

# **Wage from the main job in the Economically Active Population Survey (EAPS)**

**Statistical treatment of Social  
Security Registers and Tax  
Authorities**

# Index

1	Purpose	3
2	Scope	4
2.1	Time scope	4
2.2	Population scope	4
3	Statistical concept of wage	5
4	Information gathering	5
5	Estimation phases	6
5.1	Obtaining the personal identification number	6
5.2	The link to the Record of Registration in the Social Security System	6
5.3	The link to the Record of Annual Wage for the Tax Authorities	7
5.4	The link to the Social Security Contribution Bases	8
5.5	Integration, filtering and imputation	8
5.6	Encoding	9
6	Dissemination of results	9
7	References	9

---

## 1 Purpose

Regulation (EC) no. 1372 / 2007 of the European Parliament and Councils establishes the legal obligation to include the variable “wage from the main job” in the Labour Force Survey of the European Union (whose equivalent in Spain is the Economically Active Population Survey), given that the “classification of salaries by socio-economic characteristics, as well as in relation to the different types of paid employment are essential to analysing and making sense of the employment market and the changes occurring in the structure of the economically active population”.

In this way, the analytical potential of the survey is improved, by introducing as a classification variable the level of remuneration in the analysis of the characteristics of the main job for the group of workers employed by others or wage-earners.

The regulation only requires that the variable “wage from the main job” is encoded in deciles and allows information gathering by means of an interview or else using administrative records. In the Economically Active Population Survey, after carrying out a range of qualitative tests in the year 2003, the second option was chosen.

The main advantage of using administrative registers for estimating the variable are, on the one hand, that it does not increase the response workload of informant, and on the other, the survey is not affected by a lower response rate overall.

The main drawback is that it takes longer to gather the data, dependent on when administrative registers are available for linking. An added drawback is the possible variation of the rules of said registers over time.

The apparently most straightforward solution of including the variable in the EAPS questionnaire, also gives rise to validity problems. In fact, where information is obtained directly from informants, Ormeron and Ritchie (2006) have shown a significant rounding effect, given that a wage-earner is more likely not to be home (out at work) when the surveyor calls in person or phones him or her and is given the response regarding wage by interviewing another family member.

Conversely, the same writers (2007) believe that using administrative registers means that information gathered is more reliable than that obtained directly from the household informant, since the method is based on documentary evidence.

This last hypothesis was confirmed in a pilot survey carried out by the INE in 1990 (Pilot Survey on Earnings and Underemployment).

---

## 2 Scope

---

### 2.1 TIME SCOPE

The variable wage from the main job is contained within additional structural variables or a subsample providing information on an annual average over a reference year, by grouping subsamples corresponding to around 40.000 families ending their collaboration in the survey in each of the four EAPS calendar quarters comprising said year.

---

### 2.2 POPULATION SCOPE

The variable is estimated for all persons working for others or employees pursuant to the definition established in section 3.3.A) of the survey methodology (see [http://www.ine.es/en/daco/daco43/resumetepa\\_en.pdf](http://www.ine.es/en/daco/daco43/resumetepa_en.pdf)).

Next the definition appearing in said methodology is reproduced:

“It comprises all persons in the following categories:

- *Working: persons who, during the reference week (the one prior to the one in which the interview is held) have worked, even sporadically or occasionally, for at least one hour, in exchange for a salary or other form of remuneration in cash or in kind.*
- *Employed but not working: persons who, having worked in their current job, are absent during the reference week and are closely linked to the job.*

*Persons absent from work due to illness or accident, holidays or maternity leave are considered to be employed persons.*

*Those absent due to having discontinued permanent contracts, due to being seasonal workers or due to waiting to start a new job are regarded as unemployed.*

*Those absent for reasons other than the previous are regarded as wage-earners if the employer pays them at least 50 per cent of their salary or if they expect to be able to return to the job in the next three months.*

*Apprentices who have received remuneration in cash or in kind and students who have worked full-time or part-time in exchange for remuneration are considered wage-earners and classified as working or not working, based on the same grounds as the rest of persons employed by others.*

*Active members of the armed forces are considered among the wage-earning population.*

*The following persons are not considered wage-earners:*

- *Entrepreneurs and independent workers, even if under contract as wage-earners in their own company.*

- *Unpaid family workers (family assistance) and members of cooperatives who work in these.*
- *Persons absent from work, on leave or suspended, weakly linked to their job.*
- *Seasonal, occasional or discontinuous workers employed by others in the period of least activity, who did not work during the reference week.”*

---

### **3 Statistical concept of wage**

The concept of monthly wage from the main job has been agreed on a European level, as a reference definition enabling comparison between the different countries and at the same time flexible, making it possible to estimate using different methods. In the Spanish case, it is referred to the monthly wage from the main job which is realized in the reference week of the survey, before deductions on income from work and obligatory contributions by the worker to social security schemes (i.e. gross monthly wage).

Conventional criteria have been added with regard to special cases. Thus, extra payments and other annual supplements are considered in the definition, dividing their amount by 12. Regular overtime payments, tips and commissions, referring in this case to the calendar month prior to the reference week are also included.

In addition, if it is a temporary job lasting less than one month, the wage will be an estimate of the wage that would be received for a duration of a whole month for a similar job (the person, at least in principle, could carry out other jobs of that nature in the same month). Conversely, if the work is part-time, it should not be transformed to full-time equivalent (this is the main job, and therefore the one providing the most significant income).

---

### **4 Information gathering**

As shown in point 1, instead of directly asking persons from the sample, the preference has been to obtain information by using administrative registers, but given that there is none that complies with the strictly-established definition, due to having different methodologies and limited coverage, the solution adopted has been to obtain information from several registers and, based on additional data, try to find the best estimators of the target variable.

Therefore, it is what has come to be known in statistical literature as a “derived variable”, since the need for information may not be covered immediately by direct reference to the available information. This variable is available through chaining a set of different sources that provide the information required.

Following this line of methodology, records of Income Tax returns from the tax authorities<sup>1</sup> have been used for estimating wage on the one hand, that is, on the other hand, records from the general file of affiliations and registrations and the file of contribution bases of the Social Security General Treasury.

The process followed leading to estimating and encoding of salaries is described below.

---

## 5 Estimation phases

---

### 5.1 OBTAINING THE PERSONAL IDENTIFICATION NUMBER

The survey has access to the following personal details of wage-earners, which are gathered in the interview: name, surname(s), sex, place and date of birth and postal address of the family dwelling. Using this information, we log into the central file of the Municipal Register and obtain the personal identification number of the wage-earner, which is necessary for creating the successive links to administrative sources. Specific “record linkage” techniques, which are both deterministic and probabilistic are used for searching for personal identification numbers.

---

### 5.2 THE LINK TO THE RECORD OF REGISTRATION IN THE SOCIAL SECURITY SYSTEM

The link to the file of workers registered as working makes it possible to detect the registration corresponding to the main job in the survey reference week and obtain two-fold information:

- The identification number of the main employer in the reference week with which to carry on making the following crosses with the tax authorities and other administrative registers.
- The number of days worked over the whole reference year in order to be able to determine the monthly wage based on a an annual total that existed in the registers or else the number of days worked over the month containing the reference week if the total existing in the register is monthly.

Therefore, firstly registrations corresponding to special contribution agreements (where contributions are paid but no work is carried out) are discarded.

Then the registration corresponding to the reference week is allocated; where there was more than one registration for the same worker in the week, the one whose characteristics of the main job -province and activity of the establishment, type of working day and duration of the contract, time worked, etc. is allocated - they are similar to statements in the EAPS questionnaire. If there continue to occur

---

<sup>1</sup> The State Tax Administration Agency, the Tax Administration of Navarre and from the month of reference 2012 the Tax Administration of Bizkaia. No information has been available from the rest of the Tax Administrations of Basque Country.

registrations with the same characteristics in the reference week, the contract is taken that has the greatest part-time coefficient in the registration with Social Security and lastly, if they continue to coincide, the registration with the longest duration that year.

Once the registration corresponding to the main job in the reference week has been established, we obtain the number of days' registration in the year and other types of circumstances that may affect the estimate of the monthly wage based on the annual total, such as the type of working day or the contribution group in the reference week as compared with those outside said week.

Likewise, for public sector wage-earners not registered with the Social Security System but in their own liabilities, in view of the likely stability of the work relationship, it will be possible to estimate the number of days' registration in the year by means of information obtained from the remainder of the survey, although the employer identification number will not be revealed, whereby it will be taken as unique or at least the main one with regard to income. This hypothesis is validated after crossing with the tax authorities.

---

### 5.3 THE LINK TO THE RECORD OF ANNUAL WAGE FOR THE TAX AUTHORITIES

With the "personal identification number of the wage-earner", other personal data (name, surnames and birth date) and, if available, the "tax identification code of the employer" a link is established to the annual income tax returns in order to obtain the "full annual performance" contained in keys A (Employees working for others in general), L-14 (Peace mission extraordinary gratuities), L-15 (Performance from jobs abroad) and from the month of reference 2012, the key G (Performance of professional and business activities) in the cases of wage-earners registered in especial Social Security schemes without any employer assigned.

Once the data has been linked and validated, an initial estimate can be made of the monthly wage by dividing full annual performance by 12 and then multiplying it by the quotient between the number of days in the year and the number of days in the year registered in Social Security with the main employer in the reference week.

A series of limitations may be noted in the preceding wage estimate method:

- We may have some extra-wage components (compensation for dismissal other than that legally established, delays, etc.), included within full work performance for the reference year.
- This entails an estimate regarding wage over the whole reference year and not over the month of the reference week, there existing the possibility of the working conditions changing during the year in the same company (type of working day, employment, etc.) which may have a bearing on wage in other months of the year.

- Tax information is included for the Tax Administrations of the Basque Country and Navarre with the sole exception of Álava (this information is expected to be available from the 2016 reference year).

---

#### 5.4 THE LINK TO THE SOCIAL SECURITY CONTRIBUTION BASES

Having highlighted the preceding limitations, we next calculate salaries taking information from Social Security, in order to improve both coverage and possibilities for filtering when estimating the monthly wage of the main job.

Since the contribution bases are registered for each calendar month of the reference year, the wage estimate may be obtained by means of the sum of contribution bases in the reference year divided by 12 and multiplied by the quotient between the number of days in the actual year registered with Social Security with the main employer in the reference week, as in the previous case.

It is also possible to obtain the estimate by means of the contribution base for the month corresponding to the reference week, in which case said base should be multiplied by the quotient between the number of days of the month and the number of days registered in the reference month, as shown in the registrations file or in the contribution file of the Social Security General Treasury.

The following may be highlighted as limitations in the estimate by means of Social Security contribution bases:

- Contribution bases have upper and lower limits which hinder wage estimation, particularly in the case of the upper limit
- It is not applicable for wage-earners in systems outside the Social Security, for example, those belonging to civil servants' mutualism.
- The contribution base for common contingencies does not include the part of the wage corresponding to overtime, therefore the contribution base for accidents at work and professional diseases is used, which does incorporate them, wherever possible.

---

#### 5.5 INTEGRATION, FILTERING AND IMPUTATION

As has been highlighted, in some cases it is possible to estimate the wage by means of several methods, in view of the information available in the administrative registers, which enriches the possibilities for filtering. In the unlikely event of existing discrepancies between the different methods, it is necessary to firstly determine which is the best wage estimate between all those available for subsequently validating it.

Therefore, the salaries finally estimated are obtained by combining all sources used and do not correspond exactly with the information received from any of them.



Lastly, for those wage-earners for whom it has not been possible to establish their wage by means of administrative registers, or for those whose estimate has not been considered sufficiently viable by any of the methods highlighted, this is imputed, using distribution of salaries by type of working day and by occupation standardised to three digits.

---

## 5.6 ENCODING

As a final step, salaries are ordered and encoded in deciles of "01" to "10", with decile "01" corresponding to the 10 per cent group of wage-earners receiving the lowest wage, and decile "10" to the 10 per cent group of wage-earners receiving the highest wage.

---

## 6 Dissemination of results

As is usual in dissemination of INE statistical results, tables and micro-data files are provided. The period of years published includes since 2006 for a better temporal follow-up. Although the European regulation establishing the legal obligation to introduce the variable in the EAPS sets 2009 as the first reference year.

The tables show the distribution of wage-earners according to the decile in which they are found, crossed with their demographic characteristics (sex, age, nationality, residence, qualification attained, sector of study, etc.) and the employment circumstances of the main job (occupation, type of working day, establishment activity, duration of the contract, etc.) highlighted in the EAPS questionnaire.

The tables also publish average wages for each subgroup resulting from the cross-checking of employment circumstances with demographic characteristics.

---

## 7 References

1. European Union Official Journal (2007). Regulation (EC) no. 1372/2007 of the European Parliament and Council, of 23 October 2007, modifying Regulation (EC) no. 577/98 of the Council, relating to the organisation of a sampling survey regarding the economically active population in the Community.
2. García Martínez, M.A. and Orche Galindo, E.J. (2011). "Integrating administrative data into the LFS data collection. The Spanish experience obtaining the variable INDECIL from administrative sources". Workshop on LFS Methodology, Wiesbaden, Germany, May 2011.
3. National Statistics Institute. Spain (1990). Economically Active Population Survey. Pilot Survey on Earnings and Underemployment.

4. National Statistics Institute. Spain (2008). Economically Active Population Survey. Methodology 2005. Description of the survey, definitions and instructions for completion of the questionnaire.
5. National Statistics Institute. Spain (2008). Economically Active Population Survey. Methodology 2005. Subsample variables.
6. Orche Galindo, E.J. and Bueno Maroto, H. (2011) "Obtaining Statistical Information in Sampling Surveys from Administrative Sources: Case Study of Spanish LFS 'Wages from the Main Job'". ESSnet Data Integration Workshop, Madrid, Spain, November 2011.  
[http://www.ine.es/e/essnetdi\\_ws2011/ppts/Orche\\_Bueno.pdf](http://www.ine.es/e/essnetdi_ws2011/ppts/Orche_Bueno.pdf)
7. Ormerod C and Ritchie F (2006). "Measuring low pay: Focus points and rounding". Office for National Statistics.
8. Ormerod C and Ritchie F (2007). "Linking ASHE and LFS: can the main earnings sources be reconciled?". Economic & Labour Market Review. Vol 1 No 3. Marzo 2007.
9. Tonder J.K (Coordinator) - UNECE (2007): "Register-based statistics in the Nordic countries. Review of best practices with focus on population and social statistics".
10. Wallgren A. – Wallgren B. (2007). Statistics Sweden. Register-based Statistics. Administrative Data for Statistical Purposes. Ed. John Wiley & Sons, Ltd.