) 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 until now method does not solve it, because we did not succeed in understanding its cause.

VI. The cost of the census remains high

45. As regards census costs, the goal of the redesign was not to reduce the cost of the operation but to smooth it over time. The sampling rate in large municipalities was actually calculated so that the survey volume would be equal to one-seventh of that of an exhaustive survey, in reference to the "usual" seven-year frequency of the last general censuses. INSEE consequently spends some €33 million on the census every year, roughly one-seventh of what an exhaustive survey would probably have cost. Of the annual €33 million, INSEE transfers €22 million to municipalities, which collect the data. The rest is spent on the printing of census forms, the capture of questionnaires by optical scanning, the communication campaign, INSEE staff travel, and the compensation of some interviewers.

46. The smoothing 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.

47. However, the total cost of operation is high, including salaries of INSEE dedicated to the census: 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. By considering that it is necessary to conduct 5 surveys to establish complete figures of census, and by taking of account the initial costs of installation of the device, one lead to total costs of the census of 300 million euro, which corresponds to the cost of the last traditional census, that of 1999. The comparison is not entirely fair, because the rolling census establishes figures every year, which was not the case of the traditional method. But in a period of pressure on public finances the issue of medium-term sustainability of funding and carrying out the census is on the table.

48. The issue of sustainability arises both in terms of specific expenditure for the state and in terms of workload for the statistical institute.

49. While expenses are now more regular, they match each year seventh of those of a traditional general census, which remains high: the census alone accounts for over a third of INSEE expenditures. As French population is rising by nearly 0.6% per year, his cost tends to increase each year, because the rules governing the INSEE allocation to municipalities actually stipulate that this funding is strictly proportional to population. At the same time, the operating budget of the INSEE (excluding payroll) decreased by 10% between 2001 and 2013, and census must be included in the common effort to curb spending.

50. Beyond this financial aspect, there are also human-resource issues involved. At INSEE, 400 staffers are assigned full-time to census work: 40 provide the design and the central control of operations (including methodological work and data processing), and 360 work in the regional offices of the INSEE. These share their work between the preparation of surveys (about 40 people year equivalent), tracking surveys (30 persons), reception and control of collected questionnaires (90 people), post-collection treatment (60) and updating the Local building register (140). During the collection period, another 450 staffers are responsible for training, accompanying, and supervising municipal personnel conducting census operations. Our experience shows only a small decrease in this human-resource requirement, which currently mobilizes a tenth of manpower of INSEE. However, like all government agencies, the Institute is experiencing a steady, significant decline of its workforce.

51. At a lesser level, many municipalities are complaining of their difficulties in hiring sufficient enumerators. This situation is reported by one-half of localities that need to hire more than ten enumerators.

52. Again, the issue of mid-term sustainability is on the table.

VII. The data processing workload is not diminishing year after year

53. 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 at the moment, and the situation is unlikely to improve in the years ahead—especially if we want to settle the quality issues that are emerging. The actual methodological work of establishment of the standard calculation method is completed since 2008, and corresponding computer programs as well. But the effective reduction in the properly methodological needs was compensated by an increase in the load of processing data, because of need for permanent adaptations.

54. 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 micro data 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

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

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

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

55. 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. The treatment of these modifications varies according to the position of the event date relative to the census collection cycle. Each “event” of this kind therefore requires several years of ad hoc calculations because the components of the calculation (surveys, sampling frame, extrapolation frame) cannot be mobilized as for other municipalities.

56. 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. This topic is particularly challenging because the entire system, from the collection up to computing the population is initially based on the fact that a municipality is "large" or "small". This is true in the statistical sense, but also in the organization of information systems and computer programs. In total, it took several years of work by experienced statisticians and computer scientists to design and implement all treatments tailored to the situation. As the method is based on a 5 year cycle, it is necessary to make a specific calculation for a transitional period of five years. In 2013, 55 municipalities require specific calculation due to a crossing of the threshold of 10 000 inhabitants. Obviously, controls and validations steps are further for these territories.

57. It was also necessary to integrate the change of the classification of activities in 2008 (new French Classification of Activities in coherence within the European Communities NACE rev. 2).

58. Lastly, the census questionnaire is not set in stone: in 2011, we modified it for the first time since 2004, in a marginal way, to adapt it to the 2011 EU census regulations. The publication of detailed results based on data collected with different questionnaires is obviously a difficult, costly, and sometimes ultimately impossible statistical challenge. Sometimes it was necessary to reduce the level of detail of published data to focus on the possible aggregation of modalities in both classifications. Sometimes, to build the new terms of a variable from the answers in the old nomenclature, we conducted imputations based on a statistical model.

59. A further change in the questionnaire was planned for 2016 or 2017 after a round of discussions with census users in 2011-2012. But considering the cost of the last evolutions conducted, no decision is taken on the extent and the timing of the next evolution.

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

60. 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. In small municipalities, we extrapolate or interpolate by making proportionality hypotheses. In large municipalities, we apply five-year moving averages for the variables to a known number of dwellings in the median year. But, with respect to these trends, some phenomena are “accidents” inadequately addressed by the basic method.

61. 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. In those situations, we must collect efficient information from the staff of the institution, and then implement a specific correction in the calculation process.

62. The second example is the demolition of a large building preceded by a period in which the dwellings are gradually vacated. If the building, almost empty, is investigated just before its destruction, the average number of inhabitants per housing in the town will decline, reflecting the reality. When the next year, the building is destroyed, the total number of dwellings in the municipality decreases, but the impact of this building on the average size of households in the town remains. The population is thus underestimated. So be careful to identify these situations and, again, develop a specific calculation.

63. Another situation worth mentioning 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 during the last survey.

64. Last example, discovered recently: In a small town of 2 000 inhabitants, the average number of persons per housing was 2.7 in the 2005 survey, and only 2.3 in 2010. There is no error in the collection. The "annual household size coefficient" which is an average of 0.993, here is 0.967. Our method extends this trend, leading to estimate a population decline between 2010 and 2011, even as the number of units increases significantly during 2010. This exceptional situation arises from the construction, between the two census surveys, of 210 apartments (in a town that has only 1000!) specifically designed to accommodate students, and therefore particularly small. The extrapolation method, which usually allows very good estimates, here is failing because the recent past, atypical, has no reason to reproduce. Once identified these specific situations, appropriate calculations are made.

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

65. Lastly, an inevitable number of quality accidents occurs in the census—as with all very large-scale statistical operations repeated every year and from which results are expected for more than 36,000 territorial units. Accidents include errors (involuntary omissions and double counts) in the sampling and editing frames (namely, the Local building register and institution register), 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. The municipalities themselves, upon receipt of the intermediate data concerning them, report some errors. However, INSEE takes steps to ensure that municipalities are treated fairly. The number of municipalities for which such calculation is necessary because of errors tends to decrease with time and experience of all stakeholders: 40 communes were recovered in 2009 and 2010, 20 in 2011, 10 in 2012. It is always the buildings inventory (RIL), which justifies more specific calculations, in the case of buildings, which were not present in the directory of the base year, are identified late.

66. 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.

VIII. Annual surveys do not establish annual series

67. The new census produces data every year at all geographic levels. As soon as we introduced the system, we clearly announced that these successive local data would not constitute annual series in the usual statistical sense: only results compiled from independent data sets (i.e., at least five years apart) are strictly comparable.

68. Why is it not legitimate to compare two annual figures for the same municipality?

69. In a small town surveyed in 2013, the figure to be released late 2013 is more recent than the figure for the end of 2012, since it is based on a comprehensive and recent survey and, as such, it presents a new interest. However, the gap between these two figures that have the base years 2010 and 2011 is not necessarily representative of the demographics of the town from 2010 to 2011. It depends, in fact, on the one hand, the quality of data from the occupancy-tax used to describe the actual evolution of the population between 2008 and 2010, and secondly, the methodological principle that distributed uniformly the evolution from the 3 years 2010 and 2013.

70. In the case of a large community, population calculations take into account the number of dwellings in the base year (2010 and 2011 respectively) and the average number of persons per measured during five household surveys. The first part of the calculation, the number of housing roughly corresponds to the evolution housing stock of the town between 2010 and 2011 (the 1st January situations are average situation in the 1st July framing: strictly speaking it 'is an average growth over two years). However, the second parameter is the difference between the number of persons per household in the 2008 survey and in the 2013 survey, which will explain its evolution in the calculation of 2010 (based on the period from 2008 to 20120) and 2011 (2009-2013). This difference corresponds to a demographic trend of the municipality (in statistics, we talk about moving averages), but it does not exactly correspond to changes between 2010 and 2011.

71. Pending publication of the results dated 2011 (i.e., obtained from 2009-2013 surveys), the local data from the census can be compared only with the figures of the last “general” census of 1999. This scientific “warning” is not always heeded or well understood by users. It generates complaints from some elected officials (in localities where annual figures are trending down) and serves as an argument for persons (now admittedly few in number) who still oppose the change of method.

72. For INSEE, the analysis of annual trends, at least in large cities, is a means to control the quality of results: any significant change (taking into account the confidence interval of the sampling process) must be explainable either by a duly verified event (concerning the number of dwellings in the municipality or neighborhood) or by a regularly observed trend. Otherwise, the system may very well have malfunctioned. Thus far, we have confined these verifications to the population figures.

73. In the census system, each annual survey is supposed to be representative at national and regional level (NUT2). The sample for each annual survey was built to form, at these higher geographic levels, an independent survey comparable from year to year. INSEE has not yet conducted a systematic analysis of the quality of the resulting series, expect at national level only, for demography and employment. Some other agencies of the French government statistics begin to analyze data from annual surveys in their areas of expertise, and are more satisfied with the results, but the annual census surveys are not playing their initially intended role in official statistics.

74. From the end of 2013, it will, finally, be possible to compare the local data with those of the first cycle of the new census, and describe what happened between 2006 and 2011, which will be a great improvement. We started to analyze trends between 2006 and 2010 at national and regional level; trends are those that predict the specialized statistical sources. We started to analyze finer geographic levels, in some jurisdictions, in view of the dissemination of 2013-2014: the results are encouraging.

75. But it will never be possible to construct annual series.

IX. Conclusion

76. The French "rolling census" has undoubtedly fulfilled these initial aims: through regular collection whose quality is very satisfactory, INSEE establishes each year official population figures for all administrative districts of the country, and distributes many detailed data whose quality is recognized. However, this device has a high cost, mainly due to the very high number of households surveyed each year. Moreover, the device, methodologically complex, requires permanently a large data processing team.

References

Cézard, M., 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 (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.

Godinot, A , 2004, “Redesigning the population census,” Courrier des statistiques, English series no. 10, 2004, pp. 5-12, (available at http://www.insee.fr/en/publications-et-services/collection.asp?id=8&numpage=2&nombre=20)