

Migrantes y migraciones: de nuevo la divergencia en las fuentes estadísticas(*)

por

MÓNICA MARTÍ Y CARMEN RÓDENAS

Departamento Análisis Económico Aplicado. Universidad de Alicante

RESUMEN

Este trabajo plantea una serie de deficiencias presentes en la *Encuesta de Migraciones* (EM), publicada anualmente por el INE a partir de la EPA. Consciente de la subestimación de la movilidad procedente de esta fuente estadística, el INE ha introducido recientemente ciertos cambios metodológicos. Sin embargo, no se han obtenido mejoras significativas y la misma continúa disponible para los investigadores.

Las razones que se encuentran detrás de la limitada calidad de la EM son de tipo estadístico y afectan a la precisión y al sesgo de las estimaciones. Se desatienden principios básicos de la estimación para dominios pequeños, se incumplen los supuestos del muestreo por conglomerados y se incurre en graves sesgos debidos, probablemente, a la falta de respuesta.

Consideramos que la investigación del hecho migratorio a través de la técnica del muestreo estratificado de poblaciones exige un tra-

(*) Las autoras queremos expresar nuestro agradecimiento a M. Trottni, E. Lazcorreta y A. León por sus valiosos comentarios, y a tres evaluadores anónimos cuyas sugerencias han contribuido a mejorar la versión final de este trabajo.

tamiento específico en la encuesta, dado el tamaño y las características de este dominio.

Palabras clave: Estimación de Migraciones, Encuesta de Población Activa, Falta de Respuesta y Estimación de Dominios Pequeños.

Clasificación AMS: 62P20, 91B82, 62D05.

1. INTRODUCCIÓN

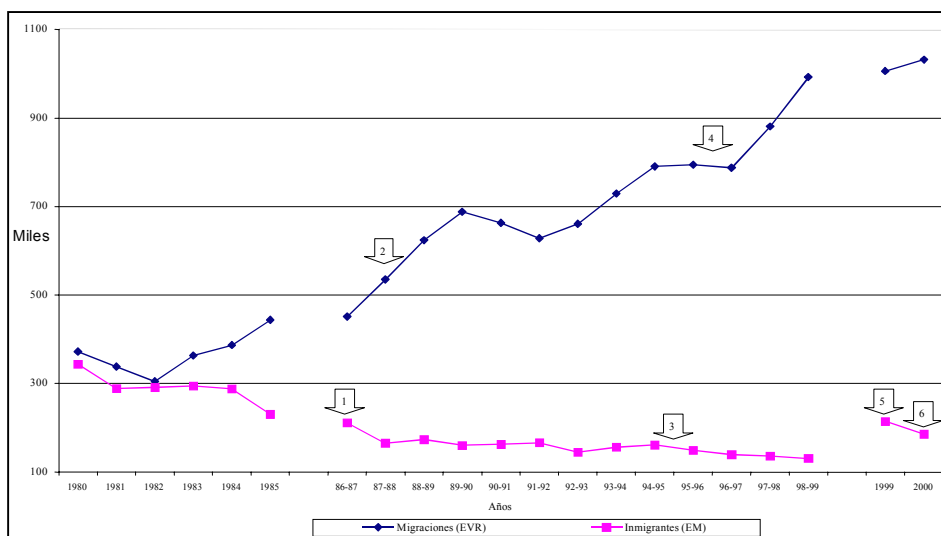
La información estadística oficial sobre los movimientos migratorios en España es elaborada con periodicidad anual por el Instituto Nacional de Estadística (INE). Este organismo ofrece dos medidas del fenómeno migratorio a partir de fuentes distintas. Por un lado, realiza una explotación específica de la Encuesta de Población Activa (EPA), conocida como la *Encuesta de Migraciones* (EM), que informa sobre el volumen de población inmigrante, así como de sus características personales, su relación con la actividad y, entre otras, del tipo de movimiento. Por otro lado, basándose en los datos de la Estadística de Variaciones Residenciales (EVR), ha publicado anualmente la serie *Migraciones* –que desde 1999 se denomina *Estadística de Variaciones Residenciales*– en la que recoge información acerca de los movimientos de altas y bajas registrados en los Padrones Municipales. Aunque ambas fuentes proporcionan información sobre el mismo fenómeno migratorio -entendido siempre como cambio de municipio de residencia-, los resultados ofrecidos por cada una de ellas se diferencian tanto en su metodología como en su contenido. Así, mientras que la primera ofrece una estimación del hecho migratorio basándose en el muestreo de la población, la información de la segunda procede de un registro administrativo. Además, mientras que la EM mide personas, esto es, *inmigrantes* y es la única en la actualidad que informa sobre sus situaciones laborales, la EVR mide movimientos, es decir, *migraciones*.

En la figura 1, se ha representado la movilidad interior captada por cada una de estas dos fuentes de información⁽¹⁾ desde el año 1980, cuando comienza a publicarse los datos de la EM, hasta el año 2000 (último dato disponible). Dado que la metodología aplicada es diferente y que, además, no miden exactamente lo mismo, cabe esperar ciertas diferencias en sus cifras en la medida en que, por ejemplo, un inmigrante puede realizar varios cambios de municipio de residencia en un año. Sin embargo, no deja de llamar la atención la formidable distancia que se observa entre las dos series a partir del año 1982. Una diferencia que, además, no se ha logrado

(1) Para comparar adecuadamente se han realizado algunos ajustes, cuyo detalle se explica más adelante.

reducir a pesar de la introducción de posteriores modificaciones metodológicas en ambas fuentes.

Figura 1
INMIGRANTES Y MIGRACIONES INTERIORES, 1980-2000



Nota: Cada flecha representa un cambio metodológico. En particular:

(1) EM, 1987: El apartado que permite estimar los inmigrantes, *situación hace un año*, se cumplimenta sólo en los segundos trimestres.

(2) EVR, 1988: Informatización del proceso de gestión estadística.

(3) EM, 1995: Renovación paulatina del seccionado de la muestra.

(4) EVR, 1996: Padrón Continuo de Habitantes.

(5) EM, 1999: Nuevo método de estimación de inmigrantes; ampliación del período de trabajo de campo y de la muestra de viviendas; mejoras en el tratamiento de la falta de respuesta.

(6) EM, 2000: Nueva actualización de la muestra de secciones.

Fuente: INE, *Migraciones, Encuesta de Migraciones* (varios años) y elaboración propia.

El objeto de este trabajo –continuación de Ródenas y Martí (1997)– es explicar esta diferencia y su resistencia a los últimos cambios metodológicos. Así, en el primer epígrafe se describen las sucesivas mejoras en el tratamiento de la información de la EVR. Los cambios en el diseño de la EPA y en la forma de medir el flujo migratorio en la *Encuesta de Migraciones* son examinados en el segundo epígrafe. A continuación, tras comparar los nuevos resultados de ambas fuentes y comprobar que la EM continúa muy alejada de la EVR, en el tercer epígrafe se discute si realmente es posible estimar, con un nivel de precisión tolerable y sin sesgos, una subpoblación como los inmigrantes a través de una muestra diseñada y dimensionada para obtener resultados generales y dar información de las principales categorías poblacionales en relación con el mercado de trabajo, como es el caso de la

EPA. El trabajo termina con un epígrafe en el que, además de recoger las principales conclusiones, se plantean algunas sugerencias para mejorar la información estadística sobre las migraciones.

2. LOS CAMBIOS EN LA ESTADÍSTICA DE VARIACIONES RESIDENCIALES

Los datos sobre las migraciones se obtienen de la Estadística de Variaciones Residenciales que elabora el INE a partir de la explotación de la información generada por la gestión de los cambios de residencia en los Padrones Municipales de Habitantes. Concretamente, la EVR proporciona información mensual sobre las migraciones clasificadas por origen y destino, así como su distribución en función de algunas variables personales como sexo, edad, lugar de nacimiento y titulación académica. Al tratarse de un registro administrativo, la veracidad de la información facilitada va a depender del grado de cumplimiento entre los nuevos vecinos del requisito legal de empadronamiento, y la calidad de la misma de los medios y la diligencia municipal en el tratamiento de la información.

Legalmente existe la obligación de empadronamiento y éste es un requisito para poder acceder a muchos servicios sociales o, simplemente, para ejercer determinados derechos democráticos como el voto. Sin embargo, su incumplimiento no impide el desarrollo de la vida de los ciudadanos en sus nuevos municipios de residencia, por lo que es bastante probable que muchos traslados queden sin registrar por sus protagonistas. La EVR presenta en este sentido una información con un cierto sesgo a la baja. No obstante, esta desviación se está reduciendo con el paso del tiempo en la medida en que se exige con mayor frecuencia a los ciudadanos la presentación del certificado de empadronamiento para acceder a los diferentes servicios sociales o municipales. De hecho, como señala el INE en la serie *Migraciones*, en los últimos años ha mejorado la cobertura de esta fuente, pues cada vez es más habitual requerir la acreditación de la residencia para poder realizar ciertas actividades u obtener ciertos beneficios.

En cuanto a la calidad en el tratamiento de la información estadística de la EVR, conviene destacar fundamentalmente dos mejoras, introducidas en 1988 y en 1996. La primera consistió en la informatización del proceso de gestión estadística. Entre 1988 y 1996, la información recogida por los ayuntamientos en un único documento de alta/baja era grabada en las Delegaciones Provinciales del INE, existiendo para cada persona que realizaba una migración un registro individualizado. Este cambio debió contribuir notablemente a mejorar la actualización y el tratamiento de la información. Sin embargo, el sistema seguía adoleciendo de un problema puntual en los años en los que se elaboraban los nuevos Padrones Municipales de Habitantes, pues ofrecía un valor para ese año anormalmente bajo. Como se explica en

Ródenas y Martí (1997:158), esta anomalía se producía porque durante los meses en los que se estaba renovando el Padrón, los nuevos residentes eran dados de alta directamente en el Padrón Municipal de su actual destino y, simultáneamente, de baja en el correspondiente Padrón de su municipio de origen y, en consecuencia, no quedaban registrados en las cifras mensuales de variaciones residenciales(2). Este inconveniente afortunadamente ha desaparecido desde 1996, cuando se ha puesto fin a una larga historia de renovaciones quinquenales del Padrón. Las renovaciones padronales se sustituyen ahora por un sistema de gestión continua del Padrón, posible gracias al desarrollo del tratamiento informático de grandes ficheros. El Padrón Continuo de Habitantes proporciona anualmente las cifras oficiales de la población española y la EVR se elabora directamente a partir de los cambios en el registro padronal.

Los cambios descritos han supuesto una mejora sustancial en la gestión de la información sobre flujos migratorios y, probablemente, los datos actuales se aproximen bastante más a la realidad, contribuyendo a acentuar la propia trayectoria creciente observada en las migraciones de la EVR a lo largo de los últimos quince años. Como la fiabilidad de esta fuente se ha reforzado con el tiempo, es muy difícil argumentar que el origen de la distancia entre la EVR y la EM se pueda encontrar en una sobrevaloración de las cifras en la primera(3). Más bien, sucede lo contrario, pues hay bastantes motivos para fundamentar la sospecha de que la EM subestima el número de inmigrantes, al tratarse de una herramienta poco recomendable para medir el fenómeno migratorio.

3. LAS TRANSFORMACIONES EN LA ENCUESTA DE MIGRACIONES

Como el propio INE advierte, la EM no es una encuesta específicamente proyectada para el estudio del fenómeno migratorio, sino que se genera aprovechando los recursos desplegados en la elaboración de la EPA. Por tanto, los cambios en el diseño de la EPA pueden repercutir en la estimación de la población migrante. Este es el caso de las sucesivas actualizaciones de la muestra, así como del cambio de

(2) Por este motivo, en la figura 1 las cifras de la EVR en los años 1981, 1986, 1991 y 1996 se han corregido interpolando los datos del año previo y del siguiente.

(3) Por ahora, no es posible verificar para el período más reciente los datos de la EVR a través de una fuente de información alternativa como podría ser el Censo o el Padrón, dado que el Censo 2001 no ha sido publicado todavía y en el caso del Padrón ni la última renovación padronal de 1996 ni el posterior Padrón Continuo incorporan datos de movilidad. No hay más remedio, entonces, que remitirse al Censo de 1991, cuya comparación ya se realizó en Ródenas y Martí (1997:159), confirmándose hasta esa fecha la semejanza entre ambas series estadísticas.

cuestionario y las modificaciones en el desarrollo de la encuesta desde el primer trimestre de 1999.

3.1 Las actualizaciones de la muestra de secciones

La unidad primaria del muestreo bietápico en el que se basa la EPA es la *sección censal*, que se define como un área geográfica perfectamente delimitada. Los criterios geográficos y socioeconómicos son los que se utilizan para estratificar y seleccionar las secciones censales que entran en la muestra(4). Dicha muestra permanece fija en la encuesta mientras que los nuevos Censos de población, Padrones o Censos electorales no aconsejen su cambio [INE (2002:25)]. Dado que la población es dinámica y cambia tanto en su cuantía como en su forma de distribuirse por el territorio, con el paso del tiempo la muestra de secciones deja de ser una representación fiel de la población. Así, hay secciones que ganan en población y otras que pierden. Las primeras están representadas en la muestra con menor probabilidad que la que les correspondería, mientras que las segundas están representadas con mayor probabilidad, ya que la probabilidad de cada sección en la muestra es proporcional a su tamaño de población.

Por esta razón, se procede periódicamente a la actualización de las probabilidades de selección o a la renovación de la muestra de secciones, a través de los nuevos Censos de población, Censos electorales y Padrones. Como se indicaba en Ródenas y Martí (1997:160 y ss.), éstos revelan las zonas de crecimiento demográfico pero el problema es que lo hacen *a posteriori* por lo que se corre el riesgo de que las nuevas secciones utilizadas no garanticen que se capten los movimientos migratorios en el momento y lugar en que se están produciendo. Por tanto, las áreas de crecimiento demográfico como consecuencia de la migración pueden pasar completamente desapercibidas hasta que se renuevan las secciones censales, e incluso es posible que para entonces ya se haya aminorado la afluencia o que los inmigrantes se dirijan hacia otros lugares(5). La solución a este problema de “envejecimiento de la muestra de secciones” reside en actualizar con asiduidad la muestra de secciones censales.

Hasta el último año abordado en este trabajo, el INE ha efectuado dos renovaciones del seccionado. Una completa entre 1995 y 1996, y otra parcial en el primer

(4) La muestra total de secciones se encuentra dividida en seis grupos, denominados turnos de rotación, renovándose cada trimestre las viviendas de las secciones que pertenecen a un determinado grupo.

(5) Otra cuestión es que en las secciones que finalmente forman parte de la muestra, al producirse en ellas ciertas incidencias (particiones, fusiones o revisiones de límites), pudiera captarse una ínfima parte del flujo migratorio reciente.

trimestre del año 2000 donde sólo se sustituyó el 4,1 por ciento de las secciones. Al analizar sus efectos⁽⁶⁾ ha aflorado un problema importante en relación con la capacidad de la EM para captar los inmigrantes, y que está relacionado con la perseverante subestimación de la EPA de la población en los tramos de edad con mayor propensión a la movilidad. Para el estudio de las migraciones resulta fundamental que, precisamente, sea el colectivo entre 25 y 34 años el que pierde más peso relativo en la EPA, al subestimarse justo el grupo de edad que más migraciones realiza. Como puede comprobarse en Ródenas y Martí (2002:43), el mayor número de movimientos calculados por la EVR son los realizados por personas con edades comprendidas entre 25 y 34 años, representando más del 30 por ciento del total de las migraciones.

Esta incapacidad de la EPA para estimar correctamente la estructura por edades de la población española aún después de las actualizaciones del seccionado sólo se explica por la falta de respuesta. La continuada deficiencia en la estimación de la pirámide de edades después de las dos actualizaciones del seccionado ha puesto de manifiesto que las dificultades cada vez mayores para entrevistar a determinados hogares, esto es, la falta de respuesta, se están constituyendo como un foco de sesgo importante en las estimaciones de la EPA. Para corregir este problema, el INE ha procedido a partir de 2002 a la reponderación de los factores de elevación a partir de la información proporcionada por el Padrón Continuo. Cabe esperar que esta reponderación de los factores de elevación sirva para mejorar las estimaciones de la EPA y de la EM.

3.2 El cambio de cuestionario y el desarrollo de la encuesta desde el primer trimestre de 1999

Desde el primer trimestre de 1999, el INE ha introducido cambios tanto en la redacción del cuestionario como en las instrucciones que siguen los entrevistadores en el desarrollo del trabajo de campo de la EPA. Algunas de estas novedades han sido inducidas por el Eurostat para armonizar las encuestas de los países pertenecientes a la UE. Concretamente, se ha adaptado la metodología de la EPA al Reglamento (CE) nº 577/98 del Consejo sobre la Encuesta Comunitaria de Fuerza de Trabajo. Dado que posiblemente estas transformaciones estén mejorando las estimaciones, se va a proceder a estudiarlas con detalle⁽⁷⁾.

(6) Por ejemplo, en García, Garrido y Toharia (1999) o en Álvarez (2000) para la primera actualización y según nuestras estimaciones para la segunda renovación –ver Tabla A del apéndice–.

(7) Para la identificación y el análisis de los diferentes cambios se ha utilizado INE (1994 y 1999b). En esta última publicación se puede encontrar, además, la relación de modificaciones que tienen su origen en las normas armonizadoras de EUROSTAT.

a) Cambio de cuestionario

En lo que atañe directamente a la medición de los inmigrantes, la innovación fundamental ha consistido en una ampliación del calendario de recogida de la información. Así, el apartado que permite estudiarlas, *situación hace un año*, se cumplimenta ahora los cuatro trimestres, en lugar de tan sólo el segundo trimestre(8) como se hacía desde 1987. Esta modificación, no obligatoria según la legislación comunitaria, era una de las posibilidades que ya se planteaba en Ródenas y Martí (1997:170) para tratar de resolver la subestimación que surgía por el cruce de la formulación de la pregunta para captar el hecho migratorio una sola vez al año y los turnos de rotación.

A partir de este momento, el INE calcula dos estimaciones diferentes del número de inmigrantes. La primera se obtiene a partir de un fichero anual que incluye únicamente los registros (individuos) de las viviendas que entran por primera vez en la muestra en cada trimestre. El cálculo anual del número de inmigrantes se basa, entonces, en una muestra total de unas 40.000 viviendas (en términos medios, 10.000 al trimestre). Una estimación que se puede considerar similar a la media anual que se obtenía con el método utilizado por el INE durante el período 1980-1985. Los resultados alcanzados con este sistema son reconocidos como los datos oficiales. Alternativamente, el INE ofrece una segunda estimación computada a partir de una media anual basada en los ficheros trimestrales completos, de forma análoga a la utilizada en la estimación de los diferentes colectivos del mercado laboral que publica habitualmente el INE en *Tablas Anuales EPA*.

Como se puede observar en la figura 1, el nuevo método oficial de estimación basado en los primeros turnos de los cuatro trimestres del año ha supuesto un salto favorable en el número de inmigrantes calculados por la EM, tanto en 1999 como en 2000. Este aumento se ha producido porque, tal y como se explicaba en Ródenas y Martí (1997), ha sido eliminada la infravaloración provocada por la disminución de la tasa de inmigración a medida que aumentaba la permanencia de las viviendas en la muestra de la EPA. Tal anomalía se producía por dos razones. En primer lugar, al formular la pregunta migratoria –*¿cuál era su municipio de residencia hace exactamente un año?*– sólo en los segundos trimestres, a mayor número de trimestres entrevistados disminuía la posibilidad de que una familia que hubiera realizado efectivamente una inmigración pudiera computarse dentro del ámbito temporal de los doce meses al que hacía referencia la pregunta. De este modo, a medida que las submuestras agotaban sus cinco y seis trimestres de permanencia,

(8) Como la EPA media las migraciones tan sólo en el segundo trimestre de cada año, para el período 1987-98, la comparación entre ambas fuentes en la figura 1, se ha realizado ajustando anualmente los datos de la EVR por segundos trimestres.

su tasa migratoria se aproximaba a cero, pues nadie que desde entonces hubiera sido entrevistado continuamente habría podido realizar su inmigración en los cuatro últimos trimestres. En segundo lugar, porque el tratamiento que recibían las incidencias tales como la detección de viviendas vacías y los grupos humanos ausentes o inaccesibles provocaban pérdidas de muestra que eran mayores a medida que avanzaban los turnos de rotación(9). La disminución de la tasa migratoria también se relacionaba con el hecho de que una buena parte de esas pérdidas de muestra estaban razonablemente constituidas por inmigrantes.

b) Otras modificaciones que pueden mejorar las estimaciones de los inmigrantes

Del resto de cambios introducidos en la EPA a partir del primer trimestre de 1999, cabe destacar fundamentalmente dos: primero, los referidos al tratamiento de las incidencias y al método de recogida de datos y, segundo, los relacionados con el diseño de la muestra.

En primer lugar, se ha articulado un nuevo tratamiento de las incidencias en el trabajo de campo en segunda o posterior entrevista y algunos cambios en el método de recogida de información, que afectan especialmente a la falta de respuesta total. Así, las viviendas no encuestables (como las vacías) se visitan ahora los trimestres posteriores por si hubieran pasado a ser encuestables. Anteriormente, las viviendas vacías dejaban de visitarse en las segundas o posteriores entrevistas, produciéndose una pérdida de muestra al no sustituirse por los grupos humanos que pudieran habitar nuevamente dichas viviendas. El actual método de estimación de las migraciones no se ve influido por este cambio pues sólo utiliza la información de las viviendas que son entrevistadas por primera vez. Sin embargo, sí hubiera tenido incidencia en el método antiguo pues es fácil imaginar que una vivienda vacía acabara siendo el destino de una familia de inmigrantes.

Asimismo, y en relación con la falta de respuesta, en la descripción de la nueva encuesta se señala que en el caso de viviendas inaccesibles o grupos humanos ausentes, se repetirán los contactos (telefónicos o personales) con la vivienda seleccionada en horas y en días diferentes [INE (1999 a:17)], mientras que anteriormente sólo se preveía vagamente que se repitieran las visitas hasta conseguir la información. Por primera vez, se menciona la posibilidad de que existan entrevistas telefónicas en días y en horarios diferentes. Es más que probable que la intensificación del uso del teléfono para contactar con los hogares haya podido incidir, como señala Pérez Infante (1999:69), en la recuperación para la muestra de parte de la falta de respuesta en segundas o posteriores entrevistas. Nuevamente, el actual

(9) El tratamiento de estas incidencias se ha modificado también, como se explica a continuación.

método de estimación no se ve influido por este cambio, pero sí lo estaría si se hubiera mantenido el anterior sistema.

En segundo lugar y en relación con el diseño de la encuesta, para lograr que la EPA fuese estrictamente una encuesta *continua*(10), esto es, realizada a lo largo de todo el año, se ha modificado el tamaño y la distribución de la muestra en el tiempo. El INE se ha visto obligado a extender el período de encuesta al mes de agosto, ampliando la muestra final de 64.000 a 65.000 viviendas visitadas aproximadamente cada trimestre, lo que supone obtener información de unas 200.000 personas. Este pequeño aumento del tamaño de la muestra ha de contribuir a una mejora en cualquiera de las estimaciones de la encuesta al reducirse el error de muestreo, incluida la estimación de los inmigrantes.

4. LOS INMIGRANTES, ¿ES POSIBLE ESTIMARLOS CON EL ACTUAL DISEÑO DE LA EPA?

Cuando en Ródenas y Martí (1997:158) se cruzaba la información procedente de las dos fuentes estadísticas que venimos analizando, se calificaba como muy poco creíble el paso de una media de 1,3 migraciones por migrante y año entre 1980 y 1985 a una media de más de cuatro entre 1986 y 1996. Tras los cambios metodológicos y las transformaciones descritas en los epígrafes anteriores, en la figura 2 se han recogido los últimos datos procedentes de la EM y de la EVR con el ánimo de evaluar los nuevos resultados. Para proceder a la comparación del modo más ajustado posible al ámbito temporal que abarca cada fuente ha sido necesario retrasar en un año los datos de la EM(11).

(10) En las condiciones y con los requisitos que establece el Reglamento (CE) nº577/98 del Consejo.

(11) Esto es así porque mientras que el intervalo de tiempo que abarca la EVR es el mismo año al que está referida esta estadística, en la EM esto no ocurre del mismo modo. En la EM, al elaborarse a partir de la pregunta retrospectiva acerca del lugar de residencia hace un año, se captura justamente la movilidad realizada en los cuatro últimos trimestres, que en la mayor parte de los casos incluye el año anterior. En realidad, como los datos de la EM "cabalgan" entre dos años –el de referencia y el anterior– lo correcto sería comparar la EVR del año t con los datos procedentes en una parte de la EM del año t-1 y en otra de la EM del año t. A los efectos generales de la figura 2 ésta no es una cuestión relevante, pero sí lo es cuando, como en INE (2001:52-59), se presenta un estudio comparativo exhaustivo de ambas fuentes y no se considera este factor de ajuste temporal.

Figura 2
INMIGRANTES Y MIGRACIONES, 1998-2000

EM			EVR		
Año	Metodología 1987-98 (1)	Metodología oficial 1999 (2)	Año	(3)	Migraciones Inmigrantes (4) = (3) / (2)
1999	130.906	214.500	1998	933.223	4,35
2000	107.559	185.600	1999	1.006.127	5,42
2001	Nd	Nd	2000	1.032.084	Nd

Fuente: INE, *Encuesta de Migraciones, Estadística de Variaciones Residenciales* (varios años) y elaboración propia.

Como puede observarse, el nuevo método de estimación de los inmigrantes utilizado en la EM a partir de 1999 ha elevado sustancialmente las cifras –ver columna (1) y (2)–, llegando casi a duplicar el flujo de movilidad. Sin embargo, y a pesar de todas las mejoras introducidas en el diseño de la encuesta, este resultado continúa muy alejado del valor proporcionado por la EVR y su actual tendencia. En consecuencia, la relación migraciones por migrantes –ver columna (4)– sigue siendo disparatada. El hecho de que prácticamente cada inmigrante realice cinco migraciones al año supera cualquier previsión que se pueda realizar a partir de otras fuentes de control, como la *Encuesta Socio-demográfica* de 1991 donde la media de migraciones por migrante a lo largo de una vida era de 1,8 [INE (1991:91)], o la encuesta del *Eurobarómetro* 54.2 de 2001, donde de quienes declaran haber cambiado de residencia –incluso dentro del mismo municipio– en los últimos 10 años, el 64,2 por ciento lo ha hecho una sola vez, mientras que sólo el 2,6 por ciento lo ha realizado cinco o más veces.

La conclusión ante estos resultados es que el reciente esfuerzo del INE introduciendo cambios metodológicos para capturar mejor las migraciones ha resultado poco fructífero. Continúa existiendo una importantísima infravaloración en la EM que, en nuestra opinión, se encuentra lejos de constituir lo que el INE denomina(12) una simple subestimación del flujo migratorio. Como se explica a continuación, el origen del problema es mucho más profundo. Es posible que la investigación del hecho migratorio a través de la técnica de muestreo de poblaciones bien no pueda

(12) Así se señala en la EM desde 1999.

aplicarse satisfactoriamente o bien exija un diseño específico de la encuesta dado el tamaño y las características de esta subpoblación o dominio(13).

Como es sabido, para realizar la inferencia estadística es condición básica disponer de estimadores insesgados y precisos. Para probar si la EM proporciona estimaciones precisas de los inmigrantes, bastará con evaluar los errores de muestreo de dichas estimaciones. Concretamente, la precisión mide la dispersión de la distribución del estimador en el muestreo o, si se quiere, la magnitud de las desviaciones respecto a la media obtenida por la aplicación repetida del procedimiento de muestreo, desviación típica del estimador o error de muestreo [Cochran (1980:38)]. A medida que aumenta la precisión del estimador, el intervalo de confianza se hace menor y, en ausencia de sesgos, las posibles diferencias entre los valores de los parámetros poblacionales y los estimados en el muestreo son menores. Para facilitar la interpretación de los resultados y las comparaciones entre estimaciones, a menudo los errores de muestreo se presentan en porcentajes de la estimación, lo que se conoce como el coeficiente de variación. Este coeficiente alcanza valores elevados cuando se estiman parámetros poco frecuentes en la población.

La información sobre el coeficiente de variación para las estimaciones de la EM no se dispone pues el INE no calcula el error de muestreo asociado a estos resultados. Por lo tanto, no es posible comprobar si éstos ciertamente se han mantenido o no en unos niveles razonables. Este hecho desconcierta bastante, pues mientras que en el ámbito provincial el INE actúa prudentemente y previene a los investigadores de posibles errores muestrales elevados, en el análisis de este dominio, cuyo tamaño de población en torno al 2 por ciento es similar al de muchas provincias españolas, no realiza tal advertencia. Es realmente sorprendente que se dedique tiempo y recursos para la elaboración anual de una publicación específica y detallada sobre este dominio en particular como es la EM y que no se desarrolle paralelamente un análisis de la calidad de estas estimaciones.

A falta de dicha información, una manera de evaluar la precisión en las estimaciones de los inmigrantes bajo el actual diseño de la encuesta, es comparar las principales componentes demográficas de los inmigrantes obtenidas a partir de la EM (edad y procedencia(14)) con la estructura de la población inmigrante que se

(13) Siguiendo a Azorín y Sánchez-Crespo (1986:151), cualquier población investigada puede dividirse en clases o divisiones, denominadas subpoblaciones o dominios, "para las que se requieren estimaciones separadas cuyo interés reside en sí mismas y no como parte de una estimación global relativa a la población en su conjunto. Estas clases suelen cruzar a los estratos y por consiguiente la selección de la muestra no puede hacerse directamente en los dominios." El caso que nos ocupa, la EM, constituye un ejemplo de dominio en la EPA.

(14) Aunque en la figura 1 sólo se comparaba las migraciones interiores, en la figura 3 se ha distinguido entre interiores y exteriores para mostrar que el problema se extiende a ambos tipos de movimientos.

deduce de una fuente alternativa como los censos de población que, en definitiva, constituyen una medida muy aproximada del universo de estudio. Desafortunadamente, al no disponer por el momento de la explotación del Censo 2001, hay que recurrir a los datos del Censo 1991.

En la figura 3, la última fila de cada uno de los tres bloques refleja, en términos relativos, la diferencia entre las estimaciones obtenidas de la muestra y la estructura obtenida del Censo. La precisión o el ajuste de las estimaciones es mayor cuando el error relativo es más pequeño. Como puede observarse, en la mayoría de las categorías establecidas los errores relativos superan considerablemente el 10 por ciento. Esto significa que el nivel de ajuste, en general, es muy bajo y que los errores de muestreo probablemente serán muy elevados. En consecuencia, con un nivel de ajuste tan bajo, no debería utilizarse la información de la EM para el estudio de esta subpoblación, ni tan siquiera de un modo aproximado.

Figura 3

ESTRUCTURA POR EDAD DE LOS INMIGRANTES DE 1 Y MÁS AÑOS RESIDENTES EN VIVIENDAS FAMILIARES Y ERRORES RELATIVOS

(Continúa)

Grupos de edad (%)	Total	De 1 a 4 años	De 5 a 14 años	De 15 a 24 años	De 25 a 34 años
Inmigrantes					
Censo 1991	100	6,15	12,94	20,42	28,99
EM 91	100	5,63	8,79	25,20	29,38
Error relativo ¹		-8,47	-32,07	23,41	1,32
Inmigrantes interiores					
Censo 1991	100	6,35	13,13	20,93	29,81
EM 91	100	6,16	8,44	25,61	30,90
Error relativo ¹		-3,06	-35,72	22,34	3,66
Inmigrantes exteriores					
Censo 1991	100	4,64	11,49	16,66	23,01
EM 91	100	2,53	10,79	22,81	20,50
Error relativo ¹		-45,44	-6,13	36,91	-10,90

(1) El error relativo es la diferencia entre la estructura derivada de la EM y del Censo, expresada en porcentaje de la estructura del Censo.

Fuente: INE, Censo 1991, Encuesta de Migraciones 1991 y elaboración propia.

**ESTRUCTURA POR EDAD DE LOS INMIGRANTES DE 1 Y MÁS AÑOS RESIDENTES
EN VIVIENDAS FAMILIARES Y ERRORES RELATIVOS**

(Conclusión)

Grupos de edad (%)						
	Total	De 35 a 44 años	De 45 a 54 años	De 55 a 64 años	De 65 a 74 años	De 75 y más años
Inmigrantes						
Censo 1991	100	12.74	6.63	5.65	3.77	2.71
EM 91	100	9.93	5.47	4.58	4.55	6.49
Error relativo ¹		-22.06	-17.45	-19.07	20.57	139.15
Inmigrantes interiores						
Censo 1991	100	12.47	6.03	4.89	3.52	2.86
EM 91	100	9.07	4.26	3.79	4.28	7.49
Error relativo ¹		-27.30	-29.43	-22.39	21.58	162.16
Inmigrantes exteriores						
Censo 1991	100	14.68	11.01	11.25	5.59	1.66
EM 91	100	14.94	12.55	9.12	6.08	0.67
Error relativo ¹		1.73	14.02	-18.89	8.85	-59.73

(1) El error relativo es la diferencia entre la estructura derivada de la EM y del Censo, expresada en porcentaje de la estructura del Censo.

Fuente: INE, Censo 1991, Encuesta de Migraciones 1991 y elaboración propia.

Una explicación del hecho de que la EM mida mal a los inmigrantes (interiores y exteriores), pasaría por recordar los factores que determinan la precisión de una estimación, esto es, el tamaño de la muestra y la heterogeneidad u homogeneidad de la variable estudiada. Respecto al tamaño muestral, cabe señalar que muestras dimensionadas para extraer conclusiones generales o medir determinados parámetros no se pueden utilizar para el análisis de caracteres muy poco frecuentes, pues el número de entrevistas realizadas a personas que dispongan de dicho rasgo no será suficiente para dar soporte a la estimación, ya que se puede superar ampliamente lo que se podría llamar el error permitido. Lógicamente, como señala Rodríguez Osuna (1991:16), sólo se pueden estimar los rasgos poblacionales cuya presencia en la muestra sea suficiente para proceder a la inferencia estadística.

Como se puede comprobar en Ródenas y Martí (2002:40), los inmigrantes son un ejemplo de rasgo poblacional poco frecuente dentro del universo formado por la población española. El tamaño de este dominio medido a través de la tasa migratoria o el porcentaje de migraciones interiores de la EVR sobre la población a 1 de

julio de cada año(15), representa aproximadamente entre un 1,2 y un 2,6 por ciento de la población española. Cuando un dominio alcanza este tamaño, lo que Purcell y Kish (1979:367) califican como *dominio menor*, no siempre puede aplicarse satisfactoriamente el muestreo tradicional. De hecho en la EPA, a pesar de establecerse en la muestra de secciones censales una afijación proporcional corregida por provincias, no se ofrece a nivel provincial todos los resultados detallados, sino tan sólo aquellas características que los coeficientes de variación de los estimadores lo permiten(16).

A las dificultades propias en la estimación de un dominio de carácter menor, en el caso de las migraciones se suman otros dos inconvenientes ligados al diseño de la EPA. En primer lugar, los inmigrantes constituyen una clase que cruza probablemente todos los estratos de la población de manera bastante uniforme y, en segundo lugar, se trata de una característica que se distribuye de forma muy homogénea dentro de la unidad última de muestreo, la vivienda. En relación con el primer inconveniente, en cualquier manual de métodos de muestreo(17) se insiste en que las variables utilizadas para la estratificación deberían estar correlacionadas con las variables objeto de la investigación y que la eficacia del muestreo estratificado depende de si se posee información suficiente para definir y separar los grupos o estratos correctamente. La estratificación, en tanto que logra reunir en cada estrato unidades homogéneas entre sí y heterogéneas en relación con el resto de los otros estratos, reduce la varianza de los estimadores, gana en precisión y, en definitiva, contribuye a reducir los errores de muestreo.

Pues bien, observando la distribución de los inmigrantes entre los diferentes estratos de la muestra(18) EPA de varios años, figura 4, en ninguno alcanza un porcentaje superior al 6 por ciento de la población muestral del estrato, ya que se trata de una característica poco correlacionada con los criterios de estratificación. Así pues, desde el momento en que los actuales criterios de estratificación utilizados en la EPA no consiguen concentrar en unos pocos estratos a la población migrante, la eficacia del muestreo estratificado – entendida como su contribución a reducir los errores de muestreo y, por tanto, mejorar la estimación– para esta variable es muy baja.

(15) Se está suponiendo una migración igual a un inmigrante. Aunque ciertamente un inmigrante puede realizar varios movimientos en un año, el indicador adecuado (inmigrantes/población) tendrá un valor muy próximo al ofrecido con un sesgo, si acaso, a la baja, lo que reafirmaría todavía más la consideración del colectivo de inmigrantes como un dominio menor.

(16) No obstante, el INE sí dispone de técnicas específicas de estimación para áreas pequeñas, como se explica en López Paños (2000).

(17) Por ejemplo, Azorín y Sánchez-Crespo (1986:83), García Barbancho (1981:413) o Rodríguez Osuna (1991:27).

(18) Nótese que se está analizando la distribución de los inmigrantes entre los estratos de la muestra. No se sabe nada de cómo se distribuirá esta variable entre los estratos de población, aunque cabe suponer que de forma muy similar.

Figura 4**DISTRIBUCIÓN DE LOS INMIGRANTES POR ESTRATOS**

(Continúa)

Porcentaje que representan los inmigrantes en la población del estrato				
	0 %	Más de 0 y menos de 1%	Más de 1 y menos de 2%	Más de 2 y menos de 3%
EPA 2º trimestre 1991				
Nº estratos	52	164	34	9
Porcentaje	20,00	63,08	13,08	3,46
Porcentaje acumulado	20,00	83,08	96,15	99,62
EPA 2º trimestre 1996				
Nº estratos	70	168	28	2
Porcentaje	25,93	62,22	10,37	0,74
Porcentaje acumulado	25,93	88,15	98,52	99,26
EPA 2º trimestre 2000				
Nº estratos	109	125	33	2
Porcentaje	40,37	46,30	12,22	0,74
Porcentaje acumulado	40,37	86,67	98,89	99,63

(Conclusión)

	Más de 3 y menos de 4%	Más de 4 y menos de 5%	Más de 5 y menos de 6%	Total
EPA 2º trimestre 1991				
Nº estratos	0	1	0	260
Porcentaje	0,00	0,38	0,00	100
Porcentaje acumulado	99,62	100	100	100
EPA 2º trimestre 1996				
Nº estratos	2	0	0	270
Porcentaje	0,74	0	0	100
Porcentaje acumulado	100	100	100	
EPA 2º trimestre 2000				
Nº estratos	0	0	1	270
Porcentaje	0,74	0,74	0,37	100
Porcentaje acumulado	99,63	99,63	100	

Fuente: INE, Encuesta de Población Activa y elaboración propia.

El segundo inconveniente se deriva del hecho de que la movilidad sea una característica que se distribuye de forma muy homogénea dentro de la unidad última

de muestreo, la vivienda. Como se puede contemplar en las diagonales principales de la figura 5, cuando aparece el fenómeno de movilidad en la vivienda esta característica se extiende con frecuencia a todos sus miembros(19). Para que el muestreo por conglomerados(20) sea igual de preciso en una determinada característica que el muestreo aleatorio simple, se requiere que no exista correlación en esa variable entre los miembros del conglomerado, es decir, que no exista homogeneidad en la variable para todos los miembros de la vivienda. Justo lo contrario de lo que sucede con los inmigrantes.

Figura 5

**DISTRIBUCIÓN PORCENTUAL DE LAS VIVIENDAS CON ALGÚN INMIGRANTE
SEGÚN TAMAÑO DE LA VIVIENDA**

(Continúa)

Tamaño de la vivienda (individuos)	EPA 2º trimestre 1991								
	Inmigrantes por vivienda								
	1	2	3	4	5	6	7	11	
1	5,2								
2	19,4	43,0							
3	21,4	17,2	55,6						
4	19,4	6,3	16,7	68,8					
5	17,0	14,1	11,1	6,3	86,7				
6	8,3	10,2	7,4	15,6	6,7	100			
7	4,6	4,7	3,7	3,1	6,7		100		
8	3,7	2,3		3,1					
9	0,2	0,8	3,7	3,1					
10	0,4	0,8							
11	0,4	0,8							100
12									
13			1,9						
Total	100	100	100	100	100	100	100	100	100

Fuente: INE, Encuesta de Migraciones y elaboración propia.

(19) La figura 5 refleja, además, una de las consecuencias de la falta de respuesta en la EPA. El bajo porcentaje del cruce del grupo "un inmigrante por vivienda" con el de "un individuo por vivienda" probablemente no responde a la realidad y sea superior, pues las viviendas habitadas por hogares unipersonales es una de las categorías especialmente afectada por los problemas de falta de respuesta y de localización del grupo humano.

(20) Cuando la unidad de muestreo comprende dos o más unidades de estudio o últimas se denomina conglomerado. En la EPA, la unidad de muestreo es la vivienda que incluye a uno o más individuos.

DISTRIBUCIÓN PORCENTUAL DE LAS VIVIENDAS CON ALGÚN
INMIGRANTE SEGÚN TAMAÑO DE LA VIVIENDA

(Continuación)

Tamaño de la vivienda (individuos)	EPA 2º trimestre 1996					
	Inmigrantes por vivienda					
	1	2	3	4	5	6
1	8,3					
2	21,8	27,5				
3	23,3	28,2	46,7			
4	18,3	19,8	20,0	40,0		
5	17,3	9,9	26,7	20,0	50,0	
6	5,3	10,7	6,7	40,0		
7	3,3	1,5				
8	1,3	1,5				100
9	1,3					
10	0,3				50	
11		0,8				
12						
13						
Total	100	100	100	100	100	100

(Conclusión)

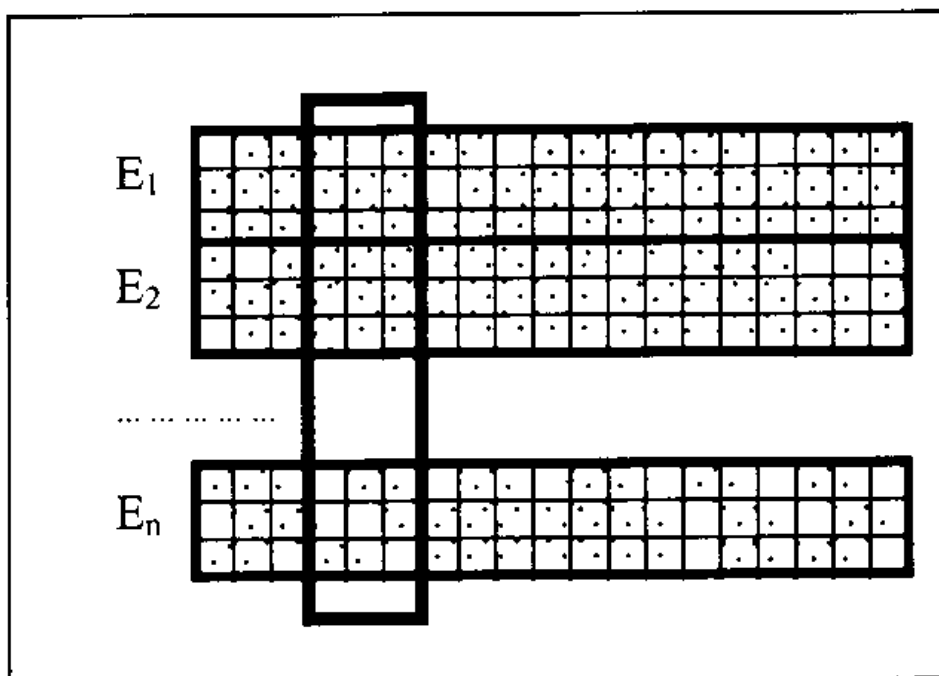
Tamaño de la vivienda (individuos)	EPA 2º trimestre 2000							
	Inmigrantes por vivienda							
	1	2	3	4	5	6	8	
1	14,2							
2	28,4	56,7						
3	19,3	15,5	78,4					
4	15,2	15,5	13,5	78,3				
5	14,5	7,2	5,4	13,0	71,4			
6	7,1	3,1		8,7	14,3	100		
7	1,0	2,1	2,7		14,3			
8	0,3							
9								
10							100	
11								
12								
13								
Total	100	100	100	100	100	100	100	

Fuente: INE, Encuesta de Migraciones y elaboración propia

Una explicación gráfica puede aclarar intuitivamente las dificultades de estimar un dominio de estas características. Es lo que se ofrece en las figuras 6, 7 y 8 donde se supone que se ha clasificado al universo poblacional en estratos (E_i). Cada estrato está constituido por las viviendas correspondientes, que se representan mediante casillas. A su vez, dentro de cada vivienda los individuos que poseen la característica estudiada están recogidos con un punto. Y, por último, la muestra de viviendas seleccionadas –y de población entrevistada– se encuentra dentro de la barra vertical que atraviesa todos los estratos.

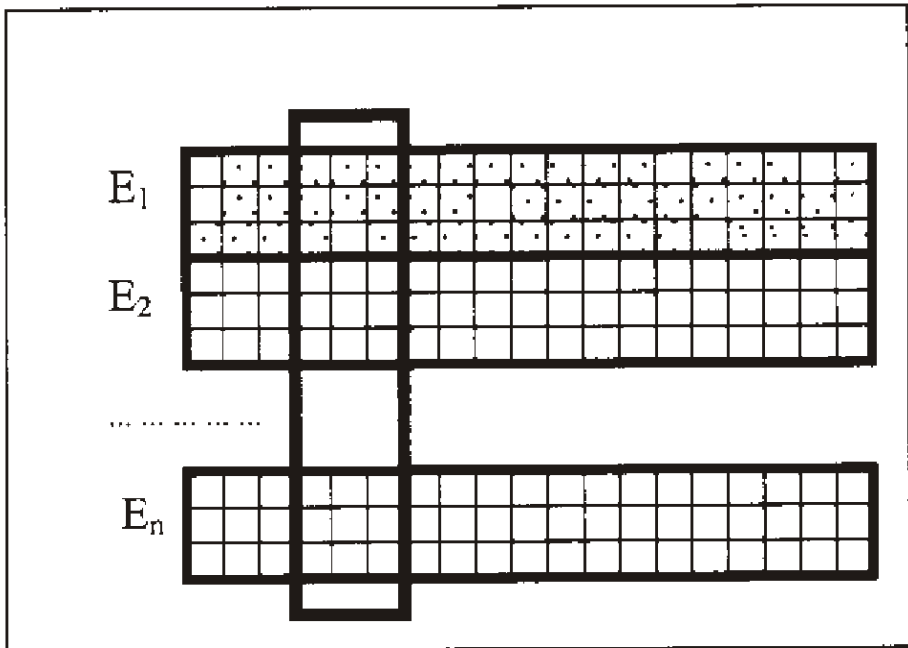
Imaginemos que queremos estimar un *dominio principal* que cruza por igual todos los estratos y que, además, es una característica muy heterogénea en la unidad última de muestreo. Éste sería el caso, por ejemplo, del sexo, los tramos de edad o el estado civil. Como lo normal es que las viviendas estén constituidas por personas de diferente sexo, edad y estado civil, la muestra que se extraiga en cualquier estrato estará compuesta con toda garantía de una buena porción de individuos que cuentan con esa característica, como se aprecia en la figura 6.

Figura 6



Si la característica que se quiere estimar es un *dominio menor* que se distribuye de forma homogénea entre las unidades últimas de muestreo, pero se considera como un *dominio de estudio* y se planea el diseño de la encuesta generando un estrato que contenga las viviendas en las que algún individuo dispone de esa característica, la estimación no revestiría ningún problema de precisión. Como se observa en la figura 7, los individuos correspondientes se concentrarían en un estrato y tal característica se recogería adecuadamente en la muestra.

Figura 7

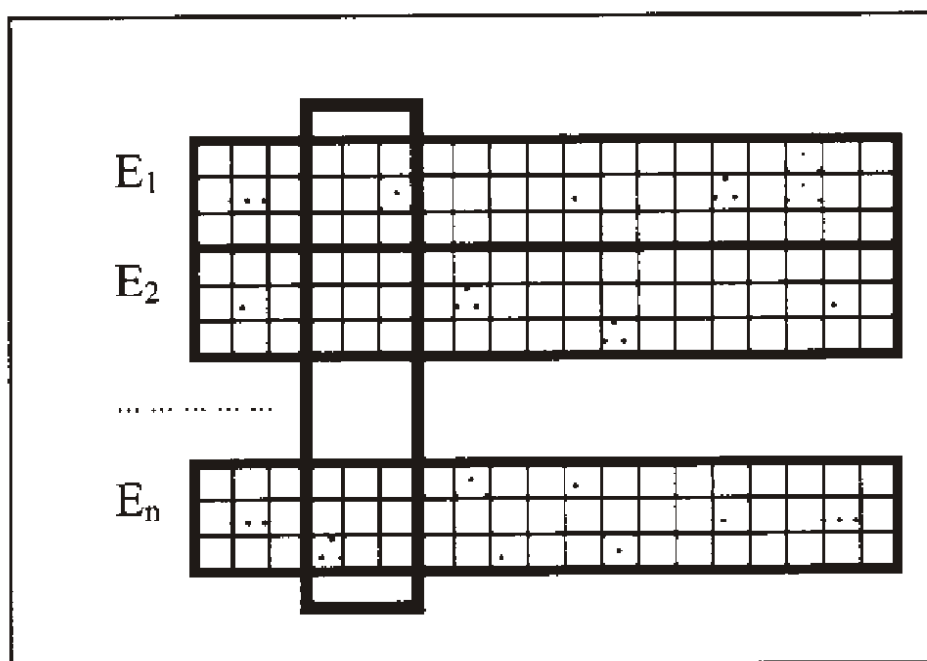


Ahora bien cuando ese mismo dominio, como sucede con los inmigrantes o con los ciudadanos de nacionalidad extranjera(21), no es considerado *de estudio*, aparecen las dificultades. Como se recoge en la figura 8, la mayoría de las unidades muestrales no contendrán dicha característica y es muy probable, entonces,

(21) Para los individuos con nacionalidad extranjera, se comprobó en la EPA de 1991 que constituían una categoría poblacional que presentaba similares dificultades de estimación que los inmigrantes, pues se trata de una característica que cruza por igual todos los estratos de la muestra y se distribuye de forma homogénea dentro de la vivienda. Así, mientras que para el Censo de 1991 en España había 353.367 personas de nacionalidad no española, la EPA estimaba 143.723 personas, es decir, sólo un 40 por ciento del valor del Censo.

que la muestra esté formada por individuos que no pertenezcan a esa categoría o que sólo unos pocos formen parte de la misma. En ausencia de sesgos, cuando en la figura 8 se tomen muestras sucesivas del universo poblacional, las estimaciones de la característica estudiada presentarán una elevada variabilidad o, dicho de otro modo, la desviación típica calculada a partir de las estimaciones muestrales será alta. El investigador se enfrentaría a un problema de precisión que haría que las estimaciones tomaran valores por encima del real, por debajo del mismo y, en ocasiones, adoptarían el valor real de la característica investigada.

Figura 8



Como el diseño de la EPA y sus criterios de estratificación son ajenos a la variable migratoria, el tamaño de este dominio es pequeño, se distribuye de forma uniforme entre los estratos y se trata de una característica homogénea dentro de la vivienda –unidad última de muestreo–, la figura 8 sería la más aproximada a la representación del tratamiento estadístico de esta variable. Por tanto, se esperaría que, en algún momento, la estimación del número de migrantes a partir de la EPA se acercara al valor real. Si consideramos la EVR como una aproximación acepta-

ble a las migraciones reales de la población residente en España, esto nunca ha sucedido, como se aprecia en la figura 1.

La única explicación de este hecho es que la EM no sólo adolece de un problema de precisión estadística sino, también, de sesgo en la estimación, razón por la cual nunca alcanzará valores similares a los de la EVR. No bastaría, entonces, con incrementar el tamaño muestral o establecer diseños específicos que aumenten la fracción de muestreo de este dominio para mejorar la calidad de las estimaciones de esta fuente, pues con ello sólo se resolvería la ausencia de precisión y no el sesgo. La amplia diferencia en los niveles a los que se sitúan una y otra fuente estadística debe proceder del impacto que provoca el sesgo de no respuesta que presenta la muestra de viviendas de la EPA. Y, hoy por hoy, este sesgo se origina básicamente por la falta de respuesta en los hogares encuestados en primer turno(22).

La falta de respuesta, bien por la ausencia de personas en las viviendas o bien por la negativa o el rechazo a ser encuestado, se trata de un error ajeno al muestreo de difícil cuantificación. De todos modos, el INE ha publicado y difundido hasta 1997 la evaluación de la calidad de los datos de la EPA, gracias a ello se conocen ciertas características de las viviendas con negativa o ausencia, pero no las relacionadas con la residencia habitual en el año anterior. Sí que se sabe que entre 1991 y 1997 la falta de respuesta ha afectado aproximadamente al 11 por ciento de las viviendas encuestadas en todos los turnos de entrevista y, algo más, al 14 por ciento de las viviendas del primer turno. Parte de la falta de respuesta se recupera posteriormente con entrevistas a las mismas viviendas, sin embargo, cuando se trata de negativas en primera entrevista, estas unidades muestrales son sustituidas aleatoriamente por otras de la misma sección censal.

La distinción en el tratamiento de la falta de respuesta según el turno es muy importante a los efectos de la actual estimación de las migraciones, ya que actualmente éstas se calculan exclusivamente a partir de los primeros turnos de entrevista en los cuatro trimestres. Las negativas de personas con movilidad nunca aparecerán en la estimación y difícilmente –dada la baja tasa migratoria– la sustitución aleatoria se hará por una vivienda con efectivos de idénticas características de movilidad a los de la negativa. Si la falta de respuesta no imputada o corregida mediante sustituciones, tal y como indica Álvarez (2000), es una causa importante de la infraestimación en el grupo de edades entre 20 y 49 años, todavía es más

(22) Posiblemente, al sesgo que surge de la falta de respuesta habría que añadir el que procede de la ejecución del propio trabajo de campo. No es improbable que ante la inminencia de un cambio de residencia entrevistador y entrevistado decidan considerar “mejor” abandonar la encuesta antes de su inicio por la imposibilidad de continuarla en el curso de los trimestres siguientes. En este caso, no se computaría ni siquiera una falta de respuesta por negativa, pero el sesgo que genera los problemas de nivel entre la EVR y la EM sí aumentaría.

relevante en el caso de las migraciones pues la probabilidad de captar otras características –como sexo, grupos de edad, situaciones laborales– es mayor que la de encontrar personas con episodios de movilidad. La mayor parte del sesgo posiblemente se origina aquí, en la medida en que la falta de respuesta esté correlacionada con la característica migratoria de la vivienda entrevistada y ésta se sustituya por otra vivienda con una reducida probabilidad de presentar movilidad.

No existe evidencia que confirme la hipótesis anterior, pero resulta razonable en términos de cuantificación. La actual diferencia entre la EVR y la EM se sitúa en torno a una media de 750.000 migrantes. Por tanto, el sesgo en términos de los efectivos de la muestra de la EPA equivale a unos 2.200 registros individuales. Presumiendo una media de tres personas por familia y una familia por vivienda, lo anterior implica que a lo largo del año son algo más de 700 viviendas con migrantes las que no se capturan por la falta de respuesta, unas 175 al trimestre. Si, como se muestra en las evaluaciones de la calidad de la EPA del INE, la falta de respuesta en primera entrevista afecta a unas 1.600 viviendas al trimestre, sólo haría falta que poco más de la décima parte de esas viviendas estuvieran formadas por migrantes para que la estimación a partir de la EPA se aproximara sustancialmente a la de la EVR.

Para corregir el sesgo procedente de la falta de respuesta en la estructura de la población por edades estimada por la EPA, en el primer trimestre de 2002 el INE ha aplicado coeficientes correctores en los factores de elevación de los registros individuales(23). Utilizando como variables auxiliares y de control a la población de 16 y más años por grupos de edad y sexo a nivel de provincia y de CC.AA se consigue de este modo, según el INE, otorgar a los entrevistados de cada tramo de edad y sexo su peso o representatividad en el conjunto de la población. Como las variables auxiliares están actualizadas gracias al Padrón Continuo estas correcciones se pueden realizar incluso con frecuencia anual. Así, suponiendo que la distribución de las variables en la muestra de población de la EPA representa a la población total adecuadamente, con la corrección de los factores de elevación se devolvería a los datos obtenidos la estructura real que debieran tener.

En el caso de las migraciones está técnica sería fácil de aplicar siempre que se dispusiera de una fuente estadística de control. Podrían servir los sucesivos censos de población que incluyen suficiente información migratoria, pero la amplitud de los

(23) En el primer trimestre de 2002 el INE ha corregido los factores de elevación con las nuevas proyecciones de población, que incluyen el incremento demográfico experimentado por la población española como resultado de la llegada de inmigrantes exteriores. Asimismo, y para tratar de paliar el efecto de la falta de respuesta también se han aplicado técnicas de reponderación de los factores de elevación. Esta posibilidad estaba prevista en el artículo 3 del Reglamento (CE) nº577/98 del Consejo como una mejora técnica deseable a implantar y se aplica ahora cuando por fin se dispone del Padrón Continuo que ofrece cifras de población frecuentes y de calidad.

intervalos intercensales hace pensar que no podrían mantenerse con validez las correcciones de los factores de elevación para los migrantes a lo largo de diez años. Y, en el caso de ser aceptables, podrían aplicarse siempre que las características de la exigua submuestra de inmigrantes en la EPA fuesen representativas de toda la población migrante. Alternativamente podría utilizarse la EVR, con frecuencia anual pero con falta de información sobre muchas de las variables contempladas en la EPA –especialmente importante es la carencia de datos sobre la situación familiar y laboral de los migrantes–. En ese caso, la reponderación eliminaría el sesgo en el total anual de migrantes estimados, pero si la submuestra de la EPA constituye un grupo de características particulares respecto al conjunto de la población migrante, la reponderación no anularía el sesgo en las distintas características de la población migrante. Para que dicha reponderación fuera útil, permitiendo la investigación y el estudio de este dominio, sería necesario contar en la EVR con información auxiliar adicional, especialmente la referente a la relación con el mercado de trabajo de quienes se dan de alta en un nuevo municipio. Pero esto es redundante. Si a través de un registro administrativo cada vez más fiable como la EVR se puede estudiar la relación entre migraciones y mercado laboral, la tarea de estimar lo mismo a partir de la EPA deja de tener sentido.

Sólo en el caso de que la EVR –al tratarse de una estadística procedente de un registro que no admite legalmente otras variables más que las estrictamente necesarias para su finalidad– no pudiera ampliar la información recabada, tendría sentido seguir elaborando la EM siempre reponderada por la EVR y comprobando en los años censales si la estimación de las variables sociolaborales es adecuada.

CONCLUSIONES Y SUGERENCIAS

Comenzábamos este trabajo destacando la formidable distancia que se observaba entre los valores alcanzados por la *Encuesta de Migraciones* y la *Estadística de Variaciones Residenciales* en la medición del fenómeno migratorio a partir del año 1982. Una diferencia que, además, no se ha logrado reducir en el último quinquenio, cuando se han incorporado diversas modificaciones metodológicas en ambas fuentes de información con el ánimo de mejorar la forma de medir el hecho migratorio en cada una.

En particular, los progresos introducidos en el tratamiento de la información en la EVR apuntan a un reforzamiento en la fiabilidad de esta fuente, por lo que existen bastantes motivos para sospechar que la EM pueda tratarse de una herramienta poco recomendable para medir el fenómeno migratorio. De hecho, ni la actualización de la muestra de la EPA, ni los diversos cambios introducidos tanto en el calendario y los métodos de recogida de información, como en el diseño de la

encuesta o en el tratamiento de las incidencias han logrado, ni tan siquiera, un leve acercamiento entre ambas series. La relación migraciones por migrante sigue siendo muy poco creíble.

Las dificultades de la EM para estimar esta subpoblación tienen su origen en diversas causas. Para comenzar, los inmigrantes constituyen un dominio de estudio de carácter menor dentro del universo formado por la población española. Asimismo, se trata de una característica poblacional que probablemente cruce todos los estratos de la población de forma uniforme, de manera que la estratificación no contribuye en absoluto, a una mejora en la precisión de la estimación. Y, finalmente, la movilidad es un rasgo que se distribuye de forma homogénea dentro de las unidades últimas de muestreo, las viviendas. A los inconvenientes anteriores que afectarían a la precisión de las estimaciones, habría que sumar los problemas de sesgo generados probablemente por la falta de respuesta en los primeros turnos de entrevista en la encuesta, que incide de manera particular en este dominio.

Una opción para afrontar los problemas de sesgo en la estimación de la población migrante es averiguar el estatus migratorio de las viviendas –y registros individuales– con falta de respuesta en primera entrevista. Pero esto resulta muy difícil. Aunque el INE incorporara la variable de movilidad en el “cuestionario de negativas” que se elabora para cada vivienda, el porcentaje de cuestionarios que los encuestadores de la EPA consiguen con información directa de las personas no alcanza la mitad, el resto procede de fuentes indirectas que previsiblemente no imputan migraciones. Además, restaría obtener la movilidad de los ausentes en primer turno de entrevista y asignarla al trimestre correspondiente.

Semejantes complicaciones no son fáciles de abordar con éxito y la experiencia en otros países ha concluido con el abandono de encuestas similares. Concretamente, hasta 1987 la Oficina Australiana de Estadística publicaba la *Encuesta de Migración Interna* (EMI), una encuesta que formaba parte del programa suplementario de la *Encuesta de Fuerzas de Trabajo* (EFT), similar a la EPA española. Como explica Cook (1996:13), la encuesta migratoria se dejó de elaborar en mayo de 1987 porque su calidad se consideró subestándar y, como en el caso español, se contradecía con las estimaciones procedentes de otras fuentes administrativas. Actualmente, la Oficina Australiana de Estadística utiliza las transferencias en la base de datos MEDICARE que elabora la Comisión de Aseguramiento de la Salud para estimar las migraciones interiores. Es lo mismo que sucede con otros países de nuestro entorno, donde se ha de recurrir a fuentes indirectas procedentes, por ejemplo, del sistema impositivo, médico o educativo para estimar la movilidad interior, enfrentándose casi todos a serios problemas de calidad y confianza en los datos.

En nuestra opinión y en la medida en que el Censo 2001 ratifique satisfactoriamente los datos de variaciones residenciales procurados por la EVR, esta estadística debería ser la de referencia para los movimientos migratorios. De hecho, la EVR es un registro administrativo directo de las migraciones firmemente implantado y lo óptimo sería potenciarlo recogiendo más características de los migrantes; especialmente, las relacionadas con la situación de las personas en el mercado de trabajo, información que dejó de solicitarse en 1987. Sin embargo, ésta parece ser una opción difícil de llevar a cabo pues por imperativo legal de protección de los datos personales, la información del Padrón hoy se reduce a las variables imprescindibles para la finalidad con la que se crea. Por eso, habría que recurrir a una solución intermedia.

La propuesta consiste en mantener la elaboración de la EM pero aplicando la técnica de reponderación a los estimadores con el objeto de ajustar las estimaciones de la encuesta a la información procedente de fuentes externas como la EVR, utilizando como variables de control el sexo, la edad, la titulación académica y la región. Además, en los años censales se comprobaría si la estimación del resto de variables sociolaborales es adecuada. De este modo, podrían corregirse los principales problemas de sesgo. En todo caso, el INE debería calcular los errores de muestreo de las nuevas estimaciones para evaluar la precisión.

REFERENCIAS

- ÁLVAREZ, F. (2000): «La Encuesta de Población Activa: cómo mide el mercado laboral», *Estadística Española*, vol.42, nº146, págs. 205-261.
- AZORÍN, F. Y J.L. SÁNCHEZ-CRESPO (1986): *Métodos y aplicaciones del muestreo*, ed. Alianza Universidad Textos, Madrid.
- COMISIÓN UE (2001): *Eurobarómetro 54.2*, Dirección General de Empleo y Asuntos Sociales.
- COOK, T. (1996): «Evaluation of administrative data sources for use in quarterly estimation of interstate migration between 1996 and 2001», *Demography Working Paper* 1996/1, Australian Bureau of Statistics .<http://www.abs.gov.au>
- DÍEZ HEDO, J.A. (1999): «Encuesta de migraciones», *Fuentes Estadísticas*, nº39, noviembre, págs. 14-15. <http://www.festadisticas.fguam.es>.
- GARCÍA, C., GARRIDO, L. Y L. TOHARIA (1999): «Empleo y paro en España: algunas cuestiones candentes», en Faustino Miguélez y Carlos Prieto (dirs.), *Las relaciones de empleo en España*, ed. Siglo XXI, Madrid, págs.23-50.

- GARCÍA BARBANCHO, A. (1981): *Estadística Elemental Moderna*, Ed. Ariel, Barcelona.
- INE (varios años): *Migraciones*, Madrid.
- INE (1991): *Encuesta socio-demográfica, 1991*. Tomo II, Resultados nacionales, vol. (2). Movimientos migratorios y vivienda, Madrid.
- INE (1994): *Encuesta de Población Activa*. «Tratamiento de la información», Madrid.
- INE (1999a): *Encuesta de Población Activa*. «Descripción de la Encuesta, definición de las instrucciones para la cumplimentación del cuestionario», Madrid.
- INE (1999b): *Encuesta de Población Activa*. «Tratamiento de la información», Madrid.
- INE (2000): *Encuesta de Población Activa. Primer trimestre de 2000*, nota de prensa. <http://www.ine.es>.
- INE (2001): *Encuesta de Migraciones, 1999*, Madrid.
- INE (2002): *Informe Técnico*, Madrid.
- LÓPEZ PAÑOS, R. (2000): «Estimaciones para áreas pequeñas», *Estadística Española*, vol.42, nº 146, págs. 291-338.
- PÉREZ INFANTE, J.I. (1999): «La medición del empleo en la contabilidad nacional y en la EPA: análisis de los cambios metodológicos recientes», *Cuadernos de Información Económica*, nº 152/153, págs.59-71.
- PURCELL, N. Y L. KISH (1979): «Estimation for Small Domains», *Biometrics*, 35, págs. 365-384.
- RÓDENAS, C. Y M. MARTÍ (1997): «¿Son bajos los flujos migratorios en España?», *Revista de Economía Aplicada*, vol. V, nº 15, págs.155-171.
- RÓDENAS, C. Y M. MARTÍ (2002): «Migraciones 1990-1999: ¿qué ha sucedido en la última década?», *Revista Valenciana de Economía y Hacienda*, vol. III, nº 6, págs. 37-58.
- RODRÍGUEZ OSUNA, J. (1991): *Métodos de Muestreo*, Cuadernos metodológicos, ed, CIS, Madrid.

APENDICE

TABLA A
ESTRUCTURA POR EDADES DE LA POBLACIÓN. 1999-2000

<i>Grupos de edad (años)</i>	<i>Padrón 1999</i>	<i>EPA 2 trim. 99</i>	<i>Error Relativo⁽¹⁾</i>	<i>Padrón 2000</i>	<i>EPA 2 trim. 00</i>	<i>Error Relativo⁽¹⁾</i>
0-4	4,20	4,22	0,33	4,21	4,19	-0,58
5-9	4,99	5,06	1,31	4,91	5,03	2,51
10-14	5,59	5,92	5,78	5,44	5,75	5,81
15-19	6,98	7,27	4,12	6,64	6,77	1,91
20-24	8,35	8,38	0,42	8,18	8,30	1,52
25-29	8,17	7,38	-9,62	8,33	7,50	-9,95
30-34	8,12	6,97	-14,20	8,20	7,03	-14,28
35-39	7,68	7,06	-8,07	7,87	7,19	-8,57
40-44	6,90	6,89	-0,04	7,11	6,94	-2,37
45-49	6,20	6,26	1,02	6,20	6,35	2,40
50-54	6,01	6,23	3,56	6,04	6,22	2,98
55-59	4,95	5,30	7,04	5,26	5,60	6,38
60-64	5,09	5,24	2,91	4,72	5,14	8,78
65-69	5,26	5,75	9,43	5,22	5,58	6,83
70-74	4,42	4,82	9,03	4,43	4,88	10,10
75-79	3,31	3,59	8,54	3,43	3,75	9,45
80-84	2,05	2,07	0,91	2,04	2,12	3,75
85 y más	1,72	1,59	-7,87	1,77	1,67	-6,03

(1): El error relativo es la diferencia entre la estructura derivada de la estimación de la EPA y los datos del Padrón, expresada en porcentaje de la estructura del Padrón.

Fuente: INE, Padrón (1999 y 2000), EPA (2º trimestre 1999 y 2000) y elaboración propia.

MIGRANTS AND MIGRATIONS: AGAIN THE DIVERGENCE IN THE STATISTICAL SOURCES

ABSTRACT

This paper analyzes some problems on the *Migration Survey* (MS) which INE gets yearly from the Spanish *Labour Force Survey* (LFS). Aware that the mobility that comes from this statistic source is underestimated, INE has recently incorporated some methodological changes. However, significant improvements have not been achieved and the source is still at the disposal of researchers.

The reasons underlying the limited quality of the MS are statistical, and affect the precision and bias of the estimations. Basic principles of the estimation for small domains are neglected, requirements regarding clustered samples are not fulfilled, and serious bias comes about because of a, probably, lack of response.

We consider that research into the fact of migration using stratified sampling of population would require a special treatment, given the size and characteristics of this domain.

Keywords: Migration Estimation, Labour Force Survey, Non-response and Small Domain Estimates.

AMS Classification: 62P20, 91B82, 62D05.