



**Economic and Social
Council**

Distr.
General

ECE/CES/2006/SP/8
8 May 2006

ENGLISH ONLY

ECONOMIC COMMISSION FOR EUROPE

STATISTICAL COMMISSION

CONFERENCE OF EUROPEAN STATISTICIANS

Fifty-fourth plenary session
Paris, 13-15 June 2006
Item 6 of the provisional agenda

**SEMINAR ON POPULATION AND HOUSING CENSUSES
SESSION III**

Spanish 2001 and 2011 Censuses and the Internet

Submitted by the National Statistical Institute of Spain

I. GENERAL IDEAS ABOUT THE SPANISH 2002 AND 2011 CENSUSES

A. Introduction

1. There has been a long history of censuses in Spain. The first counts were already being taken for fiscal purposes in the 16th century. However, the beginning of the census dates to the second half of the 18th century, with the 1768, 1787 and 1797 censuses.
2. During the last century, census of dwellings and population resident in the country were undertaken regularly by law every ten years from 1900 to 1970, and then in March 1981 and March 1991. Building census were undertaken one year before and used to carry out those for population and dwellings.
3. For budgetary reasons, the 2001 Population and Housing Census was postponed from 1 May to 1 November. It was an extensive survey in which around 45 000 staff participated, specifically enumerators who visited 21 million addresses. The project was organised and conducted by the Instituto Nacional de Estadística (INE), with the co-operation of the

Ayuntamientos (municipalities) and, in some respects, of the Statistical Institutes of the Comunidades Autónomas (autonomous communities).

4. Ten years after the last census, many innovations have been introduced during the long period of preparation. These innovations include integration of the enumeration of population, dwellings and buildings, use of personalised questionnaires using data from administrative registers, a new relationship to the population register (Padrón Municipal de Habitantes), use of the Internet, and use of scanners and other features in the data-editing phase.

B. Padrón and the Census

5. The Padrón is the population register, legally established in each municipality to establish an individual's residence. Information is limited to name, identifier, residence, sex, date and place of birth, citizenship and educational attainment. It is regularly updated for births, deaths, and change of residence or other event on communication from residents.

6. The census covers the same population with reference to its usual residence, i.e. the main residence for one year in the case of multiple residences, but having the basic objective of simultaneously collecting fresh and broader information useful for statistical purposes.

7. In the 2001 Census the questionnaires were pre-printed with the Padrón data for confirmation or updating by respondents; the latter directly reported from INE to the municipalities. This facilitated a more fruitful data collection method based on the new relationship between Padrón and Census arising from a 1996 new general regulation and guaranteeing the right of access to the municipalities for the modification of personal data. At the same time, it was totally assured that properly census data were to be used only with a statistical aim.

C. Main features of the project

8. The main methodological characteristics of the 2001 Spanish Census were:
- (a) a careful evaluation of the questions to be presented in the questionnaires, with the objective of providing comparable information for the entire country and each autonomous community;
 - (b) preparatory work based on administrative registers: The Padrón is useful in locating buildings including main residences (i.e. dwellings resulting as the habitual residence of registered persons), and the cadastre in identifying remaining buildings having dwellings not used as habitual residence;
 - (c) two pilot surveys undertaken in the summer of 1999 and in the autumn of 2000 involving respectively 0.08% and 0.09% of the population;
 - (d) personalised questionnaires that already include key information extracted from registers, with most variables pre-coded, were delivered to each address by the enumerators. The return of amended information was then made available to the Padrón with no further burden to respondents;
 - (e) questionnaires were composed on separate A4 sheets with:
 - project's presentation and dwelling sheet;
 - padronal pre-compiled sheet;

- general individual data (marital status, occupation, residence at the time of the previous census, year of arrival in the municipality, autonomous community, etc.);
 - individual's current job and educational characteristics (only for people 16 or over who are working or studying);
- (f) use of two-language forms chosen from Castilian, Catalan, Galician, Majorcan, Valencian and Basque depending on the autonomous community, including provision of the possibility of completing and returning questionnaires over the Internet;
- (g) census data were complemented by previous padronal data in cases where questionnaires were uncollected. In particular, determination of the population figures took into account results of fieldwork, existing padronal data and other auxiliary information, i.e. matching the police database of personal identification numbers. Very detailed automatic searches for duplicates were also applied. This way, main sources of under enumeration and over enumeration were controlled and population figures are 'final', consolidated figures, which already take into account all the available information, so they are not affected by any later evaluation;
- (h) data processing was carried out after data capture by scanning, OCR, video keying and video-validation, automatic and assisted coding, and automatic imputation of missing and erroneous values. There were open answers for geographical variables, nationality, occupation and activity (for the latter two there was a pre-coded alternative). The open answer only had to be used (and, therefore, codified), if the occupation (or activity) had not been found in the list;
- (i) the first data dissemination way was the Internet, by means of the data warehouse, particularly. In the first two years after his setting up (February 2004), more than 2 million queries were attended, most of them just in the moment of the query and a very small minority just the day after. There is a tailored data table creation procedure and a list of predefined tables. The most used way is the tailored procedure, because of its capability of matching our customer requirements, but the predefined tables have been used also in some hundreds of thousands of data requests.

9. In line with the innovations arising out of this census round, future plans are to increase the use of administrative registers, although the design of the 2011 census is still in the first steps and there is no big decision taken yet.

D. A preliminary, non-definitive foresight of the next 2011 Spanish Demographic Censuses

10. (a) Increasing use of administrative registers.
- (b) Getting statistical data from every main dwelling will still be necessary.
- (c) Increasing use of the Internet in data collection and dissemination.
- (d) Last 2001 census data will be linked (register by register: by person, by dwelling) thanks to a new INE project named Demographical Longitudinal Study, capable also of linking with Vital and Migration Statistics; with other administrative register data and with the future 2021 and later censuses (with independence of the kind of census to be considered best in each moment).

II. REPORT ON THE COMPLETION OF THE SPANISH 2001 CENSUSES BY INTERNET AND IDEAS FOR THE NEXT 2011 CENSUSES BY INTERNET

11. We are going now to focus in the characteristics, achievements (and also failures) of the Spanish 2001 Censuses by Internet.

A. Essential summary:

12. The Census results obtained by Internet (in particular the low response rate) do not confirm – but neither do they invalidate – the possibility of Internet being, for the moment, a widespread and efficient tool for the collection of basic statistical information.

13. In any case, the INE complied with its obligation to implement all the latest available technologies towards a more fluent communication with users. To offer the possibility of the Internet helped very much to produce a very modern image of the Census.

B. Comments on the above summary:

14. Apart from questions related to the low response rate that will be analysed below, the project clearly reveals the need to envisage the use of Internet on account of the following facts:

15. The project made it clear that a reliable and robust collection system by Internet can be established against the possible problems inherent in this tool (there have been no significant difficulties derived from general system failures, intruder attacks, confidentiality protection, quality of received information, etc.).

16. Regarding the crucial issue of users authentication, the project proved the feasibility of an adequate authentication system, adapted to the statistical aims of the action based on reserved information known by both the user and the statistical office (names of the father and the mother).

17. Maybe a similar mechanism will do for the new projects (for instance, to request besides the indication of user and password, a specific item of the earlier period's statistical information that is only known by the respondent and the statistical office).

18. The project also showed the current difficulties in the use of the electronic signature (which was necessary for modified register data) and the need to be up to date on the Public Administration development of electronic signature mechanisms, particularly with the next setting up of the electronic version of the Spanish national identity card ("Electronic DNI").

19. Regarding the usefulness itself of the Internet characteristics, the results have been satisfactory: questionnaires were created for the blind, any co-official language and four foreign languages (including Arabic) were allowed, immediate compensation for respondents (providing for free the geographical distribution of surnames), usage of electronic mail by the Users' Attendance Centre, e-mail forwarding of the identification of the questionnaires completed by Internet to regional offices, co-operation with the Autonomous Communities statistical organisms towards the preparation of versions in co-official languages, usage of specific

advertisement channels for the general Census (Lottery tickets of the ONCE, i.e. the Spanish Association for the Blind, etc.).

C. Results

20. The number of households that opted for completing the census questionnaires by Internet was rather low (13768). 29,90% of households used electronic signature.

D. Possible causes of the low response

21. Thorough analyses led to the conclusion that no basic and essential cause can account for the low response and that this rate is due to a diversity of reasons, which will now be briefly commented upon.

22. In each case, it will be stated in italics why the cause under consideration cannot be fundamental in itself.

23. In each point a small modification is suggested for the future.

- Lack of publicity: More than once the action has been deemed insufficiently advertised.

24. However, as compared to the publicity carried out in the nearest partial antecedent (Switzerland), that of Spain was very similar (mainly based on the possibility of Internet answers being offered on a leaflet added to the questionnaire). In Switzerland, Internet was used by 4% of the households that were offered with the possibility of using Internet (those located in the largest communes).

25. *On the other hand, throughout the collection period by Internet, very specific publicity campaigns were exclusively devoted to the Census by Internet, in spite of which there has been no increase in the collection by Internet on later dates. The main campaigns were: Two reports on the Internet Supplement of El País; A Session devoted to the Census by Internet; a minute and a half spot on the Antena 3 television news bulletin.*

26. A change for the future could be the inclusion of a reference to the use of Internet to fill in the questionnaire in a more conspicuous place (outside of the census envelop, near the first question of the questionnaire, etc.).

- Coordination problems with the census interviewers: In fact, in some cases problems were detected that made it impossible to answer by Internet (deficient information from some interviewers, difficulties derived from the confusion in the payment of the questionnaires collected by Internet, etc.).

27. However, as soon as the problems became known, serious efforts were made towards an improved co-ordination, with the result that after a few days the census interviewers' know-how in the matter of Internet usage was quite appropriate (particular account must be taken of the

difficulty to accede efficiently to a set of over 40.000 people). Again, no significant increase was stated in collection by Internet after the above actions.

28. Taking into account, moreover, the drawbacks that may have occurred in the basic collection of paper questionnaires, the suggestion of a change for the future could consist in almost completely separating the different procedures, so that interferences would be avoided (in practice, we refer to the possibility to answer by Internet without it being necessary to receive first the pre-printed codes on the paper questionnaires).

- Delay in starting the action: The fact that the possibility to use the Internet started a week after the action with paper questionnaires may also have given rise to some problem or negative effect from the publicity viewpoint.

29. *However, it should be born in mind that the handing out of paper questionnaires was gradual, so that the households concerned by the above problem clearly were a small minority.*

30. *On the other hand, in the case of Switzerland there was a much greater initial problem (total failure of the system that resulted in a large negative publicity effect), in spite of which 4% was achieved.*

31. As a suggestion for a change in the future, it would be desirable that the system be entirely implanted before the official beginning of the action, allowing the users to experiment, to get familiar with the system, etc.

- Difficulties in the training of census interviewers: The above indicated delay was the reason why in most training courses for census interviewers, the possibility to answer by Internet was only mentioned in passing, since it was not yet functioning at the time of the courses, whose contents were already very dense.

32. *However, it must be admitted that this would not have been an insurmountable handicap if part of the households that habitually use Internet had answered.*

33. *In fact, in Switzerland the paper questionnaire was received and returned by ordinary mail, without the participation of interviewers, 4% having been obtained, as stated earlier.*

34. As a suggestion of a change for the future it would be desirable that the system be entirely implanted before the official beginning of the action, allowing the usage in all the training courses for the census operation.

- Difficulties with the electronic signature: This type of authentication was essential for those questionnaires in which important errors in register data had been found previously or for those in which the respondents indicated that such data had to be modified.

35. *The initial approximate estimate was that these circumstances could concern 20% of the households and the data obtained so far (both of the edited questionnaires and of the analysis of the interrupted access when the electronic signature was required) are coherent with that figure. This implies that (the existence of a handicap still being possible) the need to use the electronic*

signature does not account for the low response rate in the households for which the use of the signature is not compulsory.

36. *Furthermore, since the register data are of an administrative character, the use of the electronic signature was imperative for the execution of the census.*

37. As a suggestion of a change for the future, it is desirable to envisage other authentication methods that would be equivalent to the electronic signature (identification in a public office that would provide the appropriate passwords in case the register data are modified, etc.). An alternative would be the preliminary indication in the personalised questionnaires of the obligation to use the electronic signature, thus avoiding later disappointments. Obviously, all the developments of the electronic version of the Spanish national identity card (“Electronic DNI”) will need to be considered.

- Other technical difficulties: There were other technical difficulties of different types (regarding passwords, questionnaires with several households, errors in understanding the names of the father and the mother, slowness of some connections, usage of navigators incompatible with the application, etc.) that could cause some users not to try to answer by Internet.

38. *Some of these problems were almost unavoidable. However, their thorough analysis during the collection period by Internet made it possible to solve some of them (for example it was finally possible to use MAC computers). Also, in most cases, the specific difficulties of some users who were adamant about answering by Internet could be cleared.*

39. *The overall impression of those working on this support is that the cases described were few and often derived from each user’s particular problems (this was particular conspicuous when they were related with slowness). Anyway, their absence would not have allowed a spectacular response by Internet.*

40. As a suggestion of a change for the future, it would be very desirable to plan a larger experimental period that would offer the users the possibility of intensive trials. This would reveal their problems and enable solutions according to their nature, to the navigators used, etc. It is important to consider here the big increase in the number of the Internet users, using broadband connections particularly, and the improvement in the standardisation of the browsers. To be studied also is the possibility of using some of the Censuses premises by people willing to use the Internet for the filling in of the censuses forms but without Internet connection at home.

- Lack of incentive to fill in the questionnaire by Internet instead of on paper: The fact that the questionnaires were delivered and collected at home, that they were simpler and, above all, easier to complete than by Internet (without taking account of web traffic “jams” and possible costs for the users), must have led to the result that even those users who handle Internet quite naturally (for instance for their tax returns), finally preferred the paper questionnaires.

41. *In Switzerland too, the completion was easier on paper than by Internet and nevertheless the response rate was much higher.*

42. *The provision for free of the distribution of surnames was rather successful but cannot be*

compared with the incentives involved in the tax return by Internet compared with doing it on paper: instead of being free, the paper form has to be bought and delivered in a bank or at the Ministry of Finance; the very usage of Internet makes the tax return much easier than by hand on a form, negative tax returns on Internet are reimbursed earlier, etc.

43. As a suggestion of a change for the future, and as indicated before, the possibility should be considered of a completion by Internet independent (probably earlier, as occurred in Singapore) from the one on paper. In other words, we feel the need to avoid any kind of “disloyal competence” between the two data collection procedures.

- Scarce presence of Internet in the households: It is possible that the low response rate is due to a much lower presence of Internet in the Spanish households than could be expected.

44. *However, any indicator of this presence is in contradiction with the response obtained. For example: when the Census by Internet was carried out, about 300.000 Spanish households had ADSL (installed by Telefónica). As is known, these households pay around 42 euros per month to be connected with Internet at high speed and for the duration they wish. They are, therefore, very eager to use Internet (if only to justify the costs). It is not clear that we succeeded in achieving this target population.*

45. As a suggestion of a change for the future, it would be advisable to carry out specific campaigns with well-defined target populations (ADSL users, computer enterprises, and users of Internet by cable. It is important again to consider here the big increase in the number of the Internet users, using broadband connections particularly.

E. A key point: Two different ways of using the Internet in next Census data collection

47. For the moment, Internet will be just another way of collecting data, in the context of a multimode strategy, but there are two very different ways of implementing such a mixture of procedures:

- (a) to use the Internet as another data collection procedure, between others, and without using it in order to reduce the overall cost of the Census, because all the traditional procedures would be implemented as if they were to do all the job;
- (b) to prioritise the Internet (and the telephone) as the first possibility (chronologically) to fill in the forms by the households and with enough publicity to achieve a big take-up. Only the households not using those 2 possibilities (Internet and telephone interviewing) will be contacted with the traditional procedures, like the strategy used in Singapore in 2000.

48. Obviously, the second way is more complex but the saving in the costs would be huge.

49. It would be also very difficult to forecast the amount of households to use a first method and we would need a big flexibility in order to make the necessary changes if there is a strong difference between reality and our forecasts related to the number of households using the Internet to fill in the Censuses forms or reachable by telephone interviewing.

50. To finally establish the optimal strategy in those matters will be, probably, one of the more

relevant methodological and logistic challenges of the next Spanish censuses.

* * * * *