

WORKSHOP ON LABOUR FORCE SURVEY METHODOLOGY

7 th WORKSHOP ON LABOUR FORCE SURVEY METHODOLOGY

DATA PROCESSING AND DATA QUALITY

Madrid, Spain, 10 – 11 May 2012

Opening session

Introduction to the Spanish LFS

Florentina Alvarez - Spain

The Spanish Labour Force Survey (Encuesta de Población Activa)

Florentina Alvarez
National Statistical Institute

General description

The Spanish Labour Force Survey (Encuesta de Población Activa) has been conducted from 1964. It has experienced a considerable number of changes in its periodicity, methodology, questionnaire, field work, etc. It is currently a quarterly continuous survey whose target population is people living in private households.

The effective sample size is approximately 65.000 dwellings (180.000 people, 0,5% of the resident population) with the objective of giving reliable estimates of the main labour market variables for each of the 52 provinces (NUTS 3) in Spain. It has a two-stage stratified sampling, with stratification of the primary units which are censal sections. The second stage units are private households. Each primary sampling unit is visited in one of the 13 weeks of the quarter, being the sample homogenously distributed throughout the quarter. A sixth part of the households sample is renewed each quarter and, in this way, there are 5/6 overlap between two consecutive quarters. A ratio estimator is used, having the population projections as auxiliary variable, with a final calibration to adjust the results to the population distribution. In this final step the auxiliary information is: population projections by sex and five years age groups at NUTS 2 level; the population by nationality and age (16 and more, under 16), also at NUTS 2, and the population by age (the same two groups) in the provinces.

The first interview is personal, made with CAPI, and it is organised in a decentralized way, given that the Institute has delegations in all the provinces. The CATI mode is used in second and consecutive interviews, except when the household does not have telephone or when the family prefers personal interview. There are 7 CATI centres covering all the territory.

The information collected is sent weekly to the INE Central Office and edited monthly. The quarterly results are published the last Friday of the month following the reference quarter.

There are also two types of annual results: the average of the quarterly ones and the data coming from the 'subsample system' (wave approach).

Tables, series and anonymised microdata files are available for users on the INE website. Tailored tables are also prepared.

The role played by the field work in the quality of the EPA data

The CATI mode was introduced in 2005 jointly with a new questionnaire. Previous research confirmed the influence of the interviewer effect, mainly due to the differences in the way of asking the questions. The necessity of achieving standard interviews and better monitoring the data collection was clear.

Consequently, the EPA questionnaire was redesigned to make it easier to be understood by the respondents, without additional explanations, and to lighten the interviewer and respondent burden. That meant:

- new ordering of paragraphs
- new wording of some questions and different way of grouping them

Special focus was put on questions about search of employment, working time and main labour status.

Qualitative tests were developed to find the best wording for questions on hours worked and searching of work, and a pilot survey was conducted to test different models of questionnaire and the effects of the introduction of the CATI mode. The increase of non-response was one of the most important consequences of the new questionnaire.

New changes were introduced in the questionnaire after 2005, to adapt it to certain population groups with very stable situations (retired persons aged 65 years or more, people with disabilities, workers with permanent contracts, etc.) with the idea of simply confirming the data given in previous interviews and shorten the duration of the current one.

The electronic questionnaire is used in CAPI (in laptop computers) and CATI.

After the introduction of CATI mode:

- the standardisation of the interviews is assured
- it is easier to monitor the data collection, because the supervisors could listen the interview, watch the PC screen of the interviewer, etc.
- there are longer timetable availability to conduct the interviews (9 to 21 hours, in two shifts)

However, this system imposes a limit on the duration of the interviews (it should last not more than 30 minutes). It is difficult also to check empty dwellings and changes in the telephone numbers.

Other peculiarities and current tasks

The results refer mainly to persons, but **data on households** are becoming important in the last years because they show how the family relationships mitigate the effects of the high unemployment level.

Flows statistics are also produced quarterly, calculated as transition probabilities. They are established from the common sample of two periods. The production of **gross flows** is under study.

Administrative registers are used to produce the variable 'wages from the main job' to diminish the burden on respondents. On the other hand, the EPA microdata files are matched with the Claimant register and the Social Security register, to test the coherence of the unemployment and the employment figures. The

construction of a new statistic of migrations, based on the combination of EPA microdata with information of the Population Register, is in progress.

Different models to obtain **monthly data** on employment and unemployment have been tested.