



Documentos de Trabajo

2/2014

**Comparación de los ingresos del trabajo entre
la Encuesta de Condiciones de Vida y las
fuentes administrativas**

Pilar Vega y José María Méndez

El Instituto Nacional de Estadística no se identifica necesariamente con las opiniones
formuladas por los autores en este informe

Primera versión: Marzo 2014

Versión actual: Marzo 2014

Comparación de los ingresos del trabajo entre la Encuesta de Condiciones de Vida y las fuentes administrativas

Resumen

En este documento de trabajo se hace un estudio comparativo entre los datos de las rentas del trabajo recogidos mediante entrevista personal en la Encuesta de Condiciones de Vida (ECV) de 2011 y los datos provenientes de fuentes administrativas.

Tanto en los ingresos del trabajo por cuenta ajena como por cuenta propia se hace un análisis de los perceptores de ingresos, así como de las diferencias existentes entre los importes de los ingresos que el informante proporciona en la ECV y los que proporcionan las Fuentes Tributarias.

Palabras clave

Encuesta de Condiciones de Vida, ingresos del trabajo, fuentes administrativas.

Autores y Afiliaciones

Pilar Vega

Secretaría de Estado de Comercio. Ministerio de Economía y Competitividad

José María Méndez

Instituto Nacional de Estadística

Comparación de los ingresos del trabajo entre la Encuesta de Condiciones de Vida y las fuentes administrativas

Pilar Vega Vicente
José María Méndez Martín

Introducción

La utilización de registros administrativos en las operaciones estadísticas está teniendo una gran importancia en los últimos años. Desde la perspectiva de la producción estadística, el acceso a registros administrativos supone una buena oportunidad que redundará en la mejora de la calidad de los correspondientes datos y en la eficiencia en el método de recogida.

Además, el cruce a nivel microdato de información recogida mediante una encuesta y de los datos disponibles de ficheros administrativos supone también una oportunidad para comparar ambas fuentes, pudiendo estudiar y analizar las discrepancias existentes. El interés se tiene tanto desde la perspectiva de la encuesta, analizando los errores de medición, como desde la perspectiva de la Administración que produce el fichero administrativo.

En los últimos años se han realizado trabajos relacionados con este tema. Cabe destacar un taller organizado por Eurostat en diciembre de 2012 en el contexto de EU-SILC que ha dado lugar a una publicación (Eurostat [1]) en la que se recogen las experiencias de distintos países. También se puede señalar el trabajo llevado a cabo por Dauphin [2] en Francia, que utiliza datos de ingresos de la SRCV (Statistiques sur les Ressources et les Conditions de Vie) de 2005 junto con datos fiscales de la Dirección General de Finanzas Públicas y el de Donatiello, Betti y Consolini en Italia [3].

Fuera del ámbito europeo es reseñable el realizado por Cristia y Schwabish [4] que usa datos de ingresos provenientes de la encuesta Survey of Income and Program Participation (SIPP) de 1996 junto con datos de ingresos tomados de registros de la Seguridad Social en Estados Unidos (Social Security Administration's Detailed Earnings Record, DER). En este estudio se concluye que los encuestados en la SIPP infravaloran de media sus ingresos en 2800 \$ (a precios constantes de 2005). Este hecho se explica en buena medida porque los perceptores de rentas altas declaran en la encuesta ganar mucho menos. Asimismo se concluye que factores positivamente asociados con los ingresos están negativamente correlacionados con los errores de medición (diferencia entre el valor de la encuesta y el del fichero administrativo).

Otros trabajos relacionados con el estudio de estos errores de medición los encontramos en Abowd y Stinson [5] y en Gottschalk y Huynh [6], que cruzan también los datos de trabajadores procedentes de SIPP con los registros de DER.

En el caso de España se dispone de una encuesta anual de ingresos, la Encuesta de Condiciones de Vida (ECV), que es la versión española de la encuesta europea EU-SILC. Su objetivo general es la producción sistemática de estadísticas sobre la renta de los hogares y las condiciones de vida.

En relación con esta fuente estadística y su combinación con fuentes administrativas, cabe destacar el Documento de Trabajo disponible en la web del INE "Linking data from administrative records and the Living Conditions

Survey”, en el que se hace un análisis comparativo descriptivo así como se valora el impacto que tendría el uso de ficheros administrativos en algunos indicadores de la ECV.

En este artículo se hará un estudio comparativo entre la información relativa a la renta de la ECV de 2011 y los datos administrativos, en la línea de los trabajos citados anteriormente. Los ingresos están referidos al año 2010.

El enlace a nivel microdato de las personas pertenecientes a la muestra con los datos disponibles en la Agencia Tributaria o en la Secretaría de Estado de la Seguridad Social permite comparar ambas fuentes en la mayoría de los componentes de la renta. En este estudio se abordarán dos componentes básicos de la renta de los hogares como son los ingresos del trabajo por cuenta ajena y por cuenta propia. La fuente administrativa de los ingresos del trabajo por cuenta ajena ha sido el modelo 190 (un fichero disponible en la AEAT que contiene información individual de los rendimientos del trabajo), mientras que en el caso de los ingresos del trabajo por cuenta propia dicha fuente ha sido el modelo 100 (declaraciones del IRPF).

Para poder llevar a cabo el enlace entre ficheros es necesario disponer de una variable común. En el caso de la ECV el campo que permite el enlace es el NIF.

Agradecemos la valiosa colaboración prestada por la AEAT al proporcionarnos la información necesaria para este estudio. Asimismo queremos expresar nuestro agradecimiento a las distintas unidades del INE por el apoyo recibido en este proyecto.

1 Marco general del estudio

En el análisis se ha tomado de la encuesta el marco común de las personas adultas con NIF disponible y que residen en territorio español excluyendo al País Vasco, dado que no se dispone de sus datos tributarios.

Las variables o factores de clasificación que se tienen en cuenta en el análisis son las siguientes:

- Edad: puesto que estamos analizando las rentas del trabajo, consideramos a los mayores de 16 años. La edad se clasifica en cuatro categorías: ‘16-24’, ‘25-44’, ‘45-64’ y ‘mayores de 65’. Categoría base¹: ‘25,44’.
- Nivel de formación: el nivel de formación más alto alcanzado se clasifica en cuatro grupos: ‘Educación primaria o inferior’, ‘Educación secundaria primera etapa’, ‘Educación secundaria segunda etapa’ y ‘Educación superior’. Categoría base: ‘Educación secundaria primera etapa’

¹ Para llevar a cabo los análisis de regresión con variables cualitativas se ha de fijar una categoría como base o referencia con el fin de comparar las estimaciones realizadas con respecto a esa categoría. Para seleccionar las categorías base nos hemos fijado en aquellas que cuentan con un tamaño muestral relativamente elevado.

- Actividad económica: catalogamos la actividad económica de la unidad local del establecimiento donde se realiza el trabajo principal en las categorías siguientes: 'Agricultura, ganadería, caza y silvicultura', 'Pesca', 'Industrias extractivas', 'Industria manufacturera', 'Producción, distribución de energía eléctrica, etc', 'Construcción', 'Comercio, reparación de vehículos, etc', 'Transporte, almacenamiento, y comunicaciones', 'Hostelería', 'Información y comunicaciones', 'Actividades financieras y seguros', 'Actividades inmobiliarias', 'Actividades profesionales, científicas', 'Actividades administrativas y servicios auxiliares', 'Administración Pública, Defensa y Seguridad Social', 'Educación', 'Actividades sanitarias, servicios sociales', 'Actividades artísticas, recreativas', 'Otros servicios', 'Hogares que emplean personal doméstico' y 'Organismos extraterritoriales'. Categoría base: 'Industria manufacturera'.
- Tramos ingresos del trabajo AEAT: se construirán los siguientes tramos de ingresos del trabajo según la Agencia Tributaria: ' - 9.000', '9.000 – 20.000', '20.000 – 30.000', '30.000 - '. Categoría base: '9.000-20.000'.
- Sexo: se consideran dos categorías 'Hombres' y 'Mujeres'. Categoría base: 'Hombres'.
- País de nacimiento: en función del país de nacimiento clasificamos a los individuos en los siguientes grupos: 'España', 'EU-25 sin España' y 'Resto de países'. Categoría base: 'España'.
- Trabajo secundario: las personas se dividen en dos clases, 'con trabajo secundario' y 'sin trabajo secundario'. Categoría base: 'sin trabajo secundario'.
- Tamaño de la empresa: según el número de individuos que trabajan en la unidad local, se distinguen dos tipos de empresas: 'con menos de 10 trabajadores' y 'con más de 10 trabajadores'. Categoría base: 'con más de 10 trabajadores'.
- Comunidad Autónoma: contamos con 17 categorías, las comunidades autónomas españolas (se excluye País Vasco al no disponer de datos de la AEAT) y otra categoría que incluye 'Ceuta y Melilla'. Categoría base: 'Castilla y León'.

En relación con el desglose cuenta ajena/cuenta propia, la correspondencia en ambas fuentes no es total. Estas discrepancias se explican parcialmente porque hay empresarios en la ECV que forman sociedades y figuran como asalariados en la AEAT. Por otra parte puede ocurrir que autónomos en la AEAT, que trabajan con un mismo cliente, se vean como asalariados en la ECV.

ECV 2011. Distribución de personas según percepción de ingresos por cuenta ajena o cuenta propia (ECV y AEAT). Porcentajes (datos muestrales)

	AEAT				
	T1 Solo con ingresos cuenta ajena	T2 Solo con ingresos cuenta propia	T3 Con ingresos cuenta ajena y propia	T4 Sin ingresos trabajo	
TOTAL					
ECV	100,0	42,4	6,9	3,1	47,5
E1. Solo con ingresos cuenta ajena	40,7	34,9	0,4	1,9	3,5
E2. Solo con ingresos cuenta propia	5,7	1,1	3,5	0,6	0,6
E3. Con ingresos cuenta ajena y propia	0,6	0,2	0,1	0,3	0,0
E4. Sin ingresos trabajo	53,1	6,3	3	0,4	43,4

Para simplificar el estudio no utilizaremos factores de ponderación en los resultados y nos centraremos en aquellos individuos que perciben exclusivamente una determinada fuente en la encuesta, de manera que el análisis comparativo se hará por separado para los ingresos del trabajo por cuenta ajena y por cuenta propia.

En este estudio los importes de los ingresos del trabajo por cuenta ajena son importes brutos. Por su parte en el caso de los ingresos del trabajo por cuenta propia se han tomado los importes netos. Esta elección se debe a la disponibilidad y calidad de los datos.

Para cada componente de la renta se procederá a realizar los siguientes análisis:

1. **Análisis de la percepción de ingresos.** Se estudian los factores que influyen en que un perceptor de renta según la encuesta, no lo sea según el fichero administrativo.
2. **Análisis de los importes.** Se estudia en qué medida afecta una serie de variables a la diferencia existente entre los ingresos que el individuo proporciona a la ECV y los que proporciona a la AEAT.

2 Ingresos del trabajo por cuenta ajena

ANÁLISIS DE LA PERCEPCIÓN DE INGRESOS

A continuación se hace un estudio de los factores que determinarán el que una persona que se considera asalariada según la encuesta, no figure como perceptora de ingresos en los ficheros de la AEAT. La estructura que seguimos

es la siguiente, en una primera parte se describirán los datos utilizados y en una segunda, los resultados obtenidos tras realizar un estudio econométrico.

1. Análisis descriptivo de la percepción

En este apartado se abordará el colectivo formado por aquellas personas que declaran solamente ingresos del trabajo por cuenta ajena en la encuesta. Se tiene:

ECV 2011. Asalariados de la ECV, según tipo de perceptor de la AEAT. Porcentajes (datos muestrales)

Total	100,0
T1. Solo con ingresos cuenta ajena en la AEAT	85,8
T2. Solo con ingresos cuenta propia en la AEAT	0,9
T3. Con ingresos cuenta ajena y propia en la AEAT	4,6
T4. Sin ingresos trabajo en la AEAT	8,7

Se observa que casi el 9% de los asalariados de la encuesta figuran como no perceptores de rendimientos del trabajo según la Agencia Tributaria (en la encuesta de 2010 este porcentaje fue un 6,6%).

En la tabla adjunta, se presenta el porcentaje de no perceptores según la AEAT, en función de distintas variables de clasificación. Se proporciona también el tamaño muestral para valorar la precisión de dichos porcentajes.

ECV 2011. Asalariados de la ECV. Porcentaje de no perceptores según la AEAT por variables de clasificación (datos muestrales)

	Tamaño muestral	Porcentaje no perceptores
Total	11.086	8,7
Edad		
16 - 24	874	13,6
25 - 44	5.781	8,1
45 - 64	4.343	7,9
65 -	88	29,5
Nivel de formación		
Educación primaria o inferior	1.554	12,3
Educación secundaria primera etapa	2.717	8,6
Educación secundaria segunda etapa	2.606	9,2
Educación superior	3.991	5,7
No consta	218	32,6
Actividad económica		
Agricultura, ganadería, caza y silv.	230	4,8
Pesca	25	0,0
Industrias extractivas	69	5,8
Industria manufacturera	1.258	5,0
Prod. dist. energía eléct.,agua,etc.	221	2,7
Construcción	598	6,5
Comercio; reparación de vehículos, etc.	952	5,4
Transporte, almacenamiento	452	5,3
Hostelería	550	10,2
Información y comunicaciones	279	9,7
Actividades financieras y seguros	289	3,8
Actividades inmobiliarias	42	9,5
Actividades profesionales, cient.	256	4,7
Act. administrativas y serv. auxil.	337	6,5
Adm. Pública, Defensa y Seg. Soc.	1.055	4,5
Educación	822	4,4
Actividades sanitarias, servicios sociales	825	5,7
Actividades artísticas, recreativas	122	7,4
Otros servicios	350	8,3
Hogares que emplean personal doméstico	202	58,9
Organismos extraterritoriales	29	10,3
No consta	2.123	16,0
Sexo		
Hombres	5.867	6,9
Mujeres	5.219	10,6
País de nacimiento		
España	9.856	6,9
EU-25 sin España	148	16,9
Resto de países	876	21,2
No consta	206	35,0
Trabajo secundario		
Con trabajo secundario	117	11,1
Sin trabajo secundario	10.964	8,6
No consta	5	0,0
Tamaño de la empresa		
Empresas con menos de 10 trabajadores	4.510	12,2
Empresas con más de 10 trabajadores	6.114	5,2
No consta	462	19,9
Comunidad Autónoma		
Andalucía	1.267	7,8
Aragón	615	7,5
Asturias (Principado de)	453	6,0
Baleares (Illes)	402	10,4
Canarias	513	11,5
Cantabria	364	10,4
Castilla y León	728	6,5
Castilla - La Mancha	575	8,5
Cataluña	1.288	9,5
Comunitat Valenciana	909	8,0
Extremadura	423	8,3
Galicia	768	9,6
Madrid (Comunidad de)	1.290	10,4
Murcia (Región de)	406	9,4
Navarra (Comunidad Foral de)	442	5,7
Rioja (La)	409	9,5
Ceuta y Melilla	234	6

A partir de esta tabla se puede deducir lo siguiente:

- Alrededor de un 9% de perceptores de salarios de la encuesta no tienen registro en la AEAT como perceptores de rentas. Se observa que este porcentaje disminuye al aumentar la edad (considerando los tramos de edad activa) y el nivel de formación y que aumenta significativamente en los sectores 'Hogares que emplean personal doméstico', 'Hostelería', 'Actividades inmobiliarias' e 'Información y comunicaciones'.
- También destaca un porcentaje alto de no perceptores según la AEAT en las personas extranjeras y los que trabajan para 'Organismos extraterritoriales'.

2. Análisis logit de la percepción

Para estudiar la influencia que tienen las variables explicativas mencionadas con anterioridad (excepto la variable ingresos) sobre la probabilidad de que una persona sea perceptor según la AEAT o no, se hace un análisis logit¹. Para ello, se toma el colectivo formado por los individuos del marco común que declaran solamente ingresos del trabajo por cuenta ajena en la encuesta y se considera la variable de análisis siguiente:

$$\text{Variable "percepción AEAT"} = \begin{cases} 1 & \text{no es perceptor según la AEAT} \\ 0 & \text{sí es perceptor según la AEAT} \end{cases}$$

Este análisis nos permitirá conocer la influencia de cada factor en la variable recién definida, eliminando el efecto de las otras variables.

En la tabla siguiente se muestran las categorías que más influyen en la probabilidad de que un individuo con ingresos del trabajo según la ECV no sea perceptor de rentas según la AEAT (el mecanismo de selección utilizado es *stepwise* con un nivel de significación $\alpha=0,005$).

¹ Análisis logit: modeliza la probabilidad asociada a un suceso (en este caso "no ser perceptor según la AEAT") en función de unas características medibles. Nos va a permitir conocer la influencia de cada variable sobre la probabilidad de no ser perceptor según la AEAT, eliminando el efecto de otras variables.

Categorías	Coeficiente¹	Error estándar	Odds-ratio²
Hogares que emplean personal doméstico	2,83	0,17	16,98
Información y comunicaciones	0,88	0,21	2,40
Resto de países	0,83	0,13	2,29
Hostelería	0,46	0,16	1,58
Empresas con menos de 10 trabajadores	0,38	0,10	1,46
Mujer	0,36	0,10	1,44
Constante	-3,36	0,08	0,03

A partir de esta tabla se puede deducir que los factores más relevantes en la probabilidad de no ser perceptor según la AEAT son los siguientes:

- ‘Hogares que emplean personal doméstico’ (variable ‘Actividad económica’).
- ‘Información y comunicaciones’ (variable ‘Actividad económica’)
- ‘Resto de países’ (variable ‘País de nacimiento’)

Es decir, tomando inicialmente todas las variables explicativas introducidas en el modelo, encontramos que estas tres categorías son las más influyentes en la probabilidad de no ser perceptor según la AEAT, en particular la categoría ‘Hogares que emplean personal doméstico’ de la variable ‘Actividad económica’.

ANÁLISIS DE LOS IMPORTES

Una vez realizado el análisis de la percepción de ingresos, estudiaremos los errores de medición, comparando los importes recogidos en la encuesta con los de la AEAT. Para ello vamos a considerar el colectivo formado por los individuos que solamente tienen ingresos del trabajo por cuenta ajena, tanto en la encuesta como en la Agencia Tributaria.

¹ Coeficiente: un coeficiente positivo (respect. negativo), estadísticamente significativo, indica que la categoría considerada aumenta (respect. disminuye) la probabilidad de que un asalariado según la ECV no sea perceptor según la AEAT.

² Odds-ratio: si el odds ratio asociado a una categoría es superior a 1, la categoría tiene un efecto positivo sobre la probabilidad de no ser perceptor suponiendo que el resto de variables del modelo permanecen constantes. Cuanto mayor sea el odds ratio, mayor será el efecto positivo de la categoría considerada. Si el odds ratio es inferior a 1, la categoría tiene un efecto negativo sobre la probabilidad de no ser perceptor

1. Análisis descriptivo de los importes

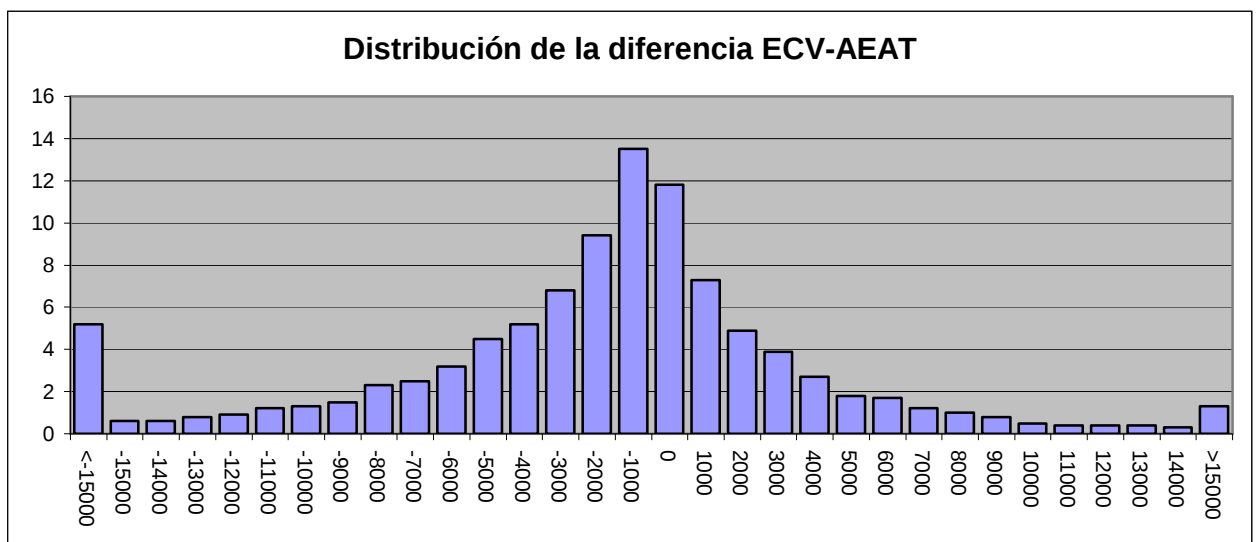
A continuación se presenta un cuadro resumen con las principales medidas descriptivas:

ECV 2011. Asalariados. Estadísticas de los importes y de las diferencias (datos muestrales)

	Media	Desviación estándar	Percentil 25	Mediana	Percentil 75
ECV	19.143	12.789	10.678	16.832	25.022
AEAT	21.214	17.188	10.434	17.724	27.906
ECV - AEAT	-2.071	10.146	-3.937	-637	1.493
ECV - AEAT	5.131	8.995	985	2.632	6.013

A partir de las mismas podemos afirmar que los salarios medios proporcionados a la Agencia Tributaria son superiores a los facilitados a la encuesta y que los datos de la Agencia Tributaria presentan una mayor dispersión. Por otra parte vemos que en los percentiles 50 y 75, el valor de la AEAT es superior al de la encuesta, en particular en el percentil 75.

Al analizar la diferencia entre el valor de la encuesta y el de la AEAT se observa una cierta asimetría hacia la izquierda.



En la tabla que vemos a continuación se hace un análisis descriptivo de la diferencia de importes según distintas variables de clasificación:

ECV 2011. Asalariados. Diferencias ECV-AEAT por variables de clasificación (datos muestrales)

	Tamaño muestral	AEAT	ECV-AEAT (Media)	ECV-AEAT (Media)
Total	9.508	21.214	-2.071	5.131
Edad				
16 - 24	748	8.535	330	2.907
25 - 44	5.032	19.156	-1.209	4610
45 - 64	3.680	26.504	-3.716	6.248
65 -	48	29.017	-3.769	8.790
Nivel de formación				
Educación primaria o inferior	1.267	14.893	-1.180	3.953
Educación secundaria primera etapa	2.339	15.290	-460	3.763
Educación secundaria segunda etapa	2.263	19.036	-1.339	4.460
Educación superior	3.514	28.904	-3.756	6.684
No consta	125	19.420	-7.146	11.125
Tramos salarios AEAT				
-9000	2.013	4.648	2.238	3.272
9000 - 20000	3.489	14.712	116	3.109
20000 - 30000	1.948	24.479	-2.112	4.076
30000 -	2058	45.352	-9.955	11.375
Actividad económica				
Agricultura, ganadería, caza y silv.	175	12.755	425	4.059
Pesca	25	19.530	-3.858	9.185
Industrias extractivas	62	23.709	-3.157	5.456
Industria manufacturera	1.132	23.059	-2.057	5.037
Prod. dist. energía eléct.,agua,etc.	202	27.971	-4.633	6.979
Construcción	531	19.925	-725	4.790
Comercio; reparación de vehículos, etc.	866	17.260	-689	4.032
Transporte, almacenamiento	404	23.124	-1.760	5.724
Hostelería	471	13.531	532	3.494
Información y comunicaciones	241	27.418	-3.739	6.582
Actividades financieras y seguros	257	41.464	-9.673	12601
Actividades inmobiliarias	36	21.743	-226	3.449
Actividades profesionales, cient.	227	26.261	-4.415	8.045
Act. administrativas y serv. auxil.	306	17.597	-1.197	3.829
Adm. Pública, Defensa y Seg. Soc.	950	29.223	-3.243	5.254
Educación	733	28.890	-2.638	4.783
Actividades sanitarias, servicios sociales	712	26.795	-3.525	6.008
Actividades artísticas, recreativas	103	19.600	-1.208	4547
Otros servicios	301	15.056	-19	3.776
Hogares que emplean personal doméstico	78	11.082	-4	2.763
Organismos extraterritoriales	26	21.178	-765	2593
No consta	1.670	11.464	-1.395	4.665
Sexo				
Hombres	5.070	24011	-2674	6026
Mujeres	4.438	18.020	-1.382	4.108
País de nacimiento				
España	8.612	21.786	-2.140	5.097
EU-25 sin España	111	23.263	-1.884	7205
Resto de países	673	14.064	-530	4.329
No consta	112	18.202	-6.200	10524
Trabajo secundario				
Con trabajo secundario	80	22350	-1476	4.806
Sin trabajo secundario	9.423	21.206	-2.076	5.135
No consta	5	19576	-2469	2.469
Tamaño de la empresa				
Empresas con menos de 10 trabajadores	3.697	14.715	-633	4.120
Empresas con más de 10 trabajadores	5.472	25.649	-2.902	5.635
No consta	339	20.513	-4.340	8.010
Comunidad Autónoma				
Andalucía	1.072	18.245	-1.636	5.083
Aragón	533	21.505	-1.811	4.041
Asturias (Principado de)	403	21.422	-1.256	4.276
Balears (Illes)	345	19.050	-755	4.629
Canarias	436	18.892	-1.933	4.604
Cantabria	317	20.881	-1.985	4.999
Castilla y León	646	21.004	-1.367	4.162
Castilla - La Mancha	467	20.753	-2.058	5.231
Cataluña	1.109	23.399	-3.166	6.577
Comunitat Valenciana	781	20.199	-1.913	4.460
Extremadura	351	17.217	-1.018	4.451
Galicia	663	19.055	-1.429	4.353
Madrid (Comunidad de)	1.118	26.858	-3.851	7.108
Murcia (Región de)	336	17.744	-1.602	4.658
Navarra (Comunidad Foral de)	375	24.611	-1.403	4.262
Rioja (La)	341	18.981	-432	3.394
Ceuta y Melilla	215	24.518	-4.943	7.642

Se puede destacar lo siguiente:

- En consonancia con el trabajo anteriormente citado de Cristia y Schwabish [1], la diferencia entre importes (ingresos ECV – ingresos AEAT) presenta valores más negativos en aquellos grupos asociados a ingresos altos, como son las personas de mayor edad (45 años en adelante), los que tienen educación superior, las personas que se dedican a “Actividades financieras y seguros”, etc.
Por su parte, la diferencia entre importes toma valores positivos en las siguientes categorías: individuos con edad comprendida entre los 16 y 24 años, primeros tramos de renta; aquellos cuya actividad económica es la ‘Agricultura, ganadería, caza y silvicultura’ y la ‘Hostelería’.
- Asimismo, el error de medición en valor absoluto también crece en los grupos con ingresos altos.

2. Análisis de regresión de los importes

Para estudiar en qué medida afectan distintas variables a la diferencia que existe entre los ingresos netos que el individuo proporciona a la ECV y los que proporciona a la AEAT, tratando de eliminar el efecto de otras variables explicativas, se va a realizar un análisis de regresión. Al igual que en el apartado anterior, vamos a considerar el colectivo formado por los individuos que solamente tienen ingresos del trabajo por cuenta ajena, tanto en la encuesta como en la AEAT.

El estudio pretende modelizar el error de medición, considerando como variables explicativas independientes las enumeradas al principio excepto los ingresos del trabajo (Edad, Nivel de formación, Actividad económica, Sexo, País de nacimiento, Trabajo secundario, Tamaño de la empresa y Comunidad Autónoma)

A continuación, mediante el mecanismo de selección de categorías influyentes *stepwise*, se presenta un análisis simplificado de la regresión del error de medición (ECV – AEAT), sobre las variables explicativas indicadas¹.

¹ En la tabla aparecen las categorías que son significativas a un nivel $\alpha=0,00001$.

Categorías	Coeficiente	Error estándar	R2 parcial¹
Constante	-1.098,26	191,68	
45-64	-2.441,43	201,79	0,022
Educación Superior	-2.552,98	206,2	0,022
Actividades financieras y seguros	-6.942,65	557,51	0,016
Empresas con menos de 10 trabajadores	1.811,53	219,92	0,01
Mujeres	1837,28	202,95	0,008
Actividades sanitarias, serv. soc..	-1.791,59	354,33	0,003

Las conclusiones que extraemos del mismo (confirmando lo mostrado en el análisis descriptivo de los importes) son las siguientes:

- Las categorías que más influyen en el error de medición en sentido negativo son los siguientes:
 - tener 'Educación superior' (variable 'Nivel de formación')
 - edad comprendida entre 45 y 64 años (variable 'Edad')
 - trabajar en 'Actividades financieras' o en 'Actividades sanitarias' (variable 'Actividad económica')

El tener dichas características hace que los importes que se den a la ECV sean inferiores a los que se proporcionan a la Agencia Tributaria.

- Las categorías que más influyen en el error de medición en sentido positivo son:
 - Trabajar en una empresa 'con menos de 10 trabajadores' (variable 'Tamaño de la empresa')
 - Y ser 'Mujer' (variable 'Sexo')

El tener las mismas hace que los importes que se proporcionan a la encuesta sean superiores a los que se dan a la Agencia Tributaria.

¹ R2 Parcial: Es una medida de la relación lineal entre la variable dependiente y cada una de las independientes una vez que se ha eliminado en ambas variables los efectos debidos al resto de variables del modelo.

3 Ingresos del trabajo por cuenta propia

ANÁLISIS DE LA PERCEPCIÓN DE INGRESOS

Al igual que en el caso de ingresos del trabajo por cuenta ajena, se hace un estudio de los factores que influirán en que una persona que se considera autónoma según la encuesta, no figure como perceptor en los ficheros de la AEAT. En este caso realizaremos exclusivamente un análisis descriptivo de la percepción de ingresos dado que no contamos con suficiente muestra para poder realizar un estudio econométrico.

1. Análisis descriptivo de la percepción

En este apartado consideraremos el colectivo formado por aquellas personas que declaran solamente ingresos del trabajo por cuenta propia en la encuesta. Se tiene:

ECV 2011. Autónomos de la ECV, según tipo de perceptor de la AEAT. Porcentajes (datos muestrales)	
Total	100
T1. Solo con ingresos cuenta ajena en la AEAT	19,1
T2. Solo con ingresos cuenta propia en la AEAT	61,2
T3. Con ingr. cuenta ajena y propia en la AEAT	9,9
T4. Sin ingresos trabajo en la AEAT	9,9

Se observa que en el caso de cuenta propia casi el 10% de los autónomos de la encuesta, aparecen como no perceptores de rendimientos del trabajo según la AEAT (en 2010, un 9,1% de autónomos fueron no perceptores)

En la tabla adjunta, se muestra el porcentaje de no perceptores según la AEAT en función de distintas variables de clasificación. También se presenta el tamaño muestral para poder valorar la precisión de dichos porcentajes.

ECV 2011. Autónomos de la ECV. Porcentaje de no perceptores según la AEAT por variables de clasificación (datos muestrales)

	Tamaño muestral	Porcentaje no perceptores
Total	1.563	9,9
Edad		19,0
16 - 24	21	19,0
25 - 44	629	10,8
45 - 64	858	8,4
65 -	55	18,2
Nivel de formación		
Educación primaria o inferior	324	11,7
Educación secundaria primera etapa	454	7,7
Educación secundaria segunda etapa	369	9,5
Educación superior	416,0	11,1
Actividad económica		
Agricultura, ganadería, caza y silv.	200	6,0
Pesca	10	20,0
Industrias extractivas	2	0,0
Industria manufacturera	119	5,9
Prod. dist. energía eléct.,agua,etc.	20	5,0
Construcción	169	7,1
Comercio; reparación de vehículos, etc.	279	6,8
Transporte, almacenamiento	86	9,3
Hostelería	156	8,3
Información y comunicaciones	37	16,2
Actividades financieras y seguros	35	2,9
Actividades inmobiliarias	11	27,3
Actividades profesionales, cient.	106	7,5
Act. administrativas y serv. auxil.	13	15,4
Adm. Pública, Defensa y Seg. Soc.	3	0,0
Educación	25	8,0
Actividades sanitarias, servicios sociales	33	9,1
Actividades artísticas, recreativas	20	15,0
Otros servicios	100	11,0
Hogares que emplean personal doméstico	17	58,8
No consta	122	25,4
Sexo		
Varones	1.069	9,0
Mujeres	494	11,7
País de nacimiento		
España	1.470	8,6
EU-25 sin España	31	19,4
Resto de países	62	33,9
Trabajo secundario		
Con trabajo secundario	18	5,6
Sin trabajo secundario	1.545	9,9
Tamaño de la empresa		
Empresas con menos de 10 trabajadores	1.475	10,2
Empresas con más de 10 trabajadores	85	2,4
No consta	3	33,3
Comunidad Autónoma		
Andalucía	163	8,0
Aragón	81	4,9
Asturias (Principado de)	70	8,6
Baleares (Illes)	58	12,1
Canarias	73	11,0
Cantabria	52	5,8
Castilla y León	112	8,9
Castilla - La Mancha	131	6,1
Cataluña	194	10,8
Comunitat Valenciana	130	7,7
Extremadura	68	11,8
Galicia	116	13,8
Madrid (Comunidad de)	104	13,5
Murcia (Región de)	59	6,8
Navarra (Comunidad Foral de)	48	18,8
Rioja (La)	78	5,1
Ceuta y Melilla	26	34,6

Se puede resaltar lo siguiente:

En torno a un 10% de perceptores de rendimientos de actividades económicas, no tiene registro en la AEAT como perceptor de renta. Los porcentajes más elevados de no perceptores se concentran en los más jóvenes (considerando sólo los tramos de edad activa), en los que tienen el menor nivel de formación y el mayor nivel, y en los extranjeros. En relación con la actividad económica los mayores porcentajes de no perceptores se encuentran en aquellas actividades con menores tamaños muestrales, de manera que no se pueden extraer resultados concluyentes.

En cuanto a Comunidades Autónomas, los porcentajes más altos se alcanzan en Navarra (aunque en este caso hay que tener en cuenta el reducido tamaño muestral), Galicia, Comunidad de Madrid, e Illes Balears.

ANÁLISIS DE LOS IMPORTES

Tras realizar un análisis de los perceptores, examinaremos los errores de medición, comparando los importes recogidos en la encuesta con los de la AEAT.

1. Análisis descriptivo de los importes

Nos quedamos con aquellas personas que declaran solamente ingresos del trabajo por cuenta propia en la encuesta y en la AEAT.

ECV 2011. Autónomos. Estadísticas de los importes y de la diferencias (datos muestrales)

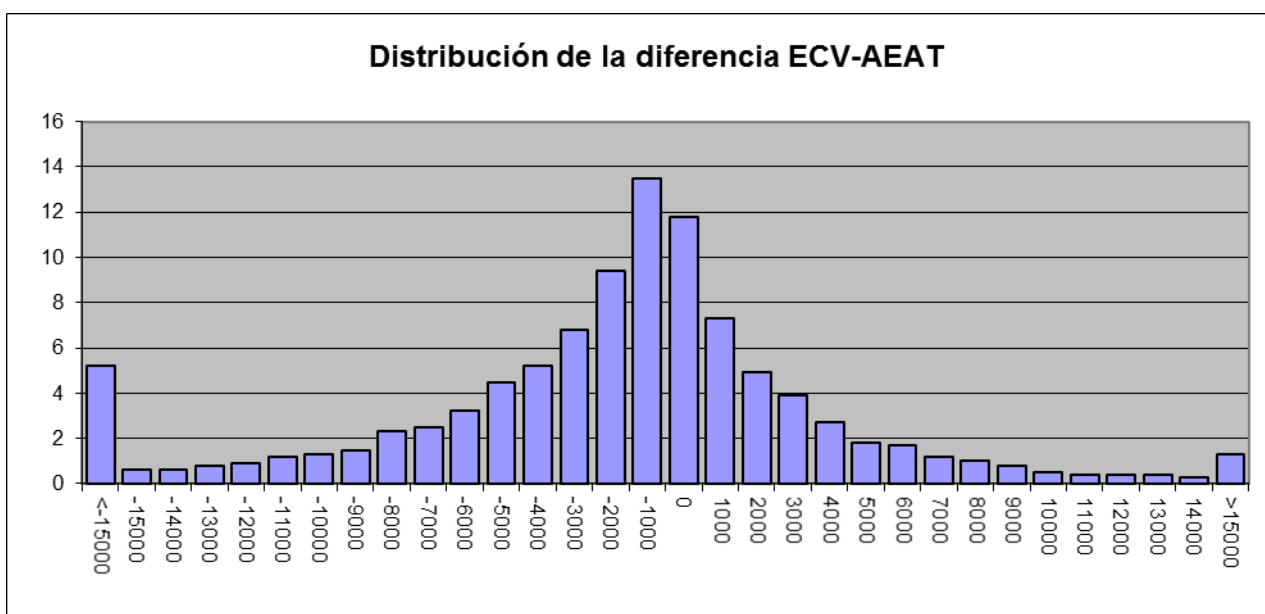
	Media	Desviación estándar	Percentil 25	Mediana	Percentil 75
ECV	9.623	16.533	3.533	7.000	12.219
AEAT	10.514	24.239	3.348	7.472	12.682
ECV - AEAT	-891	27.417	-7.818	-1.586	4.524
ECV - AEAT	12.173	24.580	2.780	6.034	12.436

Observando las medidas descriptivas se desprende que los ingresos medios que perciben los autónomos proporcionados a la encuesta son inferiores a los que se facilitan a la AEAT y que los datos de la Agencia Tributaria presentan una mayor dispersión. Por otro lado, en el percentil 25 el valor de la ECV es superior al de la AEAT mientras que en los percentiles 50 y 75 ocurre lo contrario, el valor de la ECV es inferior al de la AEAT.

La diferencia entre los importes medios dados a la AEAT y a la ECV es menor en el caso de los autónomos que en el caso de los asalariados. Sin embargo en términos absolutos la diferencia es muy importante, en particular en los niveles altos de renta como se verá después.

Hay que aclarar que en el caso de estimación objetiva (módulos) se ha tomado como ingreso de la AEAT el rendimiento neto de la actividad, que en realidad es una imputación del beneficio de la actividad.

En el siguiente gráfico se tiene la distribución de la diferencia entre el valor de la encuesta y de la AEAT. Se observa una distribución con mayor variabilidad que en el caso de los ingresos por cuenta ajena.



En la tabla adjunta se realiza un análisis descriptivo de la diferencia de importes en función de las distintas variables de clasificación:

**ECV 2011. Autónomos. Diferencias ECV-AEAT por variables de clasificación
(datos muestrales)**

	Tamaño muestral	AEAT	ECV-AEAT (Media)	ECV-AEAT (Media)
Total	951	10.514	-891	12.173
Edad				
16 - 24	8	9.665	2.191	15.746
25 - 44	369	8.844	367	9.919
45 - 64	539	11.551	-1.731	13.687
65 -	35	12.341	-1.923	11.807
Nivel de formación				
Educación primaria o inferior	221	7.888	-157	10.906
Educación secundaria primera etapa	301	8.205	495	10.201
Educación secundaria segunda etapa	211	9.916	871	11.774
Educación superior	218	16.942	-5.254	16.568
Tramos rend. Act. Econ. AEAT				
-9000	563	2.940	3.368	9.107
9000 - 20000	277	13.043	-126	10.998
20000 - 30000	55	24.275	-4.662	15.524
30000 -	56	60.627	-43.790	45.525
Actividad económica				
Agricultura, ganadería, caza y silv.	168	8.317	-317	10.147
Pesca	7	20.068	-2.901	20.868
Industria manufacturera	63	9.920	5.223	16.915
Prod. dist. energía eléct., agua, etc.	7	14.342	-7.099	9.272
Construcción	82	8.946	4.418	12.567
Comercio; reparación de vehículos, etc.	186	11.450	-3.931	11.075
Transporte, almacenamiento	56	4.748	3.836	11.299
Hostelería	102	8.514	149	10.048
Información y comunicaciones	15	7.506	4.218	9.860
Actividades financieras y seguros	16	61.346	-43.551	49.919
Actividades inmobiliarias	3	3.835	4.346	14.680
Actividades profesionales, cient.	56	11.822	-266	11.225
Act. administrativas y serv. auxil.	3	20.072	14.959	21.458
Adm. Pública, Defensa y Seg. Soc.	1	14.536	4.444	4.444
Educación	11	10.371	-978	11.413
Actividades sanitarias, servicios sociales	25	28.012	-17.495	22.878
Actividades artísticas, recreativas	13	11.082	-2.490	10.638
Otros servicios	68	6.658	3.263	8.805
Hogares que emplean personal doméstico	7	10.614	-4.043	4.900
No consta	62	6.615	-668	11.136
Sexo				
Hombres	625	11.215	-842	13.157
Mujeres	326	9.169	-986	10.287
País de nacimiento				
España	908	10.593	-1.104	12.136
EU-25 sin España	19	6.885	3.208	8.033
Resto de países	24	10.375	3.895	16.846
Trabajo secundario				
Con trabajo secundario	9	8.273	4.740	10.729
Sin trabajo secundario	942	10.535	-945	12.187
Tamaño de la empresa				
Empresas con menos de 10 trabajadores	932	10.356	-947	12.097
Empresas con más de 10 trabajadores	17	18.140	3.166	15.660
No consta	2	19.346	-9.546	17.970
Comunidad Autónoma				
Andalucía	112	10.454	-203	13.325
Aragón	53	16.717	-9.142	18.939
Asturias (Principado de)	41	4.136	7.725	14.099
Baleares (Illes)	28	7.218	5.564	11.954
Canarias	42	12.201	752	11.898
Cantabria	36	2.571	2.100	8.154
Castilla y León	86	9.940	-2.106	10.531
Castilla - La Mancha	86	7.353	1.358	8.919
Cataluña	111	15.224	-4.276	13.899
Comunitat Valenciana	78	7.945	1.551	10.570
Extremadura	42	10.822	-212	11.896
Galicia	64	7.494	-51	8.987
Madrid (Comunidad de)	54	12.430	230	15.119
Murcia (Región de)	35	11.232	-3.007	10.951
Navarra (Comunidad Foral de)	25	16.747	-8.132	15.028
Rioja (La)	47	7.365	2.963	7.472
Ceuta y Melilla	11	36.821	-30.280	30.280

Se observa que:

- Si comparamos con lo que ocurre en cuenta ajena, la diferencia entre importes toma valores positivos en un mayor número de casos. Sin embargo, hay que destacar que esta diferencia es negativa y tiene un valor elevado en el caso de individuos con una educación superior y también en los últimos tramos de rentas. En todo caso en la interpretación de los resultados hay que tener en cuenta el tamaño muestral.
- Dentro de las actividades económicas la diferencia es muy negativa en las categorías siguientes: 'Actividades financieras y seguros' y 'Actividades sanitarias'.

Referencias:

- [1] Eurostat (2013), "The use of registers in the context of EU-SILC: challenges and opportunities". Eurostat Statistical Working Papers. Edited by Markus Jäntti, Veli-Matti Törmälehto and Eric Marlier.
- [2] Dauphin, L. (2009) "La qualité des données sur les revenus : enquête versus fichier administratif". Les Journées de Méthodologie Statistique (INSEE).
- [3] Donatiello, G., Betti, G. and Consolini, P (2009). "Integrazione di dati campionari Eu-Silc con dati di fonte amministrativa". Metodi e Norme n.38 (ISTAT).
- [4] Cristia, J. and Schwabish, J. (2009) "Measurement error in the SIPP: Evidence from administrative matched records", *Journal of Economic and Social Measurement* 34, 1-17.
- [5] Abowd, J. M and Stinson, M. H. (2003), "Estimating Measurement Error in SIPP Annual Job Earnings: A Comparison of Census Survey and SSA Administrative Data".
- [6] Gottschalk, P. and Huynh, M. (2005-2007) "Validation Study of Earnings Data in the SIPP - Do Older Workers Have Larger Measurement Error?". Center for Retirement Research at Boston College Working.