

¿SON BAJOS LOS FLUJOS MIGRATORIOS EN ESPAÑA?*

CARMEN RÓDENAS

MÓNICA MARTÍ

Universidad de Alicante

La *Encuesta de Migraciones* procedente de la EPA ha venido mostrando a lo largo del último decenio una baja y decreciente intensidad del flujo migratorio en España, incluso cuando nuestra economía ha atravesado etapas de recuperación y creación de empleo. Este sorprendente escenario que reflejan los datos se torna inverosímil cuando se comparan esas cifras con otras fuentes de información continua sobre la corriente migratoria. Nuestro análisis de la metodología aplicada por el INE nos ha permitido establecer la existencia de una importante infravaloración congénita en la estimación de la movilidad a través de esta estadística.

Palabras clave: migraciones, EPA, metodología.

En España es sabido que, aunque el paro es alto en general, existe una fuerte concentración regional del desempleo. En este contexto, ha llamado frecuentemente la atención entre los investigadores que la movilidad geográfica del trabajo, según se desprende de la *Encuesta de Migraciones* (EM) de la Encuesta de Población Activa (EPA), sea tan baja y, además, decreciente, incluso cuando nuestra economía ha atravesado etapas de recuperación económica y de creación de empleo. No sólo por esta razón resultan llamativas las estimaciones de esta fuente estadística. Cuando se compara la baja intensidad y decreciente tendencia del flujo migratorio descrito por la misma con los resultados de otras fuentes alternativas surgen serias dudas acerca de si es efectivamente tan baja la movilidad interior. En nuestra opinión¹, la sospecha de que existen problemas graves con la EM, frecuentemente utilizada en los análisis microeconómicos de las migraciones, justifica que ese sea, precisamente, el objetivo del trabajo: un estudio a fondo de la información básica.

(*) Queremos agradecer al INE la disponibilidad mostrada para facilitarnos la información estadística que hemos utilizado y, en particular, a J.A. Díez por sus valiosos comentarios a los borradores previos. También nos encontramos en deuda con el grupo de mercado de trabajo del Dpto. de Fundamentos de la Economía e Historia Económica de la Universidad de Alcalá de Henares, especialmente, con el profesor Luis Toharia; con los evaluadores anónimos del trabajo y, como siempre, con los compañeros del Seminario de Investigación de nuestro Departamento. Este trabajo ha disfrutado, además, de la financiación procedente de la Generalitat Valenciana a través del Programa de Proyectos de Investigación Científica y Desarrollo Tecnológico (ref. GV-2455-94).

(1) Coincidiendo con la sombra de duda que exponen en sus trabajos Albérico y Jimeno (1993), Bentolila (1997a, b), y que ya intuía Olano (1990).

El trabajo consta de cinco partes. En la primera se plantean los *hechos*, que no son otros más que las sorprendentes diferencias que muestran las principales fuentes de información continua sobre los flujos migratorios. Tras una valoración inicial, en la segunda parte, de la *Estadística de Variaciones Residenciales* (EVR) y de la EM, la tercera parte del trabajo se dedica a la discusión de la infravaloración *congénita* de los flujos que presenta esta última. Además, se explican las consecuencias del cambio en 1987 del procedimiento de estimación de los migrantes en la EPA. En la cuarta parte se presentan las evidencias de la subestimación utilizando los datos de los ficheros anonimizados de la EPA y, para acabar, en la última parte se recogen algunas conclusiones y sugerencias.

1. MIGRACIONES Y MIGRANTES: LA INFORMACIÓN SE CONTRADICE

Hay dos fuentes de información estadística acerca de los movimientos migratorios interiores con periodicidad anual en España² que se diferencian tanto por su metodología como por sus contenidos. Así, con carácter registral, esto es, con origen en los registros administrativos de empadronamiento municipal se puede utilizar la EVR (altas y bajas por cambio de municipio de residencia)³. Procedente de encuesta, es decir, elaborada a través del muestreo de poblaciones, se dispone de la EM que se elabora desde 1980 al tiempo que se realiza la EPA. Mientras que la primera mide *migraciones* y hasta los años ochenta era la única existente, la segunda mide *migrantes* –más concretamente, inmigrantes– y en la actualidad es la única que informa sobre las situaciones laborales de los migrantes⁴.

La EM tiene carácter retrospectivo, es decir, compara el municipio actual de residencia del entrevistado con el que tenía un año antes y, por tanto, contabiliza un inmigrante en el caso de que la actual residencia no coincida con la del año anterior. El procedimiento para captar a los inmigrantes en la EM se ha modificado desde su inicio. Así, entre 1980 y 1985 sólo se preguntaba por su residencia un año atrás a los individuos de la parte de la muestra de viviendas EPA que entraba en la encuesta por primera vez en cada uno de los cuatro trimestres del año (un sexto por trimestre). A partir de tal muestra, se realizaba la estimación del número anual de migrantes. Como se señala en INE (1996a:14), desde 1986 las cifras estimadas parten de las respuestas a la pregunta retrospectiva que se realiza a todos los individuos de la muestra de viviendas EPA encuestada en los segundos trimestres de cada año⁵.

(2) En los *Padrones municipales de habitantes* y en los *Censos de población* también se puede encontrar información sobre los movimientos migratorios y los migrantes. Sin embargo, el cruce de su baja frecuencia –ya que se elaboran cada cinco y diez años, respectivamente– con el típico problema de la *memoria histórica* de los sujetos entrevistados, no las hacen ser las fuentes estadísticas más idóneas para los análisis coyunturales de paro y migraciones.

(3) Publicada habitualmente por el INE en el *Anuario estadístico* y, desde 1985, en la serie *Migraciones*. Y utilizada en los estudios de Bentolila y Dolado (1991), Ródenas (1994) o Bentolila (1997 a y b).

(4) Ésta es la fuente utilizada en los análisis de tipo microeconómico que contemplan las características personales de los inmigrantes relativas al mercado de trabajo, como el de Albérico y Jimeno (1993) o el de Antolín y Bover (1997).

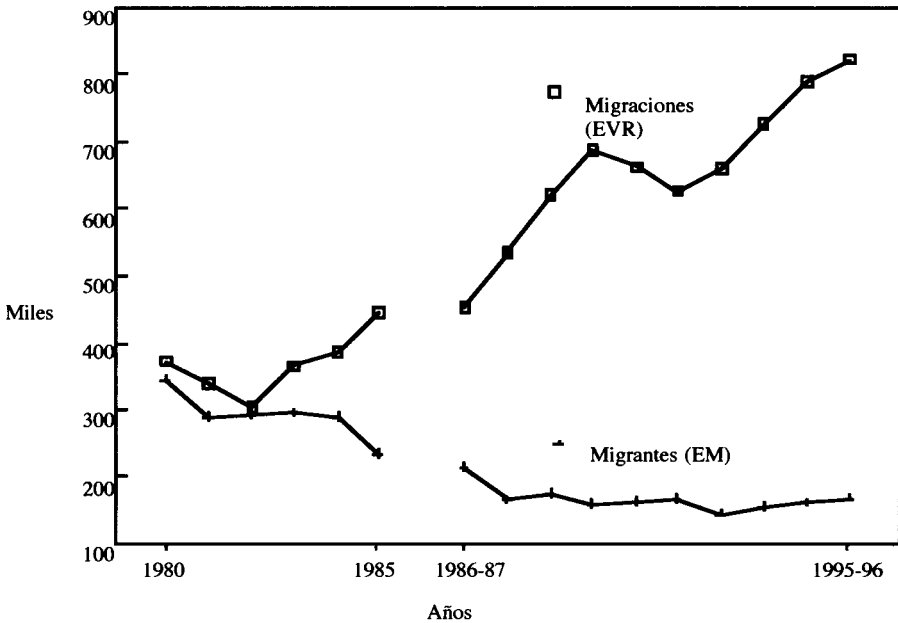
(5) Este cambio permitió al INE incorporar las directrices procedentes del EUROSTAT en relación con la Encuesta de Fuerzas de Trabajo (EFT) comunitaria. Directrices que en ningún caso limitaban al mismo a efectuar en una única ocasión anual las preguntas retrospectivas de la EPA.

Por su parte, la EVR –publicada desde 1961– se elabora a partir de los datos individuales acerca de la población migrante que los ayuntamientos remiten a la Delegación del INE en su provincia. Los datos, que proceden de la gestión de altas y bajas de los Padrones municipales, recogen mensualmente las migraciones por municipio de origen y de destino, así como su clasificación en función de un conjunto de variables individuales.

Como ambas fuentes ofrecen estimaciones de la corriente migratoria –entendida siempre como cambio de municipio de residencia–, lo lógico es esperar que dicho flujo medido a través de las migraciones (EVR) o medido a través de los migrantes (EM), presente series con alguna semejanza. Salvado el desfase temporal ajustando anualmente los datos de la EVR por segundos trimestres a partir de 1986-87 de acuerdo con el nuevo procedimiento de estimación de la EM, el gráfico 1 muestra no sólo que ambas series se mueven en niveles muy alejados, sino también que las tendencias de una y otra aparentan un comportamiento divergente.

En general, mientras que las migraciones crecen casi ininterrumpidamente desde 1982, el número de migrantes cada vez es menor. De hecho, el nivel medio de migrantes entre 1986 y 1996 supone prácticamente un tercio del existente entre 1980 y 1985. Aunque es lógico que el número de migraciones supere al de migrantes, dado

Gráfico 1: MIGRACIONES Y MIGRANTES. 1980-1996



Nota: Migrantes y migraciones interiores (todas las edades). EVR 1995/96 datos provisionales.

que una persona puede moverse más de una vez a lo largo del año, lo que llama la atención es que cruzando la información se habría pasado de una media de 1,3 migraciones por migrante y año en 1980-85, a una media de más de 4 en 1986-96.

Aceptar esto sin más es difícil, pues significaría que en España tendríamos una creciente y elevadísima concentración de la movilidad geográfica en un muy reducido grupo de efectivos, algo así como un *núcleo duro* de migrantes. El problema es que hoy eso no se podría demostrar, dado que no disponemos de las historias migratorias individuales. Cuando se utiliza la única información puntual más parecida, concretamente la *Encuesta Socio-demográfica* de 1991, la media de migraciones por migrante se deshincha drásticamente a 1,8 movimientos a lo largo de *toda su vida* [INE (1991a:91)] Con este dato de “control” las cifras medias antes citadas referidas a los *movimientos por migrante y año* resultan completamente disparatadas. Si fueran ciertas, mostrarían tanto una increíble frecuencia de los movimientos –cada tres meses–, como un extremado civismo entre los inmigrantes al darse sistemáticamente de alta en cada una de las nuevas residencias.

Por eso, más adecuado que aceptar y trabajar según la hipótesis del *núcleo duro* de migrantes, sería averiguar qué mide exactamente cada una de las dos fuentes estadísticas periódicas de que se dispone y qué no son capaces de recoger. De hecho, mostraremos que buena parte de la explicación de que el número de migrantes captados a través de la EPA haya disminuido, y que lo haya hecho al entrar nuestra economía en una fase de crecimiento económico como la iniciada en 1986, reside en el cambio de procedimiento de estimación que utiliza el INE que, precisamente, se produce en ese momento.

2. LOS INCONVENIENTES DE LAS FUENTES ESTADÍSTICAS

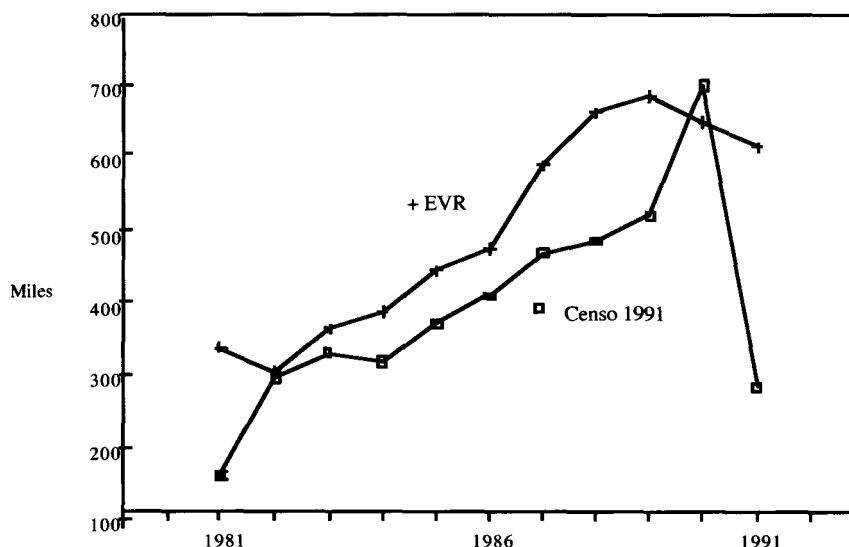
Una valoración inicial de ambas fuentes nos obligaría a decir que tanto la EVR como la EM presentan ciertos inconvenientes. La EVR tiene una serie de deficiencias que podrían plantear algunos problemas de fiabilidad. En primer lugar, ha venido mostrando hasta ahora las interferencias provocadas por la elaboración simultánea de otro registro administrativo: el Padrón municipal de habitantes. Todos los años de referencia de los Padrones se han caracterizado en la EVR por registrar mínimos volúmenes migratorios. Este *fenómeno* aparece porque durante los meses en los que se estaba elaborando cada Padrón, los nuevos residentes eran dados de alta directamente en el Padrón municipal de su actual residencia y, simultáneamente, de baja en el correspondiente Padrón de su municipio de origen. Por tanto, no quedaban registrados en las cifras mensuales de variaciones residenciales, y el volumen de movimientos para ese año completo presentaba un valor anormalmente bajo⁶. Y, en segundo lugar, la falta de inscripción padronal de una parte –quizá importante, pero no conocida– de la población que cambia de residencia podría generar cierto grado de subestimación en las migraciones proporcionadas por la EVR. En la medida en que no es imprescindible el empadronamiento –aunque legalmente sí exista la obligación– para el desarrollo de la vida y las actividades de los ciudadanos en sus municipios de residencia, muchos trasladados no serán registrados por sus protagonistas.

(6) Por eso, las cifras del gráfico 1 de la EVR en los años 1981, 1986 y 1991 se han corregido interpolando los datos del año previo y del siguiente.

Ahora bien, hay varios factores a favor de la fiabilidad de la EVR. Por un lado, el resultado de la comparación de su tendencia con la del número de migrantes del Censo de 1991. Como se muestra en el gráfico 2, en general la población censada con movilidad en cada uno de los años entre 1981 y 1991 evoluciona de forma parecida a la de la EVR.

Por otro lado, como señala reiteradamente el INE en la serie *Migraciones*, cada vez es mejor la calidad y cobertura de la EVR, tanto porque se aplican más avanzados procedimientos de control y de recogida de los datos, como porque han ido surgiendo múltiples factores que han mejorado su cobertura, al exigirse progresivamente el certificado de empadronamiento para acceder a servicios sociales y municipales básicos⁷ y para ejercer determinados derechos democráticos como el voto. Si se admite lo señalado por el INE, esto es, que las cifras de la EVR se acercan cada vez más a la realidad, entonces habría más motivos para fundamentar la sospecha de que la EM subesti-

Gráfico 2: MIGRANTES CENSO 1991 Y MIGRACIONES EVR



Nota: Para 1991 el censo sólo contabilizó los migrantes del primer trimestre.

(7) Aunque de forma discrecional, esto es, dependiendo de cada municipio y organismo, es necesaria la acreditación de la residencia para actividades como la renovación del carnet de identidad, el acceso a la enseñanza pública –primaria, media y universitaria– y sus becas, las subvenciones para la compra de viviendas públicas, el alta de la luz en el domicilio particular o el cobro de las pensiones no contributivas. Asimismo, las alteraciones o cambios en la cartilla de la Seguridad Social o la obtención de un empleo público, requieren del correspondiente certificado de empadronamiento.

ma el número de migrantes. Por tanto, que no debería utilizarse para la cuantificación de la intensidad de los flujos y que, en principio, éstos podrían o no estar lo suficientemente bien caracterizados a través de la misma.

En el caso de la EM, en la memoria explicativa que acompaña a su publicación anual, el INE ya señala que la misma no está diseñada expresamente para el estudio del fenómeno migratorio. Los datos de migrantes son uno de los resultados de una encuesta, la EPA, destinada a otro fin. La EM es un “subproducto” de la EPA, la cual tiene como objetivo estimar las situaciones de actividad e inactividad, paro y ocupación entre la población total, pero no las de movilidad ni los cambios de residencia. Por ello, el diseño de las preguntas a partir de las que se elabora la EM no es el mejor de los posibles⁸. En principio, en la EM sólo se captará a los inmigrantes que residen en viviendas familiares y, desde 1987, habrán de sobrevivir en la muestra hasta el momento en que son encuestados. Asimismo, nunca podrán cubrirse las migraciones intermedias –escalonadas o de ida y vuelta–, si es que las hubiera habido en el transcurso del segundo al segundo trimestre de cada año.

No acaban aquí las cosas, un análisis más exhaustivo del método que se utiliza en la EPA para estimar a los migrantes, conduciría la atención nuevamente hacia sus problemas de subestimación, en concreto hacia tres cuestiones. Primero, cómo puede enmascarse un movimiento migratorio ante el tipo de muestreo, las definiciones e instrucciones dadas a los encuestadores para la cumplimentación del cuestionario; segundo, qué efectos tiene la falta de respuesta y la pérdida de muestra y, tercero y más importante, qué proporción de inmigrantes nunca podrá ser captada debido a la formulación de la pregunta retrospectiva que actualmente sólo se realiza en el segundo trimestre. Las tres cuestiones se desarrollan bajo el epígrafe siguiente.

3. POR QUÉ LA EM SUBESTIMA A PRIORI EL NÚMERO DE MIGRANTES

3.1. *El muestreo, las definiciones y las instrucciones a los encuestadores*

La unidad primaria del muestreo bietápico en el que se basa la EPA es la *sección censal*. Los criterios geográficos, demográficos y socioeconómicos son los que sirven para estratificar y seleccionar las secciones censales que entran en la encuesta, y su renovación va asociada a los cambios en los Censos y Padrones. Aunque los nuevos Censos y Padrones revelan las zonas de crecimiento demográfico, el problema es que lo hacen *a posteriori* y, por eso, las secciones censales usadas en la EPA no garantizan tener la capacidad para captar los movimientos migratorios en el momento y lugar en los que se están produciendo. Es decir, el movimiento demográfico de las zonas metropolitanas, barrios o municipios que están creciendo gracias a los flujos de inmigración, puede pasar completamente desapercibido hasta que se renuevan las secciones censales de la EPA, y es posible que entonces ya se haya aminorado la afluencia o que los migrantes se dirijan hacia otros lugares.

Si la actualización de las secciones censales ya plantea inconvenientes, con la segunda unidad de muestreo de la EPA, esto es, con las *viviendas* seleccionadas en cada sección censal, también aparecen problemas. Es posible que no se pueda entrevistar a

(8) Ver INE (1991b:6) donde explícitamente se indica la posibilidad de una infravaloración del número de migrantes debida a la metodología.

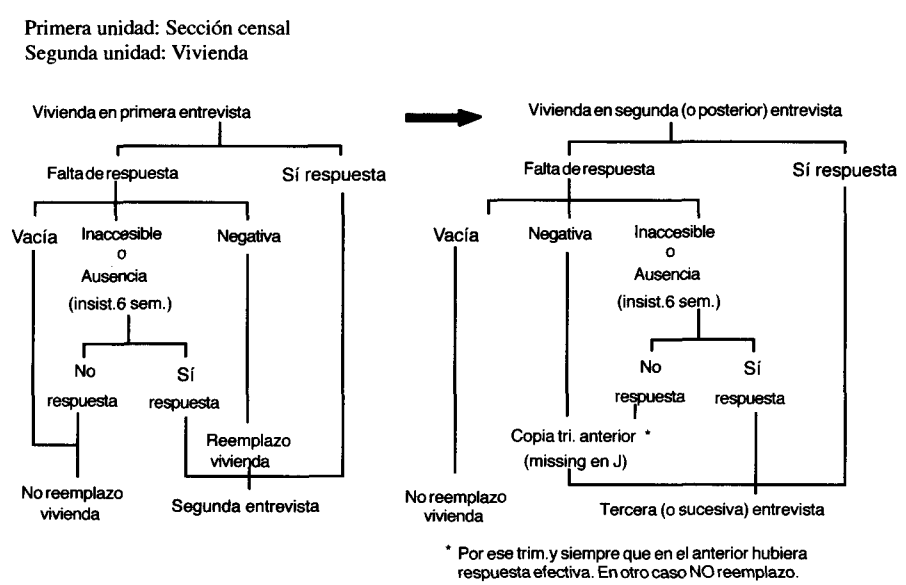
las personas que residen en la misma, pero incluso cuando ya se ha contactado con el grupo humano que habita la vivienda, un movimiento migratorio puede no ser cubierto en la EM debido a las definiciones e instrucciones que siguen los entrevistadores para la cumplimentación del cuestionario [ver INE (1993)]. En concreto, cuando hay una parte de las personas del grupo que no son encuestables como, entre otros, quienes llevan menos de tres meses en la vivienda y ocupan otra permanentemente –aunque con una clara perspectiva de abandonarla. Ese es un caso posible de un inmigrante reciente y adelantado, como el cabeza de familia, o de una parte de la misma. En casi todas estas situaciones no se puede captar, no se sigue o se pierde la “pista” a inmigrantes, incluso cuando sus viviendas han sido seleccionadas inicialmente en la muestra.

3.2. La pérdida de muestra, la falta de respuesta y la “copia”

Tanto al iniciar las entrevistas como en momentos posteriores, pueden presentarse diversas situaciones o incidencias que, a su vez, afectan a la estimación final de los movimientos migratorios. En general, nos referimos a la denominada falta de respuesta, cuyo tratamiento en la EPA se esquematiza en el cuadro 1, siguiendo INE (1993 y 1994).

A los efectos que nos interesan, cuando se analiza la conducta migratoria en los segundos trimestres y tratándose de la primera entrevista, las viviendas vacías y los grupos humanos ausentes de las viviendas, al no ser sustituidos suponen una pérdida de muestra que impide la posibilidad de captar adecuadamente el hecho migratorio al

Cuadro 1: EL MUESTREO BIETÁPICO DE LA EPA Y LOS MIGRANTES



aumentar el error de muestreo de los estimadores. Lo mismo sucede con todas las viviendas sin reemplazo en posterior entrevista y con la “copia” en el segundo trimestre –siempre que haya habido entrevista efectiva en el trimestre anterior– de los datos del grupo humano declarado ausente, inaccesible o con negativa. De todas estas situaciones, sólo disponemos de información estadística acerca de la última. Cuando se trabaja con los segundos trimestres y con las preguntas del apartado J de los ficheros anonimizados de la EPA –en el que se realizan las preguntas relativas a la movilidad y a la situación laboral hace un año–, se puede observar esa pérdida de muestra a través de los denominados registros *missing*. Disponiendo de tales ficheros para los mayores de 16 años se detecta fácilmente que ante la sencilla respuesta de si residía o no en otro lugar hace un año –la pregunta para captar a los migrantes–, siempre hay un grupo, en torno a 4 ó 5 mil personas, que no responden. Estas personas presentan un *missing* como respuesta y, cuando se les aplica el factor de elevación, equivalen en torno al millón de personas mayores de 16 años.

Tal y como se explica en INE (1994:36), los registros *missing* son el resultado del tratamiento de una parte de la falta de respuesta que se produce al realizar las encuestas. Concretamente, como se recogía en el cuadro 1, se trata de las negativas, inaccesibilidades y ausencias en segunda o posterior entrevista, cuyas encuestas se copian en el trimestre actual siempre que efectivamente la haya habido en el anterior. Pero así como los datos de actividad pueden copiarse del trimestre inmediatamente anterior, los de migraciones no están disponibles porque sólo se preguntan en los segundos trimestres y, por eso, aparece un *missing* como respuesta. En el cuadro 2 se recoge su número entre 1993 y 1996.

La justificación del INE para la “copia” es que de ese modo se evita el sesgo (en edad, sexo, etc.) que puede generar la desaparición coyuntural de estos registros en la muestra. Y es cierto que actuar así es adecuado para los períodos de vacaciones, como el tercer trimestre, tras los cuales la gente vuelve a sus casas, pero quizá no lo sea siempre. El problema es que aunque se “reencuentran” algunos *missing* en la siguiente entrevista⁹, como las preguntas de migraciones sólo se hacen en los segundos trimestres, es improbable que pueda tomarse esa observación un año más tarde, en el próximo segundo trimestre. Esto no sería muy grave en las estimaciones finales de la EPA siempre que ningún *missing* hubiera sido migrante, pero si los hay evidentemente la cosa cambia.

3.3. La pregunta en la EPA para captar el hecho migratorio y los turnos de rotación

Desde 1987, es el apartado J del cuestionario de la EPA el que se añade en los segundos trimestres de cada año y con cuyas respuestas se recoge la información sobre los migrantes, al preguntar a los entrevistados por su lugar de residencia y situación

(9) Según se nos ha indicado desde el INE, aunque no han estudiado el tema con profundidad, son pocos los *missing* que pueden recuperarse en posteriores entrevistas. Esta es una cuestión pendiente de investigación pues, del mismo modo que con los migrantes, si hay transiciones significativas en el mercado de trabajo entre los *missing* se verían alteradas, en mayor o menor medida, las correspondientes estimaciones de la EPA.

Cuadro 2: ¿CUÁL ERA SU LUGAR DE RESIDENCIA HACE UN AÑO?
NÚMERO DE PERSONAS (%)

AÑO	1993	1994	1995	1996
EL MISMO (No missing no inmigrantes)	147.801 (96%)	149.651 (96,2%)	152.935 (96,7%)	152.328 (96,5%)
DISTINTO (No missing inmigrantes)	780 (0,5%)	769 (0,5%)	813 (0,5%)	743(0,5%)
MISSING	5.312 (3,5%)	5.073 (3,3%)	4.434 (2,8%)	4.662 (3%)
TOTAL	153.893	155.493	158.182	157.733

Fuente: Ficheros anonimizados 2.º trimestres EPA (INE).

laboral exactamente hace un año. Por tanto, el fenómeno migratorio se analiza *ex-post*, con forma de inmigración y una vez realizado el movimiento.

Este tipo de pregunta retrospectiva con limitación temporal a cuatro trimestres plantea, al menos, tres problemas. En primer lugar, la respuesta es dependiente de la *memoria histórica* de los entrevistados. Aunque menos grave que en los Censos o Padrones, es posible que se produzcan inexactitudes que enturbian la información en relación al momento de la inmigración, sobre todo, cuando la contestación depende eventualmente de un sólo informante, el correspondiente a cada vivienda, y de un sólo momento, el segundo trimestre de cada año. En segundo lugar, es obvio que, aún gozando de una buena memoria, el entrevistado sólo puede responder a la pregunta que se le hace. Y esta no es *cuándo emigró por última vez*, sino *dónde estaba exactamente hace un año*. El resultado es que se eliminan de golpe todas las migraciones fuera del límite temporal de los doce meses, así como las migraciones intermedias. Y, en tercer lugar, la existencia de los turnos de rotación en la muestra EPA impide que la pregunta que se realiza a los entrevistados cubra correctamente a los inmigrantes. Es decir, parte de la muestra, aunque haya inmigrado, nunca podrá contestar afirmativamente. Veamos con un ejemplo cómo se produce la subestimación en relación con el último punto.

Las viviendas que forman parte de la muestra de la EPA en un momento concreto llevan en la misma distintos períodos de tiempo. La rotación de la muestra va renovando un sexto de la misma en cada trimestre, siendo el tiempo máximo de permanencia de cada submuestra seis trimestres. Así, en cualquier segundo trimestre una sexta parte será entrevistada por última vez (las viviendas que, con éste, ya llevan seis trimestres), otro sexto de las viviendas hará cinco trimestres que entró en la muestra por primera vez y así sucesivamente. Esto es, cuando se hace la pregunta retrospectiva en el segundo trimestre de cualquier año (t), la muestra de viviendas que va a ser entrevistada estará distribuida en submuestras con un determinado número de inmigrantes cada una de ellas del siguiente modo:

Cuadro 3: DISTRIBUCIÓN DE LA MUESTRA EPA SEGÚN PERÍODOS DE PERMANENCIA.
AÑO T (SEGUNDO TRIMESTRE)

Muestra completa EPA	Submuestra	Grupo migrantes en submuestra
* 1/6 lleva seis trimestres (incl. actual)	A	a
* 1/6 lleva cinco trimestres (incl. actual)	B	b
* 1/6 lleva cuatro trimestres (incl. actual)	C	c
* 1/6 lleva tres trimestres (incl. actual)	D	d
* 1/6 lleva dos trimestres (incl. actual)	E	e
* 1/6 lleva un trimestre (actual)	F	f

Antes de avanzar con el ejemplo y con el fin de simplificar, utilizaremos los siguientes supuestos. Primero, que todas las entrevistas se hacen en un momento concreto, por ejemplo, en el último día del segundo trimestre. Segundo, que en cada vivienda el grupo humano está formado por una sola persona, a la que encontramos y es encuestable. Tercero, que estas personas no realizan ningún movimiento migratorio desde el momento en que son captadas en la encuesta y hasta la finalización de sus seis trimestres de permanencia. Con el supuesto primero se uniformiza el tratamiento de los períodos de referencia, mientras que los supuestos segundo y tercero sirven para soslayar los problemas ya enunciados de no encuestabilidad e incidencias. Concretamente, con el segundo se simplifica la falta de respuesta y el tratamiento del grupo humano, y con el tercero se retiene a quienes han tenido un episodio migratorio hasta el momento de la entrevista en el segundo trimestre. Ninguno de estos tres supuestos tiene por qué cumplirse en la realidad, pero controlan la pérdida de información muestral.

Si los tres supuestos anteriores se refieren a la participación de un individuo en la muestra EPA, los dos siguientes controlan el tipo de movilidad. Así, en cuarto lugar, los movimientos múltiples (varias migraciones previas realizadas por la misma persona) se simplificarán imponiendo que cada inmigrante entrevistado sólo haya efectuado una migración antes de ser preguntado. Y, finalmente, con el quinto y último supuesto se restringe temporalmente la posibilidad de movimiento *ex-ante* a los cuatro últimos trimestres hacia atrás contados desde el momento en que la persona es encuestada por primera vez. Con ambos supuestos, se puede repartir a los inmigrantes de cada submuestra (por ejemplo, c) en cuatro subgrupos (c_1, c_2, c_3 y c_4), según el trimestre en el que hicieron el movimiento.

Por tanto, si se tuviera conocimiento *a priori* de todos los inmigrantes de la muestra del año t (2 tr.), éstos se distribuirían como en el cuadro 4.

Así, si en el último día del segundo trimestre del año t se pregunta por el lugar de residencia hace exactamente un año, quedarían fuera del cómputo como inmigrantes, $M_{nc}^{(2)}$, la suma de:

$$M_{nc}^{(2)} = \sum_{i=1}^4 a_i + \sum_{i=1}^4 b_i + \sum_{i=2}^4 c_i + \sum_{i=3}^4 d_i + e_4 \quad [2.1]$$

Cuadro 4: DISTRIBUCIÓN DE LOS EFECTIVOS MIGRANTES DE LA MUESTRA ENTREVISTADA EN EL AÑO T (2.º TR.). SUBGRUPOS DE MIGRANTES

		Año y trimestre en el que se movieron								
		t-2(2)	t-2(3)	t-2(4)	t-1(1)	t-1(2)	t-1(3)	t-1(4)	t(1)	t(2)
Año y trimestre	t-1(1)	a ₄	a ₃	a ₂	a ₁					
en el que	t-1(2)		b ₄	b ₃	b ₂	b ₁				
entraron	t-1(3)			c ₄	c ₃	c ₂	c ₁			
en la	t-1(4)				d ₄	d ₃	d ₂	d ₁		
muestra	t(1)					e ₄	e ₃	e ₂	e ₁	
	t(2)						f ₄	f ₃	f ₂	f ₁

y como inmigrantes constarían $M_c^{(2)}$:

$$M_c^{(2)} = c_1 + \sum_{i=1}^2 d_i + \sum_{i=1}^3 e_i + \sum_{i=1}^4 f_i \quad [2.2]$$

habiendo sido el total de inmigrantes en la muestra $M^{(2)}$:

$$M^{(2)} = M_{nc}^{(2)} + M_c^{(2)} \quad [2.3]$$

La subestimación no sería tan grande en realidad, pues no hay que olvidar que parte de los efectivos no cubiertos, $M_{nc}^{(2)}$, fueron entrevistados en el año anterior (2 tr.) y contabilizados como inmigrantes. Para averiguar exactamente qué parte es esa, basta cruzar el año t-1(2) con el trimestre de entrada en la muestra de las diferentes submuestras. En t-1(2) formaban ya parte de la muestra entonces entrevistada las submuestras A y B, siendo cubiertos, concretamente, los subgrupos de inmigrantes a₁, a₂, a₃ y el subgrupo b completo, pero no el a₄.

Si bien parte de los migrantes de las submuestras A y B fueron captados en t-1(2), no sucede lo mismo con los de C, D y E. Los subgrupos c₂, c₃, c₄, d₃ y d₄ ya no estarán en la muestra en t+1(2) y el e₄, aunque todavía permanecerá el próximo año, hará más de cuatro trimestres desde que realizó su migración y, por tanto, quedará fuera del ámbito temporal de la pregunta. De este modo, de los veinticuatro subgrupos de inmigrantes inicialmente planteados, a más de la cuarta parte (siete) no se les capta nunca como migrantes.

Como se ha visto, las submuestras más afectadas por la pérdida de información en t(2) son las C y D, las que llevan, respectivamente, cuatro y tres trimestres siendo encuestadas. De una forma muy simple, puede estimarse lo que esto supondría en términos de migrantes para un año concreto. Por ejemplo, si a estas submuestras en el segundo trimestre se les preguntara por su residencia hace año y medio en lugar de hace exactamente un año, se cubriría a más de la mitad de los siete subgrupos antes

perdidos, concretamente, a los inmigrantes en c_2 , c_3 , d_3 y d_4 . Así, únicamente quedarían sin recoger los inmigrantes de los subgrupos a_4 , c_4 y e_4 . Para 1995/96, período en el que según la EM se registran 168.900 migrantes interiores, atendiendo al *cuadro 4*, podremos observar que este valor correspondería a los diez subgrupos captados por la pregunta actualmente formulada. Si distribuimos esos 168.900 migrantes proporcionalmente entre los diez subgrupos y añadimos los recogidos en c_2 , c_3 , d_3 y d_4 , entonces su número se podría situar en torno a 236.460 (168.900 ya recogidos, más 16.890 por cada grupo adicional). Esto significa un incremento de un 40 por cien en el número de migrantes.

3.4. El cambio de metodología en la EM

El cálculo anterior tiene interés porque, curiosamente, esa sería una cifra ya más parecida a las de los años entre 1980 y 1985. En nuestra opinión, y entre otras razones, seguramente parte de la *dramática reducción* en el número de migrantes del período 1980-1985 al 1986-1996, responde al cambio metodológico que produjo una pérdida de cobertura importante. Como se ha señalado, el procedimiento para captar a los migrantes en la EM entre 1980 y 1985 consistía en hacer la pregunta retrospectiva sólo a la población correspondiente a la parte de la muestra de viviendas EPA que entraba en la encuesta por primera vez en cada uno de los cuatro trimestres del año, aproximadamente unas 40.000 viviendas en total. Conocido el número de inmigrantes para cada nueva submuestra, se sumaban los cuatro trimestres y se estimaba el valor anual tomando el adecuado factor de ponderación, en el que debía tenerse en cuenta que tal muestra estaba referida sólo a unas 40.000 viviendas¹⁰. Preguntando de ese modo, la encuesta podía cubrir a todos los inmigrantes que se habían movido durante los cuatro últimos trimestres y pertenecían a las cuatro nuevas partes entrantes en la muestra. La pregunta inmediata es si la cobertura era mejor entonces y si, por eso, ahora las cifras podrían ser menores.

Podemos responder a esta cuestión planteando un nuevo ejemplo en el que ajustamos el período de referencia del primer procedimiento al actualmente empleado (esto es, preguntando a las cuatro submuestras con entrada en los trimestres tercero y cuarto del año $t-1$, y en los trimestres primero y segundo del año t , en lugar del año natural), estimamos de ambas maneras y comparamos. Si con el procedimiento antiguo (P^I) se entrevistaba a 40.000 personas (viviendas), con el nuevo (P^{II}) se entrevista en torno a 60.000. Si la población total a estudiar es de, por ejemplo, diez millones, entonces, de acuerdo con el tamaño de la muestra utilizada en cada procedimiento, los registros representarán diferente número de personas. Así, en el P^I un registro equivale a 250 efectivos, mientras que en el P^{II} representa casi a 167 personas, es decir 1,5 veces menos. Estos serán los factores de elevación de cada registro al total de la población, respectivamente denominados f_e^I y f_e^{II} .

Para el año $t-1(3)$ a $t(2)$, las submuestras encuestadas por el P^I serían C, D, E y F, y recurriendo al *cuadro 4* se aprecia que con la pregunta retrospectiva se cubriría a todos los subgrupos de inmigrantes de esas cuatro submuestras. Suponiendo que todos

(10) Hay que hacer notar que no existe información publicada sobre la metodología utilizada a lo largo de esta etapa. Hemos podido reconstruir el método partiendo de las breves notas que figuran en las introducciones de la EM y con la ayuda del personal del INE.

ellos se distribuyeran proporcionalmente (esto es, que $c_i = d_i = e_i = f_i = \alpha$, para todo i), entonces la cifra estimada del total de inmigrantes $M(P^I)$ será:

$$M(P^I) = 16 \alpha f_e^I \quad [2.4]$$

Para el mismo año, con el P^{II} la cifra estimada del total de migrantes, $M(P^{II})$, será la suma de los diez subgrupos de inmigrantes cubiertos pertenecientes a las seis submuestras multiplicada por el factor de elevación correspondiente, esto es:

$$M(P^{II}) = (c_1 + \sum_{i=1}^2 d_i + \sum_{i=1}^3 e_i + \sum_{i=1}^4 f_i) f_e^{II} = 10 \alpha f_e^{II} \quad [2.5]$$

Siendo $f_e^I = 1,5 f_e^{II}$ y sustituyendo en [2.4]:

$$M(P^I) = 16 \alpha 1,5 f_e^{II} = 24 \alpha f_e^{II} \quad [2.6]$$

donde obviamente, $M(P^I)$ es mayor que $M(P^{II})$. Esto es, los inmigrantes captados con el procedimiento anterior son bastantes más que los cubiertos con el actual.

Parece claro que aquí habría una respuesta al menos a parte del por qué de esa *dramática caída* de las migraciones interiores que muestra la EM desde que se utiliza el nuevo procedimiento de estimación y, justamente, cuando se entra en una fase de crecimiento económico y de creación de empleo como la que se inicia en España en 1985. De hecho, si continuamos con el ejemplo asignando un valor a α , por ejemplo 97 migrantes, a través de P^I se habría estimado casi 389.000 migrantes, mientras que con el P^{II} la cifra estimada hubiera sido algo inferior a 162.000. Es decir, con el nuevo procedimiento sólo se captaría en torno al 40 por cien de los migrantes cubiertos con el viejo proceder. Por tanto, el procedimiento actualmente utilizado en la EPA para la cobertura de las migraciones presenta una muy inferior capacidad para capturar la intensidad de los flujos.

4. APROXIMACIÓN EMPÍRICA A LA SUBESTIMACIÓN DE LA EM

Ya con datos reales, esto es, con los ficheros anonimizados de la EM, puede hacerse una doble aproximación empírica a este problema demostrando, por un lado, cómo disminuyen las tasas migratorias por submuestras a medida que aumenta la permanencia y, por otro, estimando el número de migrantes a través de la tasa migratoria submuestral con mayor cobertura.

Una vez identificados los turnos (submuestras) a partir de la información facilitada por el INE, se puede comprobar en el cuadro 5 que a medida que aumenta la permanencia y el número de entrevistas, la correspondiente tasa migratoria disminuye¹¹.

Como ya se ha explicado, esto se produce porque a mayor número de trimestres siendo entrevistados cada vez es menor el número de subgrupos que se cubre con las preguntas del apartado J. A medida que las submuestras agotan sus cinco y seis trimestres de permanencia, su tasa migratoria se aproxima a cero, pues nadie que desde enton-

(11) Hay que aclarar que es muy difícil que en la realidad se manifieste estrictamente el esquema teórico propuesto en el ejemplo pues los supuestos acerca del momento de la entrevista –último día del 2.º trimestre–, de que una vivienda sólo es ocupada por una persona y, en general, de que no hay cambios en el grupo humano, lógicamente se rompen al ser efectivamente realizada la encuesta.

**Cuadro 5: TASAS MIGRATORIAS (%) POR TURNOS DE ENTREVISTA
(1993-1996, 2.º TR.)**
(efectivos de 16 años y más, excluidas observaciones missing)

	Turno de entrevista (submuestra):					
	1.º (F)	2.º (E)	3.º (D)	4.º (C)	5.º (B)	6.º (A)
1993	0,85	0,69	0,51	0,40	0,32	0,33
1994	0,82	0,59	0,45	0,45	0,34	0,39
1995	0,97	0,54	0,54	0,39	0,35	0,33
1996	0,95	0,52	0,42	0,34	0,33	0,33

Fuente: INE y elaboración propia.

ces haya sido entrevistado continuamente ha podido realizar su inmigración en los cuatro últimos trimestres. Por el contrario, la mayor tasa migratoria la presenta la submuestra que acaba de entrar, la entrevistada por primera vez en el segundo trimestre.

El cuadro 6 tiene por objeto estimar el número de inmigrantes a partir de los entrevistados por primera vez en el segundo trimestre –la submuestra del tipo F en el ejemplo– y, a partir de ahí, ofrecer un valor provisional para la población total¹². La ventaja de operar con esta submuestra es doble: se cubre a todos los que han inmigrado a lo largo de los cuatro últimos trimestres y, al no haber sido anteriormente encuestada, no presenta copias de registros ni, por tanto, el problema de los *missing*. Así, tomando los inmigrantes del primer turno en el segundo trimestre, se obtienen los resultados del cuadro 6.

Cuadro 6: APROXIMACIÓN AL NÚMERO DE MIGRANTES A TRAVÉS DE LA SUBMUESTRA ENTREVISTADA POR PRIMERA VEZ EN EL 2.º TRIMESTRE (1993-1996).
(efectivos de 16 años y más)

	Muestra			Estimación para la Población total (1)		
	Migrantes muestra	Tamaño Muestral	Tasa migrat. muestral	Migrantes población	Tamaño población	Tasa migrat. población
AÑO 1993						
EPA	780	153.893	0,51%	139.494	31.238.244	0,45%
Subm. 2T	225	26.333	0,85%	246.236	32.135.239	0,77%
				(incr. 76%)		

(12) Es ésta una posibilidad sugerida en el marco del seminario de investigación realizado en el Departamento de Fundamentos de la Economía e Historia Económica de la Universidad de Alcalá.

Cuadro 6: APROXIMACIÓN AL NÚMERO DE MIGRANTES A TRAVÉS DE LA SUBMUESTRA ENTREVISTADA POR PRIMERA VEZ EN EL 2.º TRIMESTRE (1993-1996) (continuación)
(efectivos de 16 años y más)

	Muestra			Estimación para la Población total (1)		
	Migrantes muestra	Tamaño Muestra	Tasa migrat. muestral	Migrantes población	Tamaño población	Tasa migrat. población
AÑO 1994						
EPA	769	155.493	0,49%	144.640	31.531.100	0,46%
Subm. 2T	217	26.441	0,82%	233.040 (incr. 61%)	32.300.742	0,72%
AÑO 1995						
EPA	813	158.282	0,51%	147.073	31.841.229	0,46%
Subm. 2T	264	27.175	0,97%	295.110 (incr. 101%)	31.912.668	0,92%
AÑO 1996						
EPA	743	157.733	0,47%	137.855	32.095.153	0,43%
Subm. 2T	252	26.403	0,95%	277.136 (incr. 101%)	32.784.424	0,85%

(1) factor de elevación submuestra 2T multiplicado por seis.

Fuente: INE y elaboración propia.

Aunque se trata de un cálculo sencillo, resulta muy ilustrativo. Puede observarse que las estimaciones finales de los flujos migratorios cambian sensiblemente, apreciándose siempre aumentos no inferiores al 60 por ciento respecto de los flujos estimados y publicados por la EM. No debe sorprender este resultado, ni el de que la tasa migratoria se doble en valor para las estimaciones de la población total, pues vienen a confirmar lo ya explicado. En concreto, que la EM actualmente subestima el número de inmigrantes ya que sólo capta de forma completa las migraciones realizadas por el sexto de la muestra de reciente entrada.

5. CONCLUSIONES Y SUGERENCIAS

El objetivo de este trabajo no ha sido otro más que el análisis en profundidad de la información básica sobre migraciones en España, a raíz de las sorprendentes diferencias que muestran las principales fuentes de información continua: la EVR y la EM. Estas diferencias cuestionan la reciente, y comúnmente aceptada, baja movilidad interior que se desprende de la última fuente.

El estudio pormenorizado de la metodología que actualmente se utiliza en la EPA para la estimación de los migrantes, nos ha permitido establecer la existencia de un serio problema de subestimación de la movilidad cuando se mide a través de esta estadística debido, sobre todo, a la formulación de la propia pregunta acerca del hecho migratorio combinada con los turnos de rotación de las viviendas en la muestra EPA.

De hecho, actualmente al ser encuestada la población de la muestra acerca de su conducta migratoria una sólo vez cada año, buena parte de los efectivos que ciertamente han experimentado movilidad no puede contestar afirmativamente debido a un problema de desajuste temporal. Y el resultado es un bajo número de inmigrantes estimados a través de la EPA. La corrección de este problema de “tiempos” podría venir de la mano de dos alternativas. Bien estimando para la población total a partir de la submuestra entrante en los segundos trimestres, teniendo en cuenta que habría que resolver los problemas de no representatividad de un sexto de la muestra. O bien, preguntando a todas y cada una de las submuestras entrantes en cada uno de los cuatro trimestres del año, como se hacía entre 1980 y 1985; aprovechando, en este último caso, para adaptarse a las futuras directrices del EUROSTAT en relación con la EFT comunitaria, pero procurando evitar la sobreestimación que se produciría si se estimara a partir de los datos de toda la muestra encuestada en todos y cada uno de los cuatro trimestres.

En cualquier caso, aun corrigiendo este problema, el número de inmigrantes captados por la EPA se mantendría todavía en niveles muy alejados en relación con las cifras de la EVR. Quiere esto decir que la subestimación procedente de la elección de las secciones censales –que en lugar de ser fijas podrían rotar, como sucede con su equivalente en otros países europeos–, la definición de migrante de la EPA, las instrucciones seguidas por los entrevistadores, la pérdida de muestra y la falta de respuesta no debe ser en absoluto despreciable. El problema es que actualmente no se puede cuantificar, lo que nos lleva a concluir que en los estudios macroeconómicos con dimensión temporal en los que se estima la intensidad de los flujos no es aconsejable el uso de esta fuente.

Por último, si se acepta esta infravaloración *congénita* de la EM, la cuestión que quedaría pendiente sería valorar su actual potencia para describir correctamente las características cualitativas e individuales de los migrantes; en particular, por su interés en cuanto a las referentes a su situación en el mercado de trabajo que son, precisamente, las que no se recogen en la EVR. Esta cuestión es pasada por alto en los estudios que utilizan la EM. El problema es que los mismos no pueden garantizar que sus estimaciones –y, en definitiva, sus conclusiones– estén libres de sesgos de importancia y de una elevada dosis de variabilidad según los años estudiados, pues trabajan con una submuestra que posiblemente no represente el universo real de migrantes. De hecho, si comparamos las características que simultáneamente recogen ambas fuentes aparecen notables diferencias en las distribuciones porcentuales según la edad, el nivel de educación y las regiones de origen y destino¹³. Estas diferencias y su sensibilidad según el año elegido, pensamos que deberían tenerse en cuenta al realizar las aproximaciones microeconómicas al objeto de estudio pues cuestionan nuevamente esta fuente estadística. En todo caso, los cambios que previsiblemente se incorporarán en

(13) Para el período entre 1993 y 1995, la EM subestima sistemáticamente los inmigrantes en edad de trabajar –en torno a un 6%–, entre 25 y 44 años de edad; por el contrario, la EM sobreestima los inmigrantes de más de 65 años atribuyéndoles más del doble de su participación en la EVR. En cuanto al nivel educativo –a pesar de las dificultades para la comparación–, la EM subestima todos los años el número de migrantes con grado medio y superior, mientras que sobreestima a quienes sólo alcanzaron el nivel de graduado escolar. Finalmente, el mapa geográfico de las migraciones dibujado por la EM difiere notablemente del reflejado por la EVR. En particular Madrid, la macrorregión Sur –Andalucía, Murcia, Ceuta y Melilla– y la macrorregión Este –Cataluña, Valencia y Baleares– son las peor captadas tanto como región de origen como de destino –en un rango de diferencias entre un 4 y un 12%–, y con una alta variabilidad según el año analizado.

la EPA a partir de 1999 –procedentes de EUROSTAT, según informa el INE–, podrían modificar esta valoración mejorando los resultados de la comparación. Si no fuera así, habría que buscar una alternativa. Una propuesta sencilla consistiría en indagar nuevamente a través de la EVR sobre la situación laboral actual y anterior de los inmigrantes, como se hacía con anterioridad a 1987, con el objeto de completar la información estadística de tipo cualitativo sobre migraciones.



REFERENCIAS BIBLIOGRÁFICAS

- Albérico Gil, L. y Jimeno, J.F. (1993): "The determinants of labor mobility in Spain: Who are the migrants?", FEDEA, *Documento de trabajo*, n.º 93-05.
- Antolín, P. y Bover, O. (1997): "Regional migration in Spain: the effect of personal characteristics and of unemployment, wages and house price differentials using pooled cross-sections", *Oxford Bulletin of Economics and Statistics*, n.º 59, págs. 215-235.
- Bentolila, S. (1997a): "Sticky labor in Spanish regions", *European Economic Review*, n.º 41, págs. 591-598.
- Bentolila, S. (1997b): "La inmovilidad del trabajo en las regiones españolas", *Papeles de economía española*, n.º 72, págs. 168-175.
- Bentolila, S. y Dolado, J.J. (1990): "Mismatch and Internal Migration in Spain, 1962-86" en F. Padoa-Schioppa (ed.): *Mismatch and Labour Mobility*, Cambridge Univ. Press, págs. 182-236.
- INE (varios años): *Anuario estadístico*.
- INE (varios años): *Migraciones*.
- INE (1991a): *Encuesta Sociodemográfica. 1991*, Tomo II Resultados Nacionales, vol. (2). Movimientos migratorios y vivienda.
- INE (1991b): *Encuesta de Migraciones 1989*. EPA.
- INE (1993): EPA. *Descripción de la Encuesta, definiciones e instrucciones para la cumplimentación del cuestionario* (ed. revisada).
- INE (1994): EPA. *Tratamiento de la información*.
- INE (1996a): EPA. *Encuesta de migraciones. 1995*.
- INE (1996b): EPA. *Estadística de flujos. Metodología*.
- Olano, A. (1990): "Las migraciones interiores en fase de dispersión", *Revista de economía y sociología del trabajo*, n.º 8-9, págs. 86-97.
- Ródenas, C. (1994): "Migraciones interregionales en España, 1960-1989", *Revista de Economía Aplicada*, vol. II, n.º 4, págs. 5-36.

Fecha de recepción del original: febrero, 1998

Versión final: mayo, 1998

ABSTRACT

The Encuesta de Migraciones (Migration Survey) taken from the Encuesta de Población Activa (Spanish Labour Force Survey) has shown a low and decreasing intensity of migration flows in Spain over the past ten years, even when the Spanish economy was growing and creating employment. This surprise scenario is even more so when we compare these data with other related statistical information. Our analysis of the methodology used by the Instituto Nacional de Estadística (Spanish National Statistical Institute) has enabled us to establish the existence of an important and significant problem whereby mobility is underestimated by the said statistics.

Keywords: migration, Spanish Labour Force Survey and methodology.