STATISTICAL COMMISSION and ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Forty-seventh plenary session
(Neuchâtel, Switzerland, 14-16 June 1999)

EVALUATION OF A NATIONAL STATISTICAL INSTITUTE

Supporting paper presented by the Spanish National Institute of Statistics

1.   The demand for statistical information is ever more important and its role
is increasingly significant for the elaboration of policies, distribution of
resources, both national and international, and the decision-making by the
economic and social entities. The NSIs have to take up this challenge on a
background of budgetary restrictions and the Public Administration's desire to
determine their activities in view of the results and to show taxpayers their
money is being used well.

2.   Output indicators and evaluations have become indispensable tools both in
helping the NSIs to improve their efficiency and to achieve a greater
accountability.

3.   The evaluation procedures and the output indicators are more or less
consolidated depending on each State's administrative situation. Every
Statistical Office may decide to complete the Public Administration's procedures
with its own, probably more specific. In order to obtain these particular
indicators, international comparisons and good practices identification are very
useful. However, it should be born in mind that the above mentioned procedures
are not in themselves the objective but only a means to achieve a better
management and accountability. That is why one cannot be too careful with the
resources devoted to the assessment and the elaboration of output indicators.

4.   This document does not tackle the general Spanish Administration's procedure which are treated, as far as evaluation is concerned, by the 'Intervención General del Estado' (*General Accounting Office*) and by the 'Inspección General de Servicios' (*General Inspection of Services*). The latter assesses the financial performance and the latter concentrates on the following up of general administrative standards. Much to the contrary, this document describes the initiatives of INE towards an improved efficiency and accountability. A distinction is made between those indicators designed to measure and augment the overall INE's results and those associated with individual statistical operations.

I. Planning

5.   Planning is an essential element for the correct performance of a complex organization, endowed with a high budget and a numerous staff.

6.   The carrying out of a medium term planning within a short term planning is a fundamental indicator for the evaluation of an NSI. The degree in which these plannings are fulfilled gives the qualitative measure of this indicator.

7.   The INE conducts a quadrennial Statistical Plan for the activities it is going to undertake during these years. This Plan is part of the National Statistical Plan which includes the overall General Government statistical activity and has to be sanctioned through a legal Government regulation after being duly briefed by several consultative entities. The annual Programmes worked out by the Plan and which are used to introduce modifications, will also have to be laid before the 'Consejo Superior de Estadística' (Higher Statistical Council), a consultative body of the State statistical services, and approved by the Government.

8.   Every year, a *Balance* is made on the actual activity of the previous year: it accounts for the fulfilment of the programmed objectives and every four years, after the termination of the National Statistical Plan, a *Memoire* is written on the latter's execution.

9.   Both the annual Balance and the quadrennial Memoire are submitted to the *Higher Statistical Council* and disseminated to the media.

II. Recent demands

10. Closely connected with planning is the identification of the INE users' demands, whose satisfaction depends on two factors: which products are available and in what format. In terms of statistical production, this means: what statistics are being carried out and how are they disseminated.

11. Most of the INE production is made up of periodical statistical operations whose need is unquestionable. These are included in the Statistical Plans, and

their continuity is discussed during their elaboration. Other requirements refer to the coverage of new areas of knowledge and to the integration of information from different sources, i.e. to the enlargement or modification of a statistic or the presentation of current results. The compliance with these recent needs depends on cost-efficiency criteria which are in themselves a quality indicator of the Statistical System as a whole.

- New areas

12. The coverage of information gaps is a constant worry. On some matters of social concern, there is an increasing demand of high quality information, examples being statistics on environment, on the information society or on the use of time.

13. Lately, the INE started research on environment.

- Information updating

14. In other cases, aperiodic surveys have to be updated , for instance, the INE is updating those on fertility or disabilities which have been carried out more than ten years ago.

- Aggregability

15. The INE has to meet an ever-growing demand of aggregated information through systems that make it possible to associate the results of different statistics and to present global outcomes. The systems in question also allow answering a great amount of ad hoc requests.

16. At present, an integrated system of industrial statistics is being created and a demographic report is worked out.

- Integrability

17. A basic measure towards the assessment of statistics on a given sector or subject consists in integrating at least their basic results into sectorial and satellite accounts describing synthetically the sector's situation.

18. In this sense, the INE is setting up the Satellite Account on Tourism, the Satellite Account on Environment Expenditure and Protection, and Accounts on Water.

- Presentation

19. The conventional dissemination on paper is increasingly being replaced by speedier, cheaper and more flexible products, such as electronic publications, CD-ROMs with 'anonymized' data files, or databases on Internet.

III. Meeting the clients' needs

20. Clients demand that the information they ask for supplies a number of characteristics besides traditional quality. Their requirements give rise to the construction of indicators.

III.1 ADVANCED DISSEMINATION DATES

21. The publication of calendars indicating the dissemination dates of statistical results is a fundamental quality indicator. Deviations from this calendar are the indicator's quality measure.

22. Since 1987, the INE publishes every year, without the slightest delay, the calendar with the exact dissemination date of the results of short-term statistics.

23. In 1999, the INE published for the first time a yearly calendar with the approximate date of the dissemination of structural statistics.

III.2 TIMELINESS

24. The lapse of time between the data's reference date and the dissemination date of the results determines their usefulness, it thus being an indicator of the statistic's punctuality or timeliness of a statistical figure.

25. In recent years, the INE achieved an ever shorter delay between the collection of primary data and the dissemination of results of all short-term statistics; for some of them this delay is so small that it can hardly be improved.

III.3 ACCESSIBILITY

26. Another qualitative indicator for the evaluation of NSIs, is the easy and equal accessibility to statistical results.

27. The INE works as follows: it incorporates the main results of all its statistics into its Web site; it regularly summons a press conference to inform about the most important structural statistics results and it publishes press communications as soon as new statistics become available.

28. However, specialized users and the media wish to receive a special treatment and the INE devotes specific resources to them. Consequently, the INE simultaneously sends by megafax to the Press, the General Government Organisms, the Autonomous Regions and the Social Entities, the main results of the short-term statistics the very day they are disseminated.

III.4 METHODOLOGIES

29. In order to facilitate the user's understanding of the statistical results, the current methodology should be widely disseminated and of easy access, this being yet another qualitative indicator to be born in mind.

30. At the INE, the public may accede to the current methodologies through the monographic publications that proved necessary because of their importance, as well as through the methodological notes in the publications of all the statistical results. They are also available in the freely accessible INE Web site.

III.5 COMPARABILITY OVER TIME

31. The maintenance of historical series for a given statistic is also a quality indicator, which the INE is particularly eager to take care of by several means.

32. When a statistical operation is modified, account is taken of all the factors that are liable to repercute on the comparability of results; either the users are kept informed for them to make their own adjustments, as occurs for structural statistics, or the INE itself established the linkage of series as is the case for short-term statistics.

33. The INE keeps the TEMPUS data base permanently updated with historical series of the main statistics which allow any type of analyses.

34. Apart from the creation of indicators suggested by the characteristics users wish to know, the way the latter apply our system is a source of very interesting indicators. In this context, the following magnitudes are significant both regarding levels and changes in relation to the previous period:

- The amount of ad hoc demands and their economic value
- The amount of accesses to the Web server
- The amount of consultations answered by the information service
- The amount of copies sold for each publication
- The amount of parliamentary consultations that have been met
- The amount of library visitors

Whenever technically possible, indicators should be broken down by subjects.

IV. Resources

35. In order to assess an NSI it is necessary to know the available human, economic and material resources.

36. The staff of an NSI is its main capital, not easy to be replaced within a short time because of its specialization. The staff's professional qualification, the way it is selected and its continuous training are NSI quality indicators.

37. The INE selects its technical staff, both of the highest and medium level, by means of competitive examinations. It also has a complete programme of continuous occupational training for all the staff categories, including courses in languages, computer science, offimatics , management statistics.

38. The INE has at its disposal a permanent personnel for interviews who are not usually specialized according to surveys, except in the larger provinces: since there are 52 provincial offices for data collecting, this activity has been ordered by territories with the advantage of the respondents being on the spot. Because of this scattered collection, it has been necessary to create global indicators to relate production capacity with the actual production, i.e. synthetic indicators allowing a common measurement of very heterogeneous statistics' working loads. Thanks to this kind of information, it is possible to take specific decisions on the distribution of different types of resources.

39. The basis of these studies is the information, collected twice a year, on the time consumed by the INE staff for each statistical activity. The resultant data, together with those stemming from the economic management system endowed with a subsystem for analytical accounting, enabled the INE to work with a new method to calculate costs. For several years already, the latter has made it possible to know not only how the budget is distributed among the different statistics, but also what has been assigned to each stage and each unit.

40. The material resources, particularly those for information and communication technologies, have become indicators of a NSI's modernization and indirectly of its efficiency. The number of posts having Pcs connected with an internal network, the number and capacity of central and departmental computers, the storing capacity on and off line, the band breadth of the internal and external communication networks, are some of the possible indicators.

V. Quality

41. To measure the quality of a statistical operation, account should be taken of all its stages, from planning to dissemination.

V.1 METHODOLOGICAL PROJECT

42. A first quality indicator is the availability of a methodological project on the way to carry out a statistic, with a special emphasis on the national and supranational legal framework within which it is elaborated and on the application of internationally approved standards.

43. The INE attaches special importance to this aspect and requires that a methodological project be closed before starting a new statistical operation. Moreover, these projects are taken to the following statistical bodies for discussion and briefing: Interministerial Statistical Commission made up of representatives of the Ministries; Interterritorial Statistical Committee with representatives of the Autonomous Regions and the *Higher Statistical Council* whose members represent the social entities.

V.2 FOLLOW-UP AND CHECKING OF THE DATA COLLECTION

44. It is fundamental to dispose of following up and checking that make it possible to quickly obtain information during the survey's data collection on the field: this allows to take the necessary measures when unforeseen situations arise that might impair the quality of the final results.

45. In view of their purpose, these indicators should be supplied weekly or fortnightly, according to the type of survey. The necessary basic information refers to the percentage of the different types of collection incidents or situations, including the collected and processed units, in relation to the regular sample. Each type of incident has different causes and the measures that may be necessary to correct them also vary.

46. An interesting information for the users is the non-response rate broken down by refusals and inaccessible units; it is published by the INE for its most important surveys.

V.3 THE ASSESSMENT OF ACCURACY

47. Unlike the above indicators, those specifically designed to assess the surveys' accuracy, are obtained after fieldwork is finished. They essentially approach sampling errors and each incident's final percentages.

48. Sampling errors show directly the estimates dispersion caused by the way sampling has been carried out, and of which users of statistics should be informed. That is why the INE not only calculates all its statistics' sampling errors but also reveals them in most of its publications. The ideal theoretical measure of data quality would consist in measuring directly the bias together with the variance of the estimator in order to calculate the average quadratic error. However, as this is not easily achieved, it is almost unavoidable that data quality be evaluated by means of response rates.

49. The latter is calculated for all statistics and disseminated in most publications. On the contrary, information on automatic imputation procedures and on its volume is usually restricted to inside use.

V.4 EVALUATION SURVEYS

50. For major surveys, it is desirable to establish a continuous programme of repeated interviews, which allows the measurement of errors of content. The main drawback of repeated interviews is its cost, both for the respondents obliged to give data they already supplied and for the INE.

51. The INE conducts evaluation surveys for the Labour Force Survey (LFS) and the Consumer Price Index (CPI). The corresponding repeated interviews are the task of 16 full-time employees of the INE central services.

52. For the LFS, there is even an indicator of the response bias of the fundamental variables generated by the net change index. The repeated interviews programme, besides providing the public with information on the quality of the LFS, has been useful in identifying the most unstable variables, whose definition has improved by modifying the wording of the questions. It also has been resorted to for analysing the effect on quality of the different collecting procedures (personal or telephone interview).

53. This type of evaluations should also include those sporadic studies that are undertaken to know minutely the collecting procedures for specific surveys. In the most recent one, it has been suggested to improve the efficiency of data collection for structural economic surveys. To this end, it has been necessary to obtain data for parameters such as the average number of calls per enterprise in relation to the final incidence, the average number of days elapsing before a questionnaire was collected, recorded and processed, as well as a general analysis of the collection cycle. It has thus been possible to evaluate, for a type of surveys with a random factor referred to the time a respondent takes to send his information, when working peaks occur and which are the best ways to distribute the work load during the collection.

V.5 EVALUATION STUDIES

54. Sometimes it is desirable to thoroughly revise a given group of statistics, especially when their information gives rise to disputes or critical social discussions. This is what happened in Spain in the matter of labour market statistics, particularly because of discrepancies between data from the Labour Force Survey and those proceeding from administrative sources such as the Social Security and Registered Unemployment. The *Higher Statistical Council* created a Working Group made up of representatives of the organisms that produce these statistics and of their main users, whether belonging to the Administration itself or to Universities and trade union and business organizations. The Working Group is about to deliver its first report, which includes technical recommendations and advice on how to bring forward the results. One of the Group's objectives is to increase the public's confidence in statistics, explaining why discrepancies may occur between the results stemming from different sources.

VI. Reduction of costs

VI.1 THE RESPONDENTS' RESPONSE LOAD

55. NSIs have to face the challenge of offering ever more information while diminishing the respondents' trouble. This may be achieved thanks to replacing data directly obtained from respondents, by administrative data, simpler forms or the use of a single one for several purposes. By designing the questionnaires for the statistics of enterprises according to their own accounts, which greatly facilitates the filling in, the response burden has diminished. Labour Force or Family Budgets statistics are used to carry out as well specific modules on other subjects. The enterprises' industrial survey will allow the obtention of information on environment.

56. In order to reduce the load of individual respondents, an endeavour is made to avoid whenever technically possible, that the same unit informs in more than one survey. This is applicable particularly to household surveys and to those on small and medium enterprises.

57. In decentralized systems like that of Spain, it is essential to avoid double-counts and to this end, agreements are reached with the Autonomous Regions. Business associations are also contacted in order to determine those data collection procedures they deem less burdensome.

VI.2 COST OF STATISTICS

58. NSIs are required to put out an ever increasing production with budgets that are, at best, frozen.

59. That is why costs should be reduced at all stages, particularly the most expensive ones, such as data collection, recording and processing.

60. In this context, an indicator to be assessed is the application of alternative methods.

61. The INE has been rationalizing the costs for years through several measures. For instance, in order to reduce data collection costs, some surveys are conducted by telephone. In others, involving consecutive interviews with the same unit, the second and the following ones are replaced by telephone calls. Respondents are offered the possibility to call free of charge for consultation, thus avoiding personal interviews.

62. With a view to lower recording and processing costs, use is being made of hand-held computers which carry out a first processing of the information recorded *in situ* by the interviewers. Pre-printed questionnaires already bearing

the identification variables when these have not changed or questionnaires prepared for optical scanning also proved to be very efficient.

63. Manual coding is being replaced by automatic or computer assisted coding.

64. Instead of specific programmes for ad-hoc requests, general systems are introduced that are easily understood by the dissemination services.

65. Dissemination on paper is drastically cut down in favour of cheaper means.

VI. Conclusion

66. Summarizing the above arguments, the following indicators are suggested:

- Indicators for a National Statistical Institute in general.
- Medium-term planning
- Memoire of the Plan's execution at the end of the period
- Short-term planning
- Annual execution balance
- Selection system for technical staff
- Staff by categories
- Training courses at several levels
- Budget by categories
- Cost of statistics
- Number of large, medium and small computer systems
- Communication systems (local networks and their interconnections)
- Volume of data managed by the systems and number of applications
- Annual purchases and budget for computer equipment (hardware and software)
- Sectorial or satellite accounts that are calculated
- Integrated systems under consideration and statistics referring to them
- Dissemination calendar with the exact dates on which short-term statistics are published
- Dissemination calendar with the approximate dates on which structural statistics are published
- INE Web page, number of accesses and volume of information supplied
- Number of information requests and consultations attended to
- New products

- Indicators for assessing the quality of specific statistics.

67. In view of the great variety of statistics the NSIs are tasked with, a preliminary selection should be made of those considered the core, which will then be used to measure the indicators designed to assess an NSI as a whole.

- Availability of the methodological project.
- Compliance with legality
- Use of standardized principles
- Indicators for collection following up and checking
- Evaluation surveys
- Sampling errors
- Response rates
- Recording errors
- Volume of automatic imputation
- Indicators for sales and ad hoc requests
- Gap between the data reference date and the day the results are disseminated
- Inclusion of methodological notes in the dissemination of results on all its possible supports, in particular the Web page.
- Evaluation of the respondents' response load
- Cost by working stages