

Economically Active Population Survey

Double encoding matrices CNAE 93 - CNAE 2009

Background

Each quarter, the Economically Active Population Survey collects, among many other variables, information regarding the activity of the establishment where employed persons work (in both their main employment, and as may be the case, their secondary employment) and, for those persons who are not employed but have professional experience, the activity of the establishment where they last worked, if they stopped working one year ago or less.¹

These variables are encoded following the corresponding official statistical classifications.

During the year 2008, the EAPS began the encoding of its activity variables with the new National Classification of Activities 2009 (CNAE-2009). A brief reference regarding the change to this classification may be viewed at:

http://www.ine.es/en/daco/daco42/clasificaciones/cnae09/cnae2009_en.pdf.

In order to build the link in the series in which the economic activity variables intervene, throughout the four quarters of 2008, the activity variables were encoded, using both the new CNAE-2009 classification and the previous CNAE-1993 classification. Both encodings have been included in the anonymised microdata files that are made available to interested users, so that they may analyse the correspondence between the two classifications over the year 2008.

For the purpose of facilitating the transition between both classification to periods other than 2008, the online page of the INE dedicated to the EAPS includes a conversion matrix, calculated from the **double encoding of the activity of the establishment of the main employment for employed persons in some of the four quarters of 2008**. It is considered that this collective is the most ideal for providing a robust estimate of the empirical correspondence between the two classifications.

In any case, we must bear in mind that **the correspondence between both classifications**, for those groupings in which there is no conceptual identity in their content (that is, when there is no one-to-one correspondence), **tends to be less robust, the further it is from the period in which the double encoding has been performed** (the year 2008, in this case), given that the activity structure itself of the Spanish economy is dynamic.

The following describes the matrix and how the retrospective series from the 2000-2007 period were calculated with CNAE-2009.

¹It also obtains information regarding the activity of the establishment for those persons with professional experience who stopped working eight years ago or less, but only in the case of those persons who are on their last interview. This information is used in the annual subsample file of the EAPS, and does not offer quarterly results. The annual file also offers the information on the activity of the establishment one year prior.

Description of the matrix

The conversion matrix has been determined for a **breakdown to two digits in both classifications**, considering this election to be the optimum point between the degree of detail necessary and the reliability of the estimates obtained from the double encoding on a 2008 annual average. **The estimates for activity groupings have been reached through aggregation of the corresponding activities comprising them to two digits.**

The matrix has been calculated as the **annual average of the four quarters of 2008, tabulating the main activity of the total employed persons on a division level (two-digit codes) of both classifications.** There are 60 divisions in CNAE-1993 and 88 in CNAE-2009. The information obtained in 2008 can be used for both retrospective estimates with CNAE-2009 and future estimates with CNAE-1993, by simply applying the adequate percentage distribution. This therefore provides, for the mentioned division breakdown level, a transition matrix from CNAE-2009 to CNAE-1993, and another from CNAE-1993 to CNAE-2009.

To facilitate the use of the matrix, it has been developed "sequentially", by virtue of the distribution of each of the categories of one classification in those of the other.

Use

The use of the matrices is, conceptually, very simple. It is a simple transformation of vectors, multiplying one vector (that of the activity categories according to one CNAE) by the resulting matrix of the other vector (that of the activity categories according to the other CNAE).

We can illustrate the use of the matrix with an example. We will use the transition matrix CNAE-1993 to CNAE-2009 corresponding to the CNAE-2009=68 activity, *Real estate activities*, to estimate the proportion of employed persons in the second quarter of 2007 who were working in this activity from CNAE-1993=70, *Real estate activities*:

The number of persons employed in this branch in Q2-2007 (INEBASE results) is:

Economically Active Population Survey CNAE-93 series (Tables of final results obtained with National Classification of Economic Activities 1993)

Employed persons, by sex and branch of activity. Absolute values and percentages as compared with the total for each sex

Units: Thousands of persons and percentage

	Both sexes	
	Absolute value	
	2007QII	
70 Real estate activities		175.8

We will use the sequential development of the matrix for CNAE-93=70;

NCEA-93 Rev.1	NCEA-2009	Percentage	Subtotals.
70	41	23.5414	100.000
70	68	60.3939	
70	81	16.0647	

The previous matrix indicates that the total for the branch of real estate activities according to CNAE-93 is distributed into 23.5414% in CNAE-2009=41 ("Construction of buildings"), 60.3939% in CNAE-2009=68 ("Real estate activities" according to CNAE-2009) and 16.0647% in CNAE-2009=81 ("Services for buildings and gardening activities").

Thus, the total persons employed in real estate activities in Q2-2007, according to CNAE-2009 would be:

Result: $175.8 \times 60.3939\% = 106.172476$

Aside from the use itself that each user may make of these matrices for the transition between classifications, it is always possible to ask the INE information services to execute the tables ad-hoc, using the customary paperwork.

The results obtained will, in any case, be those resulting from the application of this double-encoding matrix to two digits, except for the occasional rounding errors caused by precision in thousands of persons in the INEBASE database.

Madrid, 24 April 2009