

# ESTADÍSTICAS DE FLUJOS DE LA POBLACIÓN ACTIVA.

## INTRODUCCIÓN.

Bajo la denominación de 'Estadísticas de Flujos de la Población Activa' se han venido considerando todas aquellas informaciones relativas al seguimiento de la población, o grupos particulares de ella, a lo largo del tiempo, en lo referente a su relación con el mercado laboral. Esta información vincula las situaciones en las que se encuentre el entrevistado en dos momentos de tiempo sobre la base de la Encuesta de Población Activa (EPA), con el objetivo de estudiar su evolución.

El diseño de la encuesta, que se renueva por sextas partes cada trimestre, manteniéndose fijas las 5/6 partes entre dos trimestre consecutivos, (4/6 entre un trimestre y el subsiguiente, etc.) y que incluye entre sus preguntas la situación en que el entrevistado se encontraba en un momento del pasado, permite obtener estadísticas de flujos de dos maneras distintas: la que se basa en las preguntas retrospectivas, con información relativa a las dos referencias temporales de interés recogidas en un solo trimestre y que se explota junto con el resto de las variables trimestrales y la que aprovecha la parte común de la muestra entre dos realizaciones de una misma encuesta. Esta última, requiere la elaboración de ficheros especiales de flujos que permitan vincular registros comunes en trimestres distintos.

Cada uno de estos procedimientos tiene sus ventajas y debilidades. No obstante, la limitación de las preguntas retrospectivas que recargarían excesivamente el cuestionario que por otro lado, se vería muy influenciado negativamente por la falta de memoria, hacen que el análisis de la parte común de la muestra mediante ficheros especiales de flujos sea el más interesante.

En la presente publicación nos centraremos precisamente en este último análisis. La razón de ello, por un lado, es ofrecer a los usuarios algunas tablas de referencia que permitan contrastar los datos a partir de los ficheros especiales de flujo, cuya explotación es técnicamente más compleja. Por otro lado, parece razonable priorizar el aprovechamiento de la información longitudinal a partir de la muestra común EPA que es exclusivo de una encuesta continua con rotación parcial de la muestra (las preguntas retrospectivas pueden ser incluidas en cualquier tipo de encuesta).

Coincidiendo con la renovación de la Encuesta de Población Activa en el segundo trimestre de 1987, el INE inició series y estudios sobre flujos de la población respecto a la actividad económica basados en el análisis longitudinal de la muestra común en periodos sucesivos, lo que supuso una importante novedad respecto a los métodos tradicionalmente llevados a cabo hasta esa fecha.

Los análisis detallados de datos sobre flujos y la metodología utilizada se pueden encontrar en las siguientes publicaciones:

Encuesta de Población Activa. Estadística de Flujos. 2º trimestre 1987- 2º trimestre 1988. INE. Madrid 1989.

Encuesta de Población Activa. Estadística de Flujos. 1988-1990. INE. Madrid 1992

Encuesta de Población Activa. Estadística de Flujos. 1990-1992. INE. Madrid 1995

Encuesta de Población Activa. Estadística de Flujos. 1992-1994. INE. Madrid 1997

Encuesta de Población Activa. Estadística de Flujos. Metodología, INE. Madrid, 1996

Encuesta de Población Activa. Estadística de Flujos. 1984-1996. INE. Madrid, 1999

Análisis de flujos del mercado laboral a partir de datos de la Encuesta de Población Activa. Indicadores Sociales de España 2003. INE. Madrid, 2003

La presente publicación en INEBASE pretende continuar con la difusión de estadísticas de flujos ofreciendo por vez primera los resultados a partir de la EPA metodología 2005 (es decir con las poblaciones base 2001 introducidas en marzo de 2005, que llevó a revisar las series EPA para el periodo 1996-2004).

Pueden consultarse series desde 1996 separadas en bloques distintos: 1996-2000, 2001-2004 y desde 2005 en periodos de dos años. La primera división (2000 y 2001) fue debida a la implantación de la definición operativa armonizada sobre la consideración de parado del Reglamento 1897/2000 de la CE, y sus consecuencias sobre el colectivo de inactivos. A partir de 2001 los datos sobre parados e inactivos no son directamente comparables con los periodos anteriores. La segunda (2004 y 2005) obedece a las variaciones en las series que produjo el cambio de los métodos de recogida de 2005. Desde entonces, a pesar de la presentación de los resultados en periodos de dos años, los datos son homogéneos a todos los efectos.

Además de la difusión bienal de resultados, cada trimestre, conjuntamente con la publicación de la EPA, se incluye en la correspondiente nota de prensa una tabla sobre 'movilidad en relación con la actividad entre le trimestre actual y el anterior' que se calcula con la misma metodología que se describe a continuación.

## **EL ANÁLISIS DE LOS FLUJOS A PARTIR DE LA MUESTRA COMÚN DE LA EPA.**

Las estadísticas de flujos EPA basadas en el aprovechamiento de la parte común de la muestra que comprende las 5/6 partes de las viviendas entre un trimestre y el siguiente, las 4/6 partes entre un trimestre y el subsiguiente, etcétera, se fundamentan en una encuesta en *panel*, en la que cada individuo de 16 y más años cuenta con un registro con datos de todas las entrevistas.

Cuadro 1. Esquema de turnos de rotación de la Encuesta de Población Activa

Turno de rotación	Trimestre					
	T+1	T+2	T+3	T+4	T+5	T+6
1	1ª entrev.	2ª entrev.	3ª entrev.	4ª entrev.	5ª entrev.	6ª entrev.
2	6ª entrev.	1ª entrev.	2ª entrev.	3ª entrev.	4ª entrev.	5ª entrev.
3	5ª entrev.	6ª entrev.	1ª entrev.	2ª entrev.	3ª entrev.	4ª entrev.
4	4ª entrev.	5ª entrev.	6ª entrev.	1ª entrev.	2ª entrev.	3ª entrev.
5	3ª entrev.	4ª entrev.	5ª entrev.	6ª entrev.	1ª entrev.	2ª entrev.
6	2ª entrev.	3ª entrev.	4ª entrev.	5ª entrev.	6ª entrev.	1ª entrev.

El sombreado en color corresponde a los turnos de rotación que constituyen la muestra común respecto del trimestre T. Después de los seis trimestres se ha producido la renovación completa de la muestra

## **LIMITACIONES DEL PROCEDIMIENTO.**

### **1.-Pérdida de muestra.**

El enlace de información no puede establecerse para todos los registros de la muestra inicial común. Siempre existen registros de la muestra común para los que no existe información en alguno de los períodos que se pretenden analizar (negativa o ausencia de un grupo familiar, inaccesibilidad de las viviendas, altas y bajas en los hogares colaboradores, cambio del grupo familiar que habita la vivienda, etc.)

Además, el enlace se efectúa prescindiendo de los registros imputados en cada periodo. Es decir, aquellos registros que son "no respuesta" en un trimestre y contestaron en el trimestre inmediatamente anterior. Su falta de respuesta en el trimestre actual se subsana copiando los datos del trimestre precedente, solucionando así en parte la falta de respuesta de la EPA

trimestral. En el caso del análisis de flujo su inclusión haría aumentar artificialmente el número de personas que no cambia de situación de un periodo al siguiente.

## **2.-Inconsistencias de respuesta**

La estadística de flujos como producto derivado de la Encuesta de Población Activa presenta inconsistencias entre la información recabada en los distintos periodos analizados. Así, por ejemplo, es posible encontrar personas que han trabajado en un trimestre y que posteriormente declaran que están buscando su primer empleo, o que tienen un nivel de estudios inferior al declarado anteriormente. Hay que tener en cuenta que en la EPA se admite la información suministrada por otra persona del hogar distinta a la de referencia; estos cambios de informante pueden dar lugar a inconsistencias del tipo descrito.

## **3.- Significación de los resultados**

Dado que la EPA es una encuesta por muestreo, los resultados que proporciona llevan asociados un error de muestreo que, en general, es mayor cuanto menor es el colectivo y cuanto menos uniformemente esté distribuida la variable o variables a que se refiera.

En el caso de los datos de flujos el problema de los errores de muestreo asociados a colectivos pequeños se acentúa, pues al vincular dos periodos de tiempo, en realidad, se están multiplicando las categorías de análisis con el consiguiente riesgo de que algunos cruces se queden con poca muestra y por tanto las estimaciones puedan estar afectadas por fuertes errores de muestreo.

## **VALIDEZ DE LAS ESTADÍSTICAS DE FLUJOS**

Es requisito indispensable para la validez de las estimaciones, que la parte de la muestra considerada sea representativa de la población total. Todos los turnos de rotación tienen la misma distribución por estratos, de modo que la muestra perdida tendría las mismas características que la enlazada.

En segundo lugar, el conjunto de pérdidas de registros originadas por falta de respuesta y las altas y bajas de personas en uno de los trimestres, no suponen una cifra elevada respecto al potencial de registros comunes.

Por ello, ni la pérdida de muestra que supone la rotación de 1/6 de la misma cada trimestre ni la debida a las pérdidas de registros de personas de 16 y más años o a las posibles inconsistencias entre trimestres, suponen riesgo alguno en lo que respecta a representatividad y más teniendo en cuenta los bajos niveles de desagregación de las categorías que muestran las tablas de flujos, en las que tampoco se dan datos de variables de respuesta subjetiva. Al considerar colectivos relativamente importantes, es decir, con una representación muestral elevada, los errores de muestreo se mantienen en unos límites tolerables.

## **EL PROBLEMA DE LOS FACTORES DE ELEVACIÓN.**

Los factores de elevación asociados a cada individuo de la muestra equivalen al número de individuos de la población total a los que representa, por ello, las cifras totales de las estadísticas de flujos representan sólo a la parte común de la muestra en función del número de trimestres consecutivos que se estudien. Así, por ejemplo, entre dos trimestres consecutivos, la cifra total corresponde a algo menos de las 5/6 partes de la población (teniendo en cuenta las carencias citadas por negativas, ausencias, etc...), entre tres trimestres consecutivos se reducirá a algo menos de las 4/6 partes, etc. Por esta razón no se pueden interpretar como flujos brutos de entrada y salida de un trimestre a otro, ni proporcionan una estimación coherente de los flujos de trasvase en relación a los efectivos de

los dos trimestres. Su representación es coherente cuando se observa en términos porcentuales.

Para que los flujos brutos de un trimestre a otro fueran representativos de la totalidad de la población se necesitaría un nuevo factor de elevación partiendo de la muestra común, calculando igualmente el cociente entre el valor poblacional y el muestral para cada estrato. Lo cual exigiría además la obtención de nuevas proyecciones de población (para el numerador del factor) cuestión que conllevaría el problema de elegir fecha de referencia para las mismas y además daría estimaciones distintas para las distintas categorías poblacionales de las procedentes de la EPA en cada trimestre. Tampoco serían coincidentes las estimaciones para un trimestre dado y el siguiente, con las obtenidas para ese mismo trimestre enlazado con los dos siguientes, pues aunque el punto de partida es el mismo en los dos casos, tanto las muestras comunes como las proyecciones son distintas en cada caso. Volveríamos a considerar representativos, al igual que en el procedimiento anterior, únicamente los datos porcentuales.

## ESQUEMA GENERAL DE LAS TABLAS DE FLUJOS A PARTIR DE LA MUESTRA COMÚN.

Las ventajas que supone el estudio de flujos a partir de la parte común de la muestra en dos trimestres consecutivos, con factores de elevación procedentes de cada muestra trimestral, han sido decisivas en la elección de dicho método en el análisis de las tablas de flujos de la presente publicación.

Se efectúa por tanto, un seguimiento de cada persona de 16 y más años de forma que se ha agrupado la información que dicha persona ha proporcionado en los dos períodos enlazados, combinando ambos. Se calculan cruces entre las categorías de una variable de un período con categorías de la misma u otras variables en el otro, presentados en cifras relativas que bien podemos considerar en términos de probabilidad. Así, por ejemplo, en la tabla del cuadro, tendríamos que la 'probabilidad de transición entre los 'estados'  $E_i - E_j$ ' es  $p_{ij}$ .

### Esquema general de tabla de flujos a partir de muestra común

Ta	Tb			
	E1'	E2'	.....	Em'
E1	p11	p12	.....	p1m
E2	p21	p22	.....	p2m
....	....	....	....	....
En	pn1	pn2	....	pnm

$p_{ij}$ : probabilidad de 'paso/procedencia' desde  $E_i$  a  $E_j$

Ta, Tb: Periodos de referencia a y b

$E_i/E_j$ : Estados analizados

En el caso de que el factor de elevación que se utilice sea el del primer periodo, estaríamos ante un **análisis de evolución**, es decir, la situación futura en términos de probabilidad de las personas observadas en el trimestre de partida ( $p_{ij}$ , siendo "a" el período de partida que marca los factores de elevación).

Si el factor utilizado es el del período final o de llegada, tendríamos un **análisis de procedencia**, o lo que es lo mismo la situación de origen, en términos de probabilidad, en el trimestre de llegada ( $p_{ij}$ , con factores de elevación del período "b").

## TABLAS

Los resultados que ofrece esta publicación muestran el **análisis de procedencia**, es decir la distribución de la situación de partida en que confluye cada categoría estudiada en el trimestre de llegada al que se refieren los datos.

Existen tres tipos de tablas. Las denominadas de “**Movilidad**” recogen la distribución porcentual de las principales categorías (ocupados, parados e inactivos) con respecto a las mantenidas en el trimestre anterior. Las “**Nacionales**” recogen las probabilidades de procedencia entre dos categorías en un análisis longitudinal, por grupos de edad y sexo. Y, finalmente las tablas por “**Comunidades autónomas**” muestran esas mismas probabilidades en el ámbito territorial autonómico.