

El Salario del Empleo Principal en la Encuesta de Población Activa

**Explotación estadística de los
Registros de la Seguridad Social y
de las Agencias Tributarias**

Índice

1	Objetivo	3
2	Ámbito	4
2.1	Ámbito temporal	4
2.2	Ámbito poblacional	4
3	Concepto estadístico de salario	5
4	Recogida de la información	5
5	Fases en la estimación	6
5.1	La obtención del Número de Identificación Personal	6
5.2	El enlace con el Registro de Altas en la Seguridad Social	6
5.3	El enlace con las Declaraciones Salariales Anuales para el IRPF	7
5.4	El enlace con las Bases de Cotización en la Seguridad Social	8
5.5	Integración, depuración e imputación	9
5.6	Codificación	9
6	Difusión de resultados	9
7	Referencias	10

1 Objetivo

El Reglamento (CE) nº 1372 / 2007 del Parlamento Europeo y del Consejo establece la obligatoriedad de incluir la variable “salario del empleo principal” en la Encuesta de Fuerza de Trabajo de la Unión Europea (cuyo equivalente en España es la Encuesta de Población Activa), dado que considera que “la clasificación de los sueldos por características socioeconómicas, así como en relación con las diferentes formas de empleo remunerado son esenciales para analizar y comprender el mercado de trabajo y los cambios que se producen en la estructura de la población activa”.

De esta forma se mejora el potencial analítico de la encuesta, introduciendo como variable de clasificación el nivel de remuneración en el análisis de las características del empleo principal para el colectivo de trabajadores por cuenta ajena o asalariados.

El reglamento sólo exige que la variable “salario del empleo principal” esté codificada en deciles y permite que la captura de la información se efectúe mediante entrevista o bien utilizando registros administrativos. En la Encuesta de Población Activa, tras realizar diversas pruebas cualitativas en el año 2003, se ha optado por la segunda opción.

Las principales ventajas del uso de registros administrativos para la estimación de la variable son, por un lado, que no se incrementa la carga de respuesta a los informantes y por otro, que no se ve afectada la encuesta por una menor tasa de respuesta en su conjunto.

El principal inconveniente es que se necesita más tiempo para la captura del dato, dado que se depende de cuándo estén disponibles los registros administrativos para su enlace. Otro inconveniente es la posible variación de la normativa de dichos registros en el tiempo.

La solución aparentemente más sencilla de incluir la variable en el cuestionario de la EPA, plantea además problemas de validez. En efecto, en caso de obtenerse la información directamente de los informantes, Ormeron y Ritchie (2006) han demostrado un efecto de redondeo significativo, dado que el asalariado o asalariada es más probable que esté fuera del hogar (en su empleo) cuando el encuestador llega o telefona al mismo y se obtenga la respuesta sobre el salario mediante entrevista a otro miembro de la familia.

Y por otro lado, los mismos autores piensan (2007) que el uso de registros administrativos permite que la información recogida sea más fidedigna que la que se obtendría directamente del informante del hogar, dado que el método se basa en evidencias documentadas.

Esta última hipótesis se vio confirmada en una encuesta piloto que se realizó en el INE en 1990 (Encuesta Piloto sobre Ganancias y Subempleo).

2 Ámbito

2.1 ÁMBITO TEMPORAL

La variable *salario del empleo principal* se integra dentro de las variables adicionales estructurales o de submuestra que proporcionan información en media anual sobre un año de referencia, mediante la agrupación de las submuestras correspondientes a las aproximadamente 40.000 familias que finalizan su colaboración en la encuesta en cada uno de los cuatro trimestres de recogida EPA que componen dicho año.

2.2 ÁMBITO POBLACIONAL

La variable se estima para todas las personas con empleo por cuenta ajena o asalariadas conforme a la definición determinada en el apartado 3.3.A) de la metodología de la encuesta.

(ver en <http://www.ine.es/daco/daco43/resumetepa.pdf>).

A continuación se reproduce la definición que figura en la citada metodología:

“Se compone de todas las personas que entren en las siguientes categorías:

- **Trabajando:** personas que durante la semana de referencia (la anterior a aquella en que se realiza la entrevista) hayan trabajado, incluso de forma esporádica u ocasional, al menos una hora a cambio de un sueldo, salario u otra forma de retribución conexas, en metálico o en especie.
- **Con empleo, pero sin trabajar:** personas que, teniendo un empleo, estén ausentes del mismo durante la semana de referencia y mantengan un estrecho vínculo con él.

Las personas ausentes de su empleo por causa de enfermedad o accidente, vacaciones o licencia por maternidad se consideran como personas con empleo.

Los ausentes por tener contratos fijos discontinuos, por ser trabajadores estacionales o por estar a la espera de incorporarse a un nuevo empleo se consideran sin empleo.

Los ausentes por otros motivos distintos de los anteriores, se consideran asalariados si el empleador les paga al menos el 50 por ciento de su sueldo o si van a reincorporarse a su empleo en los próximos tres meses.

Los aprendices que hayan recibido una retribución en metálico o en especie y los estudiantes que hayan trabajado a cambio de una remuneración a tiempo completo o parcial se consideran como personas asalariadas y se clasifican como trabajando o sin trabajar sobre la misma base que las demás personas con empleo por cuenta ajena.

Los miembros de las fuerzas armadas figuran también entre la población asalariada.

No se considerarán asalariadas las siguientes personas:

- *Empresarios y trabajadores independientes, aunque estuvieran contratados como asalariados en su propia empresa.*
- *Trabajadores familiares no remunerados (ayudas familiares) y los miembros de cooperativas que trabajan en las mismas.*
- *Las personas ausentes de su empleo, con licencia o suspendidas que estén débilmente vinculadas a su empleo.*
- *Los trabajadores estacionales, ocasionales o discontinuos por cuenta ajena en la época de menor actividad que no hayan trabajado en la semana de referencia.”*

3 Concepto estadístico de salario

Se ha acordado a nivel europeo el concepto de “salario mensual del empleo principal”, como una definición de referencia que permita la comparación entre los diversos países y sea a la vez flexible para que se pueda estimar por diferentes métodos. En el caso de España, esta definición se refiere al salario mensual del empleo principal realizado por la persona en la semana de referencia de la encuesta, incluidas las retenciones a cuenta de impuestos sobre la renta del empleo y las cotizaciones obligatorias del trabajador a sistemas de seguridad social (salario mensual bruto, por tanto).

Se han añadido criterios convencionales en relación con casos especiales. Así, las pagas extras y otros complementos anuales son considerados en la definición, dividiendo su cantidad por 12. También se incluyen los pagos por horas extras regulares, propinas y comisiones.

Además, si es un empleo eventual de duración inferior a un mes, el salario será una estimación del que se obtendría para una duración de todo el mes por un empleo similar (la persona, al menos en principio, podría realizar otros empleos de esa índole en el mismo mes). Pero en cambio, si es un empleo a tiempo parcial no debe transformarse al equivalente a tiempo completo (se trata del empleo principal y por tanto es el que proporciona los ingresos más significativos).

4 Recogida de la información

Como se indicó en el punto 1, en lugar de preguntar directamente a las personas de la muestra se ha preferido obtener la información mediante el uso de registros administrativos, pero dado que no existe ninguno que cumpla la definición establecida estrictamente, debido a tener diferentes metodologías y coberturas limitadas, la solución adoptada ha sido obtener la información de varios registros

y, sobre la base de algunos datos adicionales, intentar encontrar los mejores estimadores de la variable objetivo.

Por tanto, se trata de lo que se ha denominado en la literatura estadística una “variable derivada”, dado que la necesidad de información no puede ser cubierta inmediatamente por referencia directa a la información disponible. Tal variable se obtiene mediante el encadenamiento de un conjunto de diferentes fuentes que proporcionan la información requerida, mediante combinación de datos.

Siguiendo esta línea metodológica, se han utilizado para la estimación del salario, por un lado, los registros de las declaraciones anuales de retenciones e ingresos a cuenta del Impuesto de la Renta de las Personas Físicas (IRPF) de las administraciones tributarias¹ y por otro, los registros del fichero general de afiliaciones y altas y el fichero de bases de cotización de la Tesorería General de la Seguridad Social (TGSS).

El proceso seguido hasta la estimación y codificación de los salarios se describe a continuación.

5 Fases en la estimación

5.1 LA OBTENCIÓN DEL NÚMERO DE IDENTIFICACIÓN PERSONAL

La encuesta cuenta con los siguientes datos personales de los asalariados, que se recogen en la entrevista: nombre, apellidos, sexo, lugar y fecha de nacimiento y dirección postal de la vivienda familiar. Con esta información, se accede al fichero central del Padrón Municipal y se obtiene el número de identificación personal del asalariado, necesario para efectuar los enlaces sucesivos con las fuentes administrativas. Se utilizan, para la búsqueda de los números de identificación personal, técnicas específicas de unión de registros o “record linkage” tanto determinísticas como probabilísticas.

5.2 EL ENLACE CON EL REGISTRO DE ALTAS EN LA SEGURIDAD SOCIAL

El enlace con el fichero de trabajadores en alta laboral permite detectar el alta correspondiente al empleo principal en la semana de referencia de la encuesta y obtener una doble información:

- El número de identificación del empleador principal en la semana de referencia con el que seguir realizando los siguientes cruces con las agencias tributarias y otros registros administrativos.
- El número de días trabajados sobre todo el año de referencia para poder determinar el salario mensual en base a un total anual que existiera en los

¹ Agencia Estatal de Administración Tributaria (AEAT), Hacienda Foral de Navarra y desde el año de referencia 2012, la Hacienda Foral de Bizkaia. No se ha podido disponer de información del resto de Haciendas Forales Vascas.

registros o bien el número de días trabajados sobre el mes en el que se integra la semana de referencia si el total existente en el registro es mensual.

Para ello, primeramente se descartan las altas que corresponden a convenios especiales de cotización (donde se cotiza pero no se trabaja). Luego, se asigna el alta que corresponda a la semana de referencia; si hubiera varias altas para un mismo trabajador en la semana, se asigna aquella cuyas características del empleo principal -provincia y actividad del establecimiento, tipo de jornada y duración del contrato, antigüedad en el mismo, etc.- se asemejen a las declaradas en el cuestionario EPA. Si aun así se siguen obteniendo varias altas con las mismas características en la semana de referencia, se toma aquel contrato con mayor coeficiente de tiempo parcial en el alta de la Seguridad Social y finalmente, si siguen coincidiendo, el alta de mayor duración en el año.

Una vez se tiene establecido el alta correspondiente al empleo principal en la semana de referencia, se obtiene el número de días de alta en el año y otro tipo de circunstancias que pueden afectar a la estimación del salario mensual en base al total anual, como puedan ser el tipo de jornada o el grupo de cotización en la semana de referencia frente a los de fuera de dicha semana. Asimismo, en el caso de los regímenes generales de la Seguridad Social se consigue el código de identificación fiscal del empleador al que pertenece dicha alta.

Para los asalariados del sector público que no estén dados de alta en el sistema de la Seguridad Social sino en sus propias mutualidades, dada la probable estabilidad de la relación laboral, será posible realizar una estimación del número de días en alta en el año mediante la información obtenida del resto de la encuesta, aunque se desconocerá el número de identificación del empleador, por lo que se supondrá que es único o por lo menos el principal en cuanto a ingresos. Esta hipótesis se valida después de realizar el cruce con las agencias tributarias.

5.3 EL ENLACE CON LAS DECLARACIONES SALARIALES ANUALES PARA EL IRPF

Con el "número de Identificación personal del asalariado", otros datos personales (nombre, apellidos y fecha de nacimiento) y en el caso en que también se conozca, el "código de identificación fiscal del empleador" se enlaza con el registro de las declaraciones anuales de ingresos y retenciones e ingresos a cuenta del IRPF de las agencias tributarias para obtener el "rendimiento íntegro anual" contenido en las claves A (Empleados por cuenta ajena en general), L-14 (Gratificaciones extraordinarias misiones de paz) , L-15 (Rendimientos trabajos en extranjero) y a partir del año de referencia 2012, la clave G (Rendimientos de actividades profesionales y empresariales) para el caso de asalariados dados de alta en regímenes especiales de la seguridad social sin ningún empleador asignado.

Una vez realizado el enlace y obtenido y validado el dato, se puede realizar una primera estimación del salario mensual dividiendo el rendimiento íntegro anual entre doce y multiplicándolo después por el cociente entre el número de días del año y el número de días en el mismo año dado de alta en la Seguridad Social con el empleador principal en la semana de referencia.

Se pueden apuntar en el método de estimación del salario anterior una serie de limitaciones:

- Podríamos tener algunos componentes extra-salariales (indemnizaciones por despido fuera del legalmente establecido, atrasos, etc.), incluidos dentro del rendimiento íntegro del empleo del año de referencia.
- Se trata de una estimación sobre el salario sobre todo el año de referencia y no sobre el mes de la semana de referencia, pudiendo haber cambiado las condiciones del empleo durante el año en la misma empresa (tipo de jornada, ocupación, etc.) que podrían alterar el salario en otros meses del año.
- Se incluye la información tributaria de las Haciendas Forales del País Vasco y Navarra, con la única excepción de Álava (se prevee disponer de esta información a partir del año de referencia 2016).

5.4 EL ENLACE CON LAS BASES DE COTIZACIÓN EN LA SEGURIDAD SOCIAL

Señaladas las anteriores limitaciones se procede a calcular los salarios a partir de la información de la Seguridad Social, para mejorar tanto la cobertura como las posibilidades de depuración en la estimación del salario mensual del empleo principal.

Dado que las bases de cotización se registran para cada uno de los meses naturales del año de referencia se puede obtener la estimación del salario mediante la suma total de bases de cotización en el año de referencia dividida por doce y multiplicada por el cociente entre el número de días de año y el número de días en el mismo año dado de alta en la Seguridad Social con el empleador principal en la semana de referencia, como en el caso anterior.

También se puede obtener la estimación mediante la base de cotización del mes correspondiente a la semana de referencia, en cuyo caso se ha de multiplicar dicha base por el cociente entre el número de días del mes y el número de días dados de alta en el mes de referencia, según conste en el fichero de altas o en el fichero de cotización de la Tesorería General de la Seguridad Social.

Se pueden señalar como limitaciones en la estimación mediante las bases de cotización en la Seguridad Social las siguientes:

- Las bases de cotización tienen topes tanto máximos como mínimos que dificultan la estimación del salario, especialmente en el caso del tope máximo.
- No es aplicable para los asalariados en sistemas fuera de la Seguridad Social, por ejemplo, los pertenecientes a mutualidades de los funcionarios.
- La base de cotización por contingencias comunes no incluye la parte del salario correspondiente a las horas extraordinarias, por lo que se utiliza la base de cotización para accidentes de trabajo y enfermedades profesionales, que sí las incorpora, siempre que sea posible.

5.5 INTEGRACIÓN, DEPURACIÓN E IMPUTACIÓN

Como se ha señalado, en algunos casos existe la posibilidad de estimar el salario mediante varios métodos, a la vista de la información disponible en los registros administrativos, lo que enriquece las posibilidades de depuración. En el caso poco frecuente de que existan discrepancias entre los diversos métodos, se ha de determinar primeramente cuál es la mejor estimación del salario entre todas las disponibles para posteriormente validar la misma.

Por ello, los salarios finalmente estimados se obtienen de la combinación de todas las fuentes utilizadas y no se corresponden exactamente con la información recibida de ninguna de ellas.

Finalmente, para aquellos asalariados para quienes no ha sido posible establecer mediante los registros administrativos su salario o para aquellos cuya estimación no se ha considerado suficientemente fiable por ninguno de los métodos señalados se realiza una imputación del mismo, utilizando para ello la distribución de los salarios por tipo de jornada y por ocupación normalizada a tres dígitos.

5.6 CODIFICACIÓN

Como paso final, los salarios se ordenan y se codifican en deciles del "01" al "10", correspondiendo el decil "01" al grupo del 10 por ciento de los asalariados que perciben el salario más bajo y el decil "10" al grupo del 10 por ciento de los asalariados que perciben el salario más alto.

6 Difusión de resultados

Como es habitual en la difusión de resultados de las estadísticas del INE, se proporcionan tabulaciones y ficheros de microdatos. El periodo de años publicado incluye desde 2006 para un mejor seguimiento temporal. Aunque el reglamento europeo que establece la obligatoriedad de introducir la variable en la EPA en 2009 como primer año de referencia.

En las tablas aparece la distribución de los asalariados según el decil en que se ubiquen, cruzado con las características demográficas que poseen (sexo, edad, nacionalidad, residencia, titulación alcanzada, sector de estudios, etc.) y las circunstancias laborales del empleo principal (ocupación, tipo de jornada, actividad del establecimiento, duración del contrato, etc.) señaladas en el cuestionario EPA.

Asimismo, en las tablas se publican los salarios promedio para cada subgrupo resultante del cruce de circunstancias laborales con características demográficas.

7 Referencias

1. Diario Oficial de la Unión Europea (2007). Reglamento (CE) n° 1372/2007 del Parlamento Europeo y del Consejo, de 23 de octubre de 2007, por el que se modifica el Reglamento (CE) n° 577/98 del Consejo, relativo a la organización de una encuesta muestral sobre la población activa en la Comunidad.
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. Instituto Nacional de Estadística. España (1990). Encuesta de Población Activa. Encuesta Piloto sobre Ganancias y Subempleo.
4. Instituto Nacional de Estadística. España (2008). Encuesta de Población Activa. Metodología 2005. Descripción de la encuesta, definiciones e instrucciones para la cumplimentación del cuestionario.
5. Instituto Nacional de Estadística. España (2008). Encuesta de Población Activa. Metodología 2005. Variables de la submuestra.
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
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.