



UNIVERSIDAD
PABLO DE
OLAVIDE
SEVILLA



REVISTA DE MÉTODOS CUANTITATIVOS PARA
LA ECONOMÍA Y LA EMPRESA (14). Páginas 76–123.
Diciembre de 2012. ISSN: 1886-516X. D.L: SE-2927-06.
URL: <http://www.upo.es/RevMetCuant/art.php?id=63>

Alternativas metodológicas para el cálculo de las series retrospectivas regionales de EPA tras el cambio en la Clasificación Nacional de Actividades Económicas

PAVÍA, JOSÉ M.

Departamento de Economía Aplicada
Universitat de Valencia
Correo electrónico: pavia@uv.es

FABUEL, FRANCISCO

Instituto Valenciano de Estadística, Estadísticas Económicas, Demográficas y Sociales
Correo electrónico: fabuel_fra@gva.es

MORILLAS, FRANCISCO G.

Departamento de Economía Aplicada
Universitat de Valencia
Correo electrónico: Francisco.Morillas@uv.es

RESUMEN

A fin de solventar la potencial pérdida de información que la entrada en vigor de la nueva Clasificación Nacional de Actividades Económicas provocaría en las series de EPA, el INE ha propuesto una metodología que permite el enlace de las series CNAE 1993 y CNAE 2009. El procedimiento de enlace se basa en trabajar con una matriz de paso nacional sobre los datos a nivel de división. Los mercados de trabajo regionales, sin embargo, difieren del nacional y, además, las series regionales no suelen estar disponibles a nivel de división. Este trabajo estudia qué impacto tendría en las series enlazadas (i) trabajar en un nivel de agregación superior al de división y (ii) utilizar una matriz de paso regional propia. Centrada en las series de ocupados publicadas para la Comunidad Valenciana, la investigación muestra que: (a) a nivel sectorial, es equivalente emplear la matriz de España o una propia y, también, el trasladar las divisiones y luego agregarlas que el trasladar directamente las series agregadas; y (b) que a nivel de rama de actividad, es preferible trabajar con una matriz propia.

Palabras clave: CNAE; Encuesta de Población Activa; INE; series temporales.

Clasificación JEL: C18; C82; J00.

MSC2010: 62P20; 62M10.

Artículo recibido el 30 de julio de 2012 y aceptado el 19 de octubre de 2012.

Alternative Approaches for Calculating the Retrospective Regional EPA Series Following the Changes in the National Classification of Economic Activities

ABSTRACT

In order to overcome the potential loss of information that the transition to the new National Classification of Economic Activities could provoke in EPA series, INE has proposed a methodology to link CNAE 1993 and CNAE 2009 series. The procedure is based on employing a national transition matrix over division data. Regional labor markets, however, differ from the national one and additionally regional series are not generally available at the division level. This paper examines the impact that in the linked series should have (i) working at a level of aggregation above division and (ii) using a regional transition matrix. Focusing employment series published for the Valencian region, this research shows that (a), at the sectorial level, results do not depend on use of either the Spanish transition matrix of Spain or the Valencian one and also they are independent of the order in which aggregation of the divisions takes place and that (b), at the branch of activity, it is preferable to work with an own regional matrix.

Keywords: CNAE; active population survey; INE; time series.

JEL classification: C18; C82; J00.

MSC2010: 62P20; 62M10.



1. Introducción

En el año 2009 entró en vigor la nueva Clasificación Nacional de Actividades Económicas, CNAE-2009, en sustitución de la, hasta entonces, vigente Clasificación, CNAE-1993. Como consecuencia de ello, se ha producido una “rotura” en las series de la Encuesta de Población Activa (EPA) publicadas en España. Para el enlace de las series, el Instituto Nacional de Estadística (INE) propone utilizar “una matriz de conversión, calculada a partir de la doble codificación de la actividad del establecimiento del empleo principal para las personas ocupadas en alguno de los cuatro trimestres de 2008”, donde la codificación se ha realizado utilizando la “desagregación de dos dígitos en ambas clasificaciones” (INE, 2009a). Así mismo, el INE recomienda enlazar tanto las series de ramas como las de sectores a partir de transferir y agregar sus divisiones (a dos dígitos).

Cuando se desciende al ámbito regional, sin embargo, el analista se encuentra con un doble problema: (a) una estructura del mercado de trabajo regional no coincidente con la del conjunto de España; y (b) una falta de disponibilidad de las series desagregadas regionales a nivel de división. Es decir, por una parte, ocurre que la utilización de la matriz de España puede provocar una distorsión (quizás significativa) en la serie enlazada; mientras que, por otro lado, se observa que para aplicar estrictamente la metodología del INE se debería reconstruir (a partir de las oleadas de microdatos de la encuesta) las series a nivel de división, lo que plantearía, una vez disponibles los datos¹, nuevas complicaciones. Para solucionar la primera de las dificultades, el analista podría optar por construir, replicando la metodología del INE, una matriz de transferencia regional propia; mientras que para solventar el segundo de los inconvenientes el investigador podría colapsar, al nivel de desagregación publicado en cada región, la matriz de transferencia y utilizar la matriz colapsada para enlazar las series agregadas disponibles.

Cada una de las aproximaciones anteriores presenta sus ventajas e inconvenientes. El presente trabajo analiza, centrándose en las series de ocupados de EPA publicadas por el Instituto Valenciano de Estadística (IVE) para la Comunidad Valenciana, el impacto del cambio de clasificación y estudia las consecuencias de emplear cada una las soluciones comentadas en el párrafo anterior. En concreto, tomando como base los datos de las series trimestrales de ocupados de la EPA publicadas por el IVE en su página web para los períodos 1999.I-2008.IV en la clasificación CNAE-1993 y 2008.I-2009.IV en CNAE-2009, se analizan las diferencias (en niveles y en evolución subyacente) entre las series que se obtendrían con cada una de las estrategias anteriores. En particular, en la sección cuarta se analizan las diferencias que se observan enlazando directamente las series agregadas, mientras que en la sección quinta se comparan las series que se obtienen utilizando la matriz de España con aquellas que se logran empleando una matriz de transferencia propia de la Comunidad Valenciana. Las secciones segunda y tercera ofrecen detalles de cómo se obtuvieron las matrices colapsadas y la matriz de paso propia de la Comunidad Valenciana. Por último, la sección sexta cierra el estudio. Varios anexos completan el documento.

¹ El INE ofrece, previo pago, la posibilidad de adquirir los ficheros de microdatos anonimizados de la EPA codificados a tres dígitos (es decir, por grupos en lugar de divisiones).

2. Matrices de transferencia por ramas y sectores

A partir de las oleadas de la EPA implementadas por el INE durante los cuatro trimestres de 2008, en que las respuestas de los encuestados fueron codificadas utilizando las clasificaciones CNAE-1993 y CNAE-2009, el INE ha construido un fichero, que ofrece en su página web², donde presenta la matrices de paso de la clasificación de actividades CNAE-1993 a CNAE-2009 y viceversa. Las matrices de paso que provee el INE han sido elaboradas utilizando las divisiones a dos dígitos de ambas clasificaciones (60 divisiones en el caso de la CNAE-1993 y 88 divisiones para CNAE-2009) a partir de, como el propio INE señala: (i) la doble codificación, en términos de la CNAE-1993 y la CNAE-2009, que se obtuvo en la EPA durante los cuatro trimestres de 2008 de la actividad del establecimiento en que trabajan los ocupados; y (ii) la correspondencia teórica entre ambas clasificaciones (para fines estadísticos); accesible vía internet.³

A partir de estas matrices de paso o transferencia, el INE ha calculado en la nueva clasificación, CNAE-2009, las series que corresponden al período comprendido entre el primer trimestre de 2000 y el cuarto de 2007,⁴ empleando para ello la metodología descrita, muy sucintamente, en INE (2009a).⁵ En concreto, el INE recomienda trabajar con las divisiones a dos dígitos para realizar el paso de una a otra CNAE y posteriormente obtener los valores por ramas o sectores mediante agregación de los valores obtenidos tras la transferencia (estrategia trasladar-agregar).

Las variables que se publican a nivel regional, como las que publica el IVE, sin embargo, no suelen estar disponibles a dos dígitos, por lo que, en general, utilizando exclusivamente la información publicada no se puede seguir la recomendación del INE⁶. Para enlazar las series en estas circunstancias, es preciso realizar una explotación específica de los ficheros de microdatos correspondientes a cada uno de los trimestres y construir las series a dos dígitos, o bien “colapsar” las matrices a ramas de actividad o sectores, como paso previo para trasladar/transferir las cifras observadas de una a otra CNAE (es decir, aplicar la estrategia agregar-trasladar).

Si bien cuando se trabaja con matrices de transferencia las operaciones de trasladar y agregar son habitualmente conmutativas y pueden ser planteadas en cualquier orden, en este caso concreto, debido a la naturaleza dinámica del fenómeno a estudiar, la conmutatividad de tales operaciones no se mantiene. Para que tal propiedad se conservase, sería preciso disponer de la matriz de transferencia correspondiente a cada trimestre y aplicar en cada período la que correspondiese. En tal sentido, cabe esperar que la aplicación de la estrategia agregar-trasladar produzca algunas desviaciones (especialmente a medida que nos alejamos del período de referencia de 2008) respecto de los valores que se habrían obtenido si se hubiese

² INE (2009b), http://www.ine.es/daco/daco42/daco4211/matriz_conversion_retro0009.xls

³ INE (2008a), http://www.ine.es/daco/daco42/clasificaciones/cnae09/cnae93rev1_cnae2009.xls

⁴ http://www.ine.es/daco/daco42/daco4211/epa_retro0009.xls

⁵ INE (2009a), http://www.ine.es/daco/daco42/daco4211/nota_epa_retro0009.pdf

⁶ Las cifras que ofrece el IVE en su página web no están disponibles a tal nivel de detalle, por lo que la estrategia trasladar-agregar recomendada por el INE no puede ser aplicada directamente en el caso de la Comunidad Valenciana. Gracias a la colaboración de los responsables del Servicio de Estadísticas Económicas, Demográficas y Sociales del IVE, no obstante, sí fue posible seguir, sin coste económico adicional, la recomendación del INE para las series de ocupados.

podido seguir la estrategia trasladar-agregar del INE. A pesar de lo anterior, la dificultad (si no imposibilidad) que existe en multitud de situaciones de aplicar el procedimiento recomendado provoca que la cuestión relevante se trasponga a evaluar la validez y limitaciones de la estrategia agregar-trasladar. Tal cuestión se aborda en el epígrafe cuarto: “¿Agregar y trasladar o trasladar y agregar?”.

En aras de la completitud, no obstante, es preciso notar también que la estrategia recomendada trasladar-agregar no está libre de problemas ya que como consecuencia de la creciente necesidad de comparación internacional y de adecuación a las nuevas realidades económicas no existe una correspondencia biunívoca ni a nivel de clases ni de secciones entre ambas clasificaciones. Como el propio INE (2009a) señala: “*la correspondencia entre ambas clasificaciones, para aquellas agrupaciones en las que no existe identidad conceptual en su contenido (es decir, cuando la correspondencia no es biunívoca), tiende a ser menos robusta cuanto más se separa del periodo en el que la doble codificación ha sido realizada (año 2008 en este caso), ya que la propia estructura de actividad de la economía española es dinámica.*” A priori, por tanto, cabrá esperar mayores discrepancias en aquellas agrupaciones en las que más cambios se hayan producido⁷. A modo de resumen cuantitativo se puede indicar que la CNAE-93 consta de 512 clases de las cuales 277 se subdividen en 2 o más clases de la CNAE-2009; mientras que la CNAE-2009 está compuesta por 629 clases, de las cuales 274 participan de 2 o más clases de la CNAE-93 (INE, 2008b). A grandes rasgos se puede indicar que la CNAE-2009 ha aumentado el nivel de detalle con el objetivo de recoger la mayor importancia que la sociedad ha dado a las actividades relacionadas con el medio ambiente y con la finalidad de capturar mejor el impacto que tienen las actividades relacionadas con las nuevas tecnologías, especialmente internet. De ahí que, como se mostrará en epígrafes posteriores, sea, por ejemplo, en la agrupación J “*Información y Comunicación*” donde se observen algunas de las mayores discrepancias al pasar de una a otra CNAE.

Para colapsar las matrices, estas deben venir expresadas en valores absolutos o en tanto por uno (que la suma de todos los coeficientes de la matriz sea unitario). Las matrices elaboradas y puestas a disposición del público por el INE, sin embargo, se ofrecen *fila estandarizadas* (la suma de los coeficientes de cada fila es la unidad⁸), lo que dificulta el proceso de agregación, aunque en contrapartida permite una fácil interpretación de los coeficientes. En concreto, en su expresión fila estandarizada el valor a_{ij} de una matriz es interpretado como el porcentaje de individuos clasificados en la categoría i -ésima en la clasificación de origen que son reclasificados en la categoría j -ésima en la clasificación de destino.

Afortunadamente, sin embargo, el INE ofrece también el peso relativo que tiene, en cada clasificación, cada una de las divisiones a dos dígitos, permitiendo con ello que se puedan transformar las matrices fila estandarizadas en matrices de coeficientes de suma global unitaria. En concreto, si denotamos por ω_i al peso, en tanto por uno, correspondiente a la i -ésima división a dos dígitos de la clasificación de origen y por $\mathbf{A} = (a_{ij})$ a la matriz fila estandarizada (en tanto por uno) que permite pasar

⁷ En <http://www.ine.es/daco/daco42/clasificaciones/cnae09/cnae2009.pdf> se puede consultar un resumen gráfico de la correspondencia entre secciones.

⁸ En realidad se ofrecen en tanto por cien, por lo que la suma de las filas es 100.

de la clasificación de origen a la de destino, la matriz de coeficientes, **C**, se obtendría mediante $\mathbf{C} = (a_{ij} * \omega_i)$.

Para hacer más ágil la exposición, a partir de este momento representaremos mediante:

- $\mathbf{A}_{n \times m}$ a la matriz fila estandarizada de orden $n \times m$ que permite pasar las cifras de CNAE-1993 a CNAE-2009;
- $\mathbf{B}_{m \times n}$ a la matriz fila estandarizada de orden $m \times n$ que permite pasar las cifras de CNAE-2009 a CNAE-1993;
- $\mathbf{C}_{n \times m}$ a la matriz de coeficientes (de suma global unitaria) de orden $n \times m$ que permite pasar las cifras de CNAE-1993 a CNAE-2009;
- $\mathbf{D}_{m \times n}$ a la matriz de coeficientes (de suma global unitaria) de orden $m \times n$ que permite pasar las cifras de CNAE-2009 a CNAE-1993;
- $\omega_{1 \times 60}$ al vector columna de orden 1×60 de ponderaciones de las divisiones a dos dígitos en la CNAE-1993; y,
- $\varpi_{1 \times 88}$ al vector de orden 1×88 de ponderaciones de las divisiones a dos dígitos en la CNAE-2009.

De forma que, con la notación anterior, para obtener las matrices de coeficientes entre divisiones no hay más que calcular:

$$\mathbf{C}_{60 \times 88} = (\mathbf{i}_{88} \otimes \omega_{1 \times 60}) . * \mathbf{A}_{60 \times 88}$$

$$\mathbf{D}_{88 \times 60} = (\mathbf{i}_{60} \otimes \varpi_{1 \times 88}) . * \mathbf{B}_{88 \times 60}$$

donde \mathbf{i}_{60} e \mathbf{i}_{88} son vectores fila de unos de órdenes respectivos 60 y 88, \otimes es el símbolo del producto de Kronecker y “.” denota el producto de matrices elemento a elemento.

Teóricamente, se debería verificar que para todos los órdenes la traspuesta de $\mathbf{C}_{n \times m}$ fuese igual a $\mathbf{D}_{m \times n}$. Es decir, que, por ejemplo, $(\mathbf{C}_{60 \times 88})^T = \mathbf{D}_{88 \times 60}$. Sin embargo, debido al hecho de que las matrices y los vectores de ponderaciones ofrecidos por el INE aparecen con solo seis decimales (en tanto por uno) y no con todos los dígitos significativos, los errores de redondeo que esto introduce provoca que no se verifique la igualdad teórica. La diferencia máxima entre los coeficientes de las matrices $(\mathbf{C}_{60 \times 88})^T$ y $\mathbf{D}_{88 \times 60}$ es, sin embargo, de escasa cuantía. En concreto, en valor absoluto, la diferencia máxima es de $7.3486 * 10^{-7}$, lo que equivale a una desviación, en el coeficiente correspondiente, de poco más del 0.02%. Así, se ha optado por utilizar (y colapsar) la matriz $\mathbf{C}_{60 \times 88}$ para pasar los valores clasificados en CNAE-1993 a CNAE-2009 y la matriz $\mathbf{D}_{88 \times 60}$ para pasar de CNAE-2009 a CNAE-1993. Esta decisión implica que si los valores observados para un trimestre se pasasen primero de CNAE-1993 a CNAE-2009 y, posteriormente, se volviesen a pasar a CNAE-1993 se producirían pequeñas (insignificantes) desviaciones comparándolos con los valores originales.

Una vez construidas las matrices de transferencia a nivel de división ($\mathbf{C}_{60 \times 88}$ y $\mathbf{D}_{88 \times 60}$), se calcularon las correspondientes matrices de paso a nivel de rama de actividad y de sector. Para ello se construyeron oportunas matrices de diseño que permiten agregar los coeficientes de las matrices que corresponden a una misma rama (o sector). En concreto, empleando los datos de la Tabla A.I (Anexo I), que detallan las

divisiones a dos dígitos que pertenecen en la clasificación CNAE-1993 a cada una de las ramas de actividad utilizadas por el IVE en sus publicaciones sobre la EPA, y la información de la Tabla A.II, que concreta las divisiones a dos dígitos que integran cada una de las ramas de actividad en la clasificación CNAE-2009 se construyeron las matrices $\mathbf{M}_{15 \times 60}$ y $\mathbf{N}_{17 \times 80}$, que permiten colapsar, agregando adecuadamente, las matrices a dos dígitos a matrices por ramas de actividad. En concreto, utilizando la ordenaciones obvias que se derivan de las Tablas A.I y A.II para las ramas de actividad, se definieron los coeficientes m_{ij} y n_{ij} de las matrices $\mathbf{M}_{15 \times 60} = (m_{ij})$ y $\mathbf{N}_{17 \times 88} = (n_{ij})$ como: $m_{ij} = 1$ si la j -ésima división de la clasificación CNAE-1993 pertenece a la rama i -ésima y $m_{ij} = 0$ en otro caso; y $n_{ij} = 1$ si la j -ésima división de la clasificación CNAE-2009 pertenece a la rama i -ésima y $n_{ij} = 0$ en otro caso.

Con las definiciones anteriores, las matrices de transferencia de coeficientes entre ramas de actividad ($\mathbf{C}_{15 \times 17}$ y $\mathbf{D}_{17 \times 15}$) se obtuvieron fácilmente a través del siguiente producto de matrices:

$$\mathbf{C}_{15 \times 17} = \mathbf{M}_{15 \times 60} \mathbf{C}_{60 \times 88} (\mathbf{N}_{17 \times 88})^T$$

$$\mathbf{D}_{17 \times 15} = \mathbf{N}_{17 \times 88} \mathbf{D}_{88 \times 60} (\mathbf{M}_{15 \times 60})^T$$

A partir de las matrices por ramas de actividad se obtuvieron las matrices por sectores agregando adecuadamente las ramas de actividad. En concreto, teniendo en cuenta que con la clasificación CNAE-1993 las ramas se agrupan en A+B, C+D+E, F y G+...+T y que con la clasificación CNAE-2009 se agrupan en A, B+D+E+C, F y G+...+T, se construyeron, empleando la lógica descrita para la elaboración de \mathbf{M} y \mathbf{N} , nuevas matrices de diseño $\mathbf{P}_{4 \times 15}$ y $\mathbf{Q}_{4 \times 17}$, de donde se derivaron las matrices de coeficientes de paso por sectores:

$$\mathbf{C}_{4 \times 4} = \mathbf{P}_{4 \times 15} \mathbf{C}_{15 \times 17} (\mathbf{Q}_{4 \times 17})^T$$

$$\mathbf{D}_{4 \times 4} = \mathbf{Q}_{4 \times 17} \mathbf{D}_{17 \times 15} (\mathbf{P}_{4 \times 15})^T$$

Finalmente, y a partir de las matrices anteriores, se obtuvieron las correspondientes matrices, $\mathbf{A}_{15 \times 17}$ y $\mathbf{A}_{4 \times 4}$ y $\mathbf{B}_{17 \times 15}$ y $\mathbf{B}_{4 \times 4}$, fila estandarizadas. A modo de ejemplo, en las Tablas I y II se ofrece, en tanto por cien, las matrices de transferencia $\mathbf{A}_{4 \times 4}$ y $\mathbf{B}_{4 \times 4}$ obtenidas. De los coeficientes obtenidos se deduce, por ejemplo, que al 6.83% de los ocupados que son clasificados en el sector Agrícola en la CNAE-1993 corresponde clasificarlos en la CNAE-2009 en el sector Servicios.

Tabla I. Matriz de transferencia (fila estandarizada) entre clasificaciones CNAE-1993 y CNAE-2009 para los principales sectores económicos, $\mathbf{A}_{4 \times 4}$.

CNAE-1993	CNAE-2009			
	Agricultura	Industria	Construcción	Servicios
Agricultura	93.1559	0.0117	0.0000	6.8324
Industria	0.0000	97.0984	0.0917	2.8099
Construcción	0.0000	0.0000	99.9363	0.0637
Servicios	0.0003	0.5025	0.3335	99.1637

Fuente: elaboración propia a partir de INE (2009b). Nota: coeficientes en porcentaje.

Tabla II. Matriz de transferencia (fila estandarizada) entre clasificaciones CNAE-2009 y CNAE-1993 para los principales sectores económicos, $\mathbf{B}_{4 \times 4}$.

CNAE-1993	CNAE-2009			
	Agricultura	Industria	Construcción	Servicios
Agricultura	99.9947	0.0000	0.0000	0.0053
Industria	0.0032	97.8373	0.0000	2.1595
Construcción	0.0000	0.1205	98.0106	1.8689
Servicios	0.4356	0.6570	0.0111	98.8963

Fuente: elaboración propia a partir de INE (2009b). Nota: coeficientes en porcentaje.

3. Matrices de transferencia de la Comunidad Valenciana

El INE recomienda utilizar las matrices de transferencia de España (construidas con las respuestas recogidas en la EPA para el conjunto del país) para pasar los datos entre clasificaciones. Esto permite que la herramienta utilizada en todos los casos sea la misma y que, en tal sentido, las series que pueden corresponder a diferentes ámbitos geográficos sean comparables. No es menos cierto, sin embargo, que la estructura del mercado de trabajo varía en el espacio (y también en el tiempo⁹) por lo que la aplicación de una matriz fija, la de España, a todos los territorios podría provocar importantes desviaciones sobre lo que deberían haber sido los valores realmente observados. El sesgo que este hecho puede introducir tendría escaso significado económico y únicamente respondería a la circunstancia de haber utilizado una estructura de reparto que no se ajusta adecuadamente a la estructura subyacente de las series estudiadas.

Para poder valorar cuál sería el impacto de utilizar la estrategia recomendada por el INE a los datos de la Comunidad Valenciana sería preciso construir una matriz de transferencia propia y analizar si las diferencias que se obtienen son o no significativas. En este apartado, se aborda la tarea de la construcción de tal matriz. Utilizando los microdatos de la base de datos depurada de las oleadas de la EPA de 2008 correspondientes al ámbito de la Comunidad Valenciana se construyó, replicando la metodología seguida por el INE para construir la matriz de España, una matriz de transferencia propia de la Comunidad Valenciana y se enlazaron las series utilizando tal matriz.

En concreto, para el cálculo de las matrices de coeficientes ${}_{CV}C_{n \times m}$ (y ${}_{CV}D_{n \times m}$) de la Comunidad Valenciana se construyeron las matrices de factores de elevación acumulados, $(FE)_{88 \times 60}$, $(FE)_{17 \times 15}$ y $(FE)_{4 \times 4}$, utilizando la correspondencia entre las clases en CNAE-2009 y en CNAE-1993 (y por trasposición de CNAE-1993 a CNAE-2009). El cálculo de los elementos de estas matrices no es complejo, aunque debido al elevado número de observaciones que hubo que tratar (más de 24 mil) se realizó de forma automática mediante un *software ad-hoc*. En concreto, denotando por $FE_{k, 88 \times 60} = (fe_{ij})_{k, 88 \times 60}$ a la matriz obtenida con factores de elevación en el trimestre k-ésimo para el paso entre divisiones de CNAE-2009 a CNAE-1993 (donde $(fe_{ij})_{k, 88 \times 60}$ representa la suma

⁹ Como se comentó previamente, el fenómeno estudiado es dinámico, lo que provoca que las operaciones de agregar y trasladar no sean conmutativas. En este caso, sin embargo, la variación en el tiempo será menos intensa que en el espacio. Es decir, entre dos trimestres (años) consecutivos los cambios en la estructura del mercado de trabajo de un territorio varía, por término medio, menos de lo que varía la estructura de trabajo entre territorios en un mismo periodo temporal.

de todos los factores de elevación correspondientes a la equivalencia entre la clase i en CNAE-2009 y la clase j en CNAE-1993, para los datos correspondientes al trimestre k), se tiene que una vez obtenidas las matrices $FE_{k,88 \times 60}$ y $FE_{k,60 \times 80}$, para $k=1,2,3$ y 4 ,¹⁰ se obtuvieron las matrices $FE_{k,17 \times 15}$ y $FE_{k,4 \times 4}$ (y $FE_{k,15 \times 17}$ y $FE_{k,4 \times 4}$), y de ahí, por agregación, las matrices anuales $FE_{a,n \times m}$, cuyo coeficiente genérico (i,j) es:

$$(fe_{ij})_{a,n \times m} = \sum_k (fe_{ij})_{k,n \times m}$$

La información obtenida en estas matrices $FE_{a,n \times m}$ es la que permite el cálculo de las matrices de coeficientes (${}_{CV}C_{n \times m}$ y ${}_{CV}D_{m \times n}$) y fila estandarizadas (${}_{CV}A_{n \times m}$ y ${}_{CV}B_{m \times n}$) de la Comunidad Valenciana. Para la obtención de las matrices de coeficientes no hay más que dividir cada elemento de la matriz correspondiente $FE_{a,n \times m}$ por la suma total de todos sus elementos; mientras que para obtener las matrices fila estandarizadas basta con dividir cada elemento por la suma total de los elementos de su misma fila.

4. ¿Agregar y trasladar o trasladar y agregar?

El Instituto Nacional de Estadística (INE) construye la matriz de transferencia entre clasificaciones CNAEs a partir de las divisiones a dos dígitos (60 en el caso de la CNAE-1993 y 88 para la CNAE-2009) y recomienda construir los valores correspondientes a la variable de interés a partir de trasladar primero las cifras de las divisiones a dos dígitos que lo componen, para en segunda etapa obtener por agregación los valores de la variable de interés.

En concreto, imaginemos que x_{93} representa el vector 4×1 de valores observados para, por ejemplo, el número de ocupados en agricultura, industria, construcción y servicios en el segundo trimestre de 2006 con la clasificación CNAE-1993 y denotemos por y_{93} al vector 60×1 del número de ocupados registrados, también en el segundo trimestre de 2006, para las 60 divisiones de la clasificación CNAE-1993.¹¹ En estas condiciones la recomendación del INE para obtener los valores en la clasificación CNAE-2009 que corresponderían a x_{93} , que notaremos como x_{09} , consiste en obtener en primer lugar (utilizando la matriz de paso a nivel de divisiones) el vector 88×1 , y_{09} , del número de ocupados para las 88 divisiones de la clasificación CNAE-2009 y de ahí, agregando, obtener el vector 4×1 x_{09} . Lo que, si utilizamos la notación matricial introducida en la sección segunda, se expresaría mediante:

$$x_{09} = Q_{4 \times 17} N_{17 \times 88} y_{09} = Q_{4 \times 17} N_{17 \times 88} A_{88 \times 60} y_{93}$$

El problema estriba en que los valores correspondientes a divisiones no suelen estar disponibles (como ocurre con los valores publicados en la página web del IVE), por lo que salvo que se haga una explotación a medida de los ficheros de microdatos que dieron origen a las series publicadas, tales variables no podrían ser trasladadas utilizando la metodología recomendada por el INE. En esta tesitura, el analista puede optar, como se indicó en la sección segunda, por aproximar sus series colapsando las matrices de transferencia y trasladando las series ya agregadas. Es decir, por aplicar, para el ejemplo del párrafo previo, la solución:

$$x_{09} = B_{4 \times 4} x_{93} = B_{4 \times 4} P_{4 \times 15} N_{15 \times 60} y_{93}$$

¹⁰ Aunque estas matrices son traspuestas, el cálculo se realizó de forma independiente para reducir las posibilidades de error en el programa construido.

¹¹ Es decir, se verifica que $x_{93} = P_{4 \times 15} N_{15 \times 60} y_{93}$.

Si para cada trimestre fuese obtenida y aplicada la matriz de transferencia que correspondiese a ese trimestre, ambas estrategias serían equivalentes y originarían las mismas soluciones. La cuestión aquí es que se utiliza una matriz fija (la obtenida como media de los trimestres de 2008) y, por consiguiente, ambas aproximaciones generan soluciones diferentes. Ahora bien, ¿hasta qué punto son diferentes las soluciones?, ¿son preocupantes las diferencias?, ¿son significativas las diferencias de seguir una u otra estrategia?, ¿pueden ser consideradas como 'válidas' las soluciones aproximadas obtenidas aplicando la estrategia agregar-trasladar? Es más, considerando que los datos observados, al haber sido obtenidos mediante una encuesta, son ya una aproximación a la evolución real subyacente del fenómeno, cabría preguntarse: ¿Cambian los perfiles cíclicos y los puntos de giro de las series?

Afortunadamente, gracias a la explotación de los ficheros de microdatos históricos de las series de ocupados publicadas en la web de IVE, en el presente estudio ha sido posible ofrecer algunas respuestas a las preguntas anteriores. En concreto, para el período 1999.I-2008.IV, las 17 series de ocupados por rama de actividad de la Comunidad Valenciana y las 4 series de ocupados por sector y ámbito geográfico (Comunidad Valenciana, Alicante, Castellón y Valencia) en la clasificación CNAE-1993 fueron trasladadas a CNAE-2009 utilizando ambas estrategias. En total, por tanto, se dispone de 33 series (17+16) para la comparación, para las cuales además hay disponibles tres conjuntos de estimaciones para 2008 (el que se observó clasificando directamente los microdatos en CNAE-2009 y los dos obtenidos trasladando desde CNAE-1993 las series agregadas y por divisiones).

Prestando especial atención a las diferencias para las series del ámbito de la Comunidad Valenciana, en primer término se compararán las series en niveles y a continuación se realizará un análisis de los perfiles cíclicos y los puntos de giro que dibujan. En todo caso, y como referente, se utilizarán las series obtenidas tras aplicar la estrategia trasladar-agregar, ya que es la recomendada por el INE. Asimismo, y para completar los análisis, se compararán los niveles de los valores realmente observados en 2008 en CNAE-2009 con los correspondientes valores trasladados.

4.1. Comparación en niveles

A fin de hacer un análisis más sistemático, en primer término se realizarán los análisis por ramas de actividad para, posteriormente, abordar el estudio por sectores. Dentro de cada grupo, además, primero se realizarán las comparaciones entre las series recodificadas, para finalmente comparar ambos conjuntos de valores con las observaciones directas recolectadas en 2008.

En todo este subapartado, no obstante, se ha utilizado un mismo estadístico de comparación: la diferencia absoluta relativa en porcentaje, D_r ; la cual, si notamos por x al valor de referencia y por y al valor asociado de comparación, se ha calculado para cada trimestre y variable mediante:

$$D_r = 100 \frac{|x - y|}{x}$$

Medias anuales y globales se han calculado de los valores logrados para el estadístico D_r para realizar los análisis.

4.1.1. Ramas de actividad

En la Tabla III se ofrecen los valores medios (anuales y globales) de las diferencias relativas en valor absoluto de los valores obtenidos, para cada una de las 17 series de ocupados de la Comunidad Valenciana por ramas de actividad. Las mayores diferencias se localizan en las ramas B+D+E, J y L (ver el Anexo I, para una descripción de las ramas), siendo asimismo destacables las diferencias que se observan para las ramas C, M, N y R+S+U. Aunque en general las mayores diferencias relativas se producen para las ramas que en niveles tienen menor tamaño y, por tanto, presentan una mayor desviación relativa para una misma diferencia absoluta, lo cierto es que a la luz de las cifras anteriores, ambas estrategias no parecen equivalentes a nivel de rama de actividad, especialmente para las ramas B+D+E, J y L. Se ha de notar, no obstante, que una parte de estas diferencias son consecuencia directa de las diferentes estructuras productivas que tienen la Comunidad Valenciana y España. De hecho, estas tres ramas destacan por estar entre aquellas donde las diferencias de estructura son más evidentes (ver la sección 5). Las diferencias en todo caso son notables ya que además coincide que estas son algunas de las agrupaciones en las que más se han concentrado los cambios entre clasificaciones (INE, 2009c).

Tabla III. Diferencias relativas en valor absoluto entre las estrategias trasladar-agregar y agregar-trasladar para las series de ocupados de la Comunidad Valenciana por ramas de actividad.

Rama	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	Global
A	0.38	0.17	0.06	0.08	0.12	0.25	0.27	0.12	0.17	0.27	0.19
B+D+E	27.72	32.09	37.84	26.94	14.52	22.02	10.50	7.25	28.78	22.03	22.97
C	3.06	3.55	3.38	3.55	3.71	3.28	2.97	2.38	2.88	2.90	3.17
F	0.14	0.17	0.16	0.06	0.11	0.64	1.09	0.52	0.37	0.33	0.36
G	0.16	0.21	0.19	0.21	0.15	0.15	0.10	0.14	0.10	0.06	0.15
H	2.88	2.71	1.61	3.12	4.59	2.62	2.99	1.23	4.32	5.03	3.11
I	0.04	0.02	0.03	0.02	0.02	0.02	0.02	0.01	0.02	0.02	0.02
J	37.81	41.28	39.64	61.02	64.33	47.26	28.24	17.47	37.79	38.18	41.30
K	0.05	0.07	0.06	0.02	0.05	0.09	0.06	0.03	0.05	0.04	0.05
L	9.71	12.25	12.25	4.55	8.61	31.75	44.67	28.13	23.42	16.57	19.19
M	2.41	2.15	2.79	2.43	3.49	2.67	7.41	4.15	1.40	1.36	3.03
N	1.34	7.12	5.36	8.71	7.82	1.52	4.52	0.73	5.08	3.06	4.52
O	0.03	0.04	0.06	0.04	0.02	0.03	0.01	0.04	0.04	0.03	0.03
P	0.05	0.03	0.09	0.06	0.07	0.06	0.08	0.04	0.04	0.04	0.05
Q	0.03	0.04	0.05	0.03	0.02	0.02	0.03	0.02	0.03	0.02	0.03
R+S+U	1.23	3.73	1.10	3.00	6.31	1.70	5.94	3.40	2.93	2.46	3.18
T	0.06	0.06	0.06	0.04	0.03	0.04	0.03	0.03	0.03	0.03	0.04

Fuente: elaboración propia.

La estrategia recomendada por el INE, sin embargo, tampoco es una panacea. Como se puede deducir fácilmente observando los resultados presentados en la Tabla IV, donde se comparan los valores trasladados utilizando cada una de las estrategias (trasladar-agregar y agregar-trasladar) con las cifras efectivamente recogidas en la EPA durante 2008 en CNAE-2009. Especialmente significativas son las diferencias que se obtienen de nuevo, incluso aplicando la estrategia recomendada por el INE, para las

ramas J, L y B+D+E. En todo caso, no obstante, se puede afirmar que con carácter general y a nivel de rama de actividad es preferible utilizar la estrategia trasladar-agregar recomendada por el INE a la alternativa agregar-trasladar.

Tabla IV. Diferencias relativas en valor absoluto, para las series de ocupados de la Comunidad Valencia, de los valores obtenidos tras aplicar las estrategias trasladar-agregar y agregar-trasladar con cifras recolectadas en la EPA.

Rama	Error relativo de series obtenidas con estrategia trasladar-agregar (INE)					Error relativo de series obtenidas con estrategia agregar-trasladar				
	2008.I	2008.II	2008.III	2008.IV	2008	2008.I	2008.II	2008.III	2008.IV	2008
A	3.96	3.30	1.27	1.94	2.62	4.30	3.56	1.33	2.38	2.89
B+D+E	5.64	5.10	3.19	3.43	4.34	31.52	26.96	19.24	15.53	23.31
C	1.31	0.77	0.36	0.30	0.68	4.17	3.50	3.37	3.23	3.57
F	1.36	1.29	0.43	0.84	0.98	1.54	1.35	0.76	1.58	1.31
G	0.76	0.09	0.24	0.62	0.43	0.84	0.14	0.24	0.73	0.49
H	2.11	3.94	3.48	5.02	3.64	6.06	7.89	9.76	10.19	8.48
I	0.03	0.02	0.01	0.02	0.02	0.00	0.00	0.00	0.00	0.00
J	22.11	12.25	20.35	26.08	20.20	62.70	45.71	73.69	83.27	66.34
K	0.38	0.45	0.56	0.47	0.46	0.42	0.47	0.65	0.44	0.49
L	20.24	22.09	0.95	4.81	12.02	5.97	15.29	19.68	33.30	18.56
M	2.60	0.42	5.51	9.03	4.39	3.58	1.77	3.13	8.14	4.15
N	1.15	0.80	4.58	6.17	3.17	4.98	6.18	6.53	7.00	6.17
O	0.07	0.03	0.09	0.32	0.13	0.07	0.07	0.07	0.35	0.14
P	1.02	0.42	0.44	1.10	0.75	1.02	0.40	0.36	1.14	0.73
Q	1.09	1.32	0.94	0.58	0.98	1.09	1.35	0.91	0.56	0.98
R+S+U	3.56	0.74	1.28	0.28	1.46	6.20	4.09	0.94	1.41	3.16
T	0.03	0.05	0.02	0.01	0.03	0.00	0.00	0.00	0.00	0.00

Fuente: elaboración propia.

4.1.2. Sectores

Si la comparación por ramas de actividad deja clara que la aproximación agregar-trasladar no representa, en general, una alternativa suficientemente adecuada para estimar los valores en niveles, el panorama es muy diferente a nivel de sector. En efecto, observando los resultados mostrados en la Tabla V –donde se presentan, por sectores y para los tres ámbitos provinciales más el agregado regional, un resumen de las diferencias obtenidas (medidas a través del estadístico D_r) en el número de ocupados recodificados de CNAE-1993 a CNAE-2009– se deduce claramente que las diferencias por sectores no son muy importantes, ya que, salvo en el caso de la industria, todas ellas se mueven con discrepancias por debajo del 1% (inferiores a los niveles de error de estimación que suelen ser habituales en las encuestas). En general, por tanto, podemos considerar que a nivel de sector la estrategia aproximada (agregar-trasladar) y la estrategia recomendada por el INE (trasladar-agregar) son básicamente equivalentes. Es más, dado que la base para el cálculo de la matriz de transferencia son los valores de ocupados, pues como señala el INE se ha realizado “*tabulando la actividad principal del total de ocupados al nivel de divisiones (códigos de dos dígitos) de ambas clasificaciones*”, lo lógico es que este resultado pudiera ser extrapolado para el resto de variables publicadas.

Tabla V. Diferencias relativas en valor absoluto entre las estrategias trasladar-agregar y agregar-trasladar para las series de ocupados de la Comunidad Valenciana por sectores.

Sector	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	Global
Alicante											
Agricultura	0.36	0.28	0.18	0.14	0.24	0.51	0.50	0.11	0.30	0.36	0.30
Industria	2.45	2.63	1.18	1.10	3.16	2.43	2.65	1.97	0.85	1.18	1.96
Construcción	0.70	0.32	0.29	0.34	0.28	1.67	2.71	1.08	0.64	0.93	0.89
Servicios	1.06	0.99	0.41	0.41	1.19	1.17	1.44	0.85	0.39	0.44	0.84
Castellón											
Agricultura	0.39	0.37	0.57	0.61	0.12	0.28	0.21	0.34	0.30	0.29	0.35
Industria	2.13	1.94	1.24	1.31	2.43	2.98	1.34	1.30	1.59	2.08	1.83
Construcción	1.00	0.67	0.14	0.32	0.48	0.35	0.43	0.26	0.17	0.21	0.40
Servicios	1.03	1.04	0.88	0.90	1.42	1.85	0.62	0.65	0.74	0.82	1.00
Valencia											
Agricultura	0.39	0.46	0.32	0.21	0.16	0.37	0.42	0.22	0.26	0.68	0.35
Industria	0.58	0.25	0.50	1.33	1.51	0.35	1.40	1.20	0.99	0.88	0.90
Construcción	1.06	0.92	0.60	0.34	0.73	0.60	0.08	0.16	0.15	0.29	0.49
Servicios	0.43	0.17	0.16	0.37	0.36	0.15	0.42	0.38	0.27	0.17	0.29
Comunidad Valenciana											
Agricultura	0.38	0.17	0.06	0.08	0.12	0.26	0.27	0.12	0.17	0.27	0.19
Industria	0.90	1.23	0.78	1.24	2.20	1.32	1.77	1.47	1.04	1.16	1.31
Construcción	0.44	0.67	0.41	0.28	0.34	0.41	0.91	0.35	0.23	0.23	0.43
Servicios	0.25	0.35	0.22	0.43	0.74	0.53	0.79	0.58	0.36	0.34	0.46

Fuente: elaboración propia.

Por otro lado, analizando ahora las diferencias de los valores recodificados con los registrados en la encuesta con clasificación CNAE-2009 para 2008, ver la Tabla VI, se puede afirmar que, en general, las soluciones obtenidas siguiendo la estrategia del INE (trasladar-agregar) son ligeramente más próximas a los valores observados que las obtenidas con la aproximación agregar-trasladar; empero, ninguna solución es plenamente satisfactoria pues todavía existen diferencias, que, en este caso, se concentran principalmente en la agricultura (especialmente para la provincia de Alicante). En cualquier caso, no obstante, se puede considerar que en términos globales las soluciones se aproximan bastante a los valores estimados directamente a partir de las respuestas en la EPA.

4.2. Comparación cíclica

El análisis en niveles realizado en el subapartado anterior ha conducido a dos conclusiones básicas: (i) a nivel de ramas de actividad las estrategias trasladar-agregar y agregar-trasladar no son, en general, equivalentes y pueden conducir a diferencias significativas en niveles; sin embargo, (ii) por sectores ambas soluciones muestran resultados bastante parecidos, ya que las distancias de los resultados que se obtienen siguiendo cada estrategia no son muy apreciables. Las series de EPA, sin embargo, están sujetas a error (muestral y no muestral) ya que son estimaciones que se obtienen a partir de datos de encuesta, por lo que los valores concretos constituyen un indicador más que una medición exacta del verdadero valor que toma la variable; siendo esto especialmente cierto cuando lo que se pretende es realizar un seguimiento

coyuntural de la evolución del mercado de trabajo. La existencia de ciertas desviaciones entre valores trasladados y publicados, por tanto, no invalida necesariamente el proceso que se haya podido seguir. Desde ese punto de vista, en consecuencia, sería interesante analizar si las evoluciones que dibujan las series trasladadas con una y otra estrategia son similares o si, por el contrario, presentan diferencias significativas que impliquen que se obtienen perfiles cíclicos diferentes dependiendo de la estrategia utilizada.

Tabla VI. Diferencias relativas en valor absoluto respecto a las cifras recolectadas en EPA, para las series de ocupados de la Comunidad Valenciana y sus tres provincias, de los valores obtenidos tras aplicar las estrategias trasladar-agregar y agregar-trasladar.

Sector	Error relativo de series obtenidas con estrategia trasladar-agregar (INE)					Error relativo de series obtenidas con estrategia agregar-trasladar				
	2008.I	2008.II	2008.III	2008.IV	2008	2008.I	2008.II	2008.III	2008.IV	2008
	Alicante					Alicante				
Agricultura	19.45	7.64	2.75	4.76	8.65	19.38	7.46	1.98	5.26	8.52
Industria	0.36	1.03	1.09	0.72	0.80	1.67	0.45	0.20	1.75	1.02
Construcción	0.60	0.63	2.05	2.22	1.38	0.65	0.41	0.96	0.22	0.56
Servicios	0.50	0.46	0.78	0.43	0.54	0.14	0.11	0.32	0.14	0.18
	Castellón					Castellón				
Agricultura	1.36	1.03	1.11	1.92	1.35	1.76	0.94	1.50	1.62	1.45
Industria	0.38	0.27	0.16	1.19	0.50	1.65	2.27	2.59	1.46	1.99
Construcción	1.15	0.17	2.56	2.63	1.63	0.75	0.06	2.31	2.58	1.42
Servicios	0.32	0.24	0.73	0.20	0.37	0.72	1.04	1.76	1.27	1.20
	Valencia					Valencia				
Agricultura	4.11	1.90	0.59	0.61	1.80	3.55	2.62	1.26	1.37	2.20
Industria	1.53	1.32	1.60	0.96	1.35	2.12	2.07	2.66	2.03	2.22
Construcción	1.86	1.91	1.44	2.18	1.85	1.95	1.67	1.05	1.78	1.61
Servicios	1.04	0.73	0.78	0.66	0.80	1.20	0.87	0.97	0.85	0.98
	Comunidad Valenciana					Comunidad Valenciana				
Agricultura	3.96	3.30	1.27	1.94	2.62	4.30	3.56	1.33	2.38	2.90
Industria	0.96	0.40	0.56	0.54	0.62	1.88	1.56	1.79	1.85	1.77
Construcción	1.36	1.29	0.43	0.84	0.98	1.37	1.11	0.58	1.42	1.12
Servicios	0.41	0.25	0.19	0.20	0.26	0.67	0.54	0.57	0.63	0.60

Fuente: elaboración propia.

En la literatura económica, para analizar las similitudes y las sincronías entre los ciclos, se han venido utilizando distintos enfoques, como el estudio de las funciones de correlación cruzada o la comparación entre los puntos de giro que presentan las series a comparar. Debido a que en este trabajo no se desea analizar la sincronía-asincronía-desfase cíclico que presentan distintos conjuntos de series diferentes, sino que el objetivo es ver si dos *traslaciones* distintas de una misma variable dibujan la misma evolución cíclica, se ha optado por utilizar la segunda aproximación. Este procedimiento se centra en caracterizar los ciclos a través de sus fases de expansión y recesión, donde las fases del ciclo vendrán determinadas por los momentos concretos en los que la actividad económica cambia de signo: puntos de giro. Aunque, una vez caracterizados los ciclos, lo habitual es comparar las evoluciones cíclicas de las distintas

series examinando las duraciones y las amplitudes de las distintas fases¹²; en este caso, el análisis es más sencillo pues trabajamos con estimaciones (traslaciones) alternativas de una misma serie, por lo que bastará con estudiar si se detectan o no los mismos puntos de giro.

A pesar de la gran atención que ha recibido el estudio de los ciclos económicos en la literatura estadístico-económica, no parece existir, sin embargo, un consenso entre los estudiosos de la materia sobre (i) cuáles son los métodos que deben implementarse para obtener estimaciones fiables de la evolución cíclica de una serie, (ii) cómo debería definirse el ciclo y (iii) respecto al modo de medirlo. En este trabajo seguiremos la propuesta más extendida que admite que el ciclo de una serie se define a partir de las desviaciones que registra esta sobre su evolución subyacente. Por lo que, para estimar el ciclo, como punto de partida, se debe extraer de una serie su evolución subyacente; para, en segunda etapa, calcular las desviaciones o fluctuaciones que esta registra.¹³

En concreto y de acuerdo con la recomendación anterior, en este trabajo, para extraer la componente cíclica se han utilizado las tasas de variación interanuales de las señales de ciclo-tendencia de las series¹⁴. Señales que han sido obtenidas utilizando el programa TSW¹⁵ que implementa una metodología de extracción basada en la modelización ARIMA de la serie y en la asignación de las raíces asociadas a los polinomios al componente inobservado (tendencia, ciclo, estacionalidad y componente irregular) al que teóricamente corresponderían.

Para elaborar el fechado de la señal (es decir, para determinar los momentos o fechas concretas en las que se produce un cambio de fase en la señal cíclica), sin embargo, no pueden ser considerados todos los máximos y mínimos que dibuje la componente cíclica, ya que no todos ellos responden a la idea de cambios en la fase del ciclo¹⁶. Con el objetivo de eliminar estos ciclos espurios y, asimismo, garantizar ciclos de una duración adecuada, en la literatura se han propuesto varios criterios de identificación de puntos de giro¹⁷. Sin entrar en detalle de las distintas propuestas y en el grado de la arbitrariedad que acompaña a todas ellas, en este trabajo se ha considerado que para que un máximo o mínimo relativo pueda ser identificado como punto de giro se debe verificar que (i) la distancia entre dos puntos de giro del mismo signo (duración del ciclo) ha de ser superior a 15 meses (es decir, debe existir una distancia mínima de 6 trimestres), (ii) que al mismo tiempo que la distancia entre dos puntos de giro de distinto signo (duración de las fases) ha de ser superior a 5 meses (es decir, debe existir una distancia mínima de 2 trimestres), y que (iii) debe además existir entre dos puntos de giro del mismo signo un punto de giro de signo contrario.

¹² Véase, por ejemplo, Niemira y Klein (1994).

¹³ Véase, por ejemplo, Kaiser y Maravall (1999).

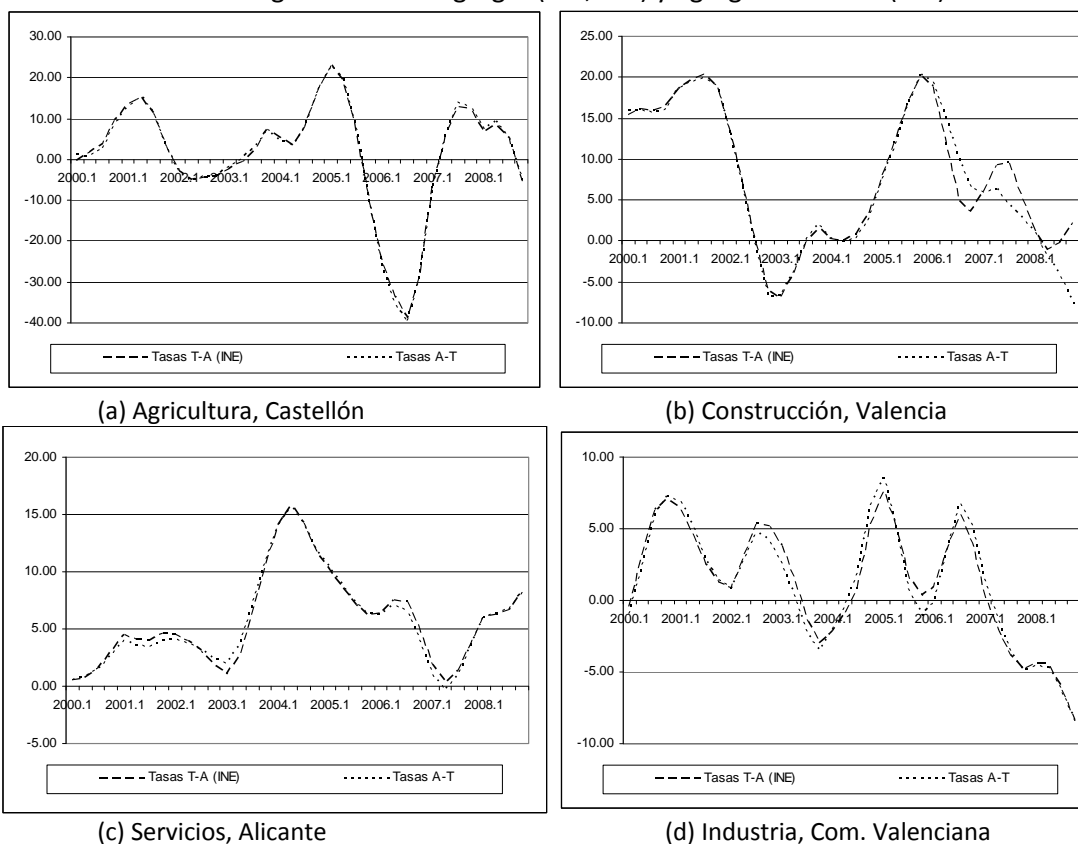
¹⁴ Aunque existen una gran variedad de procedimientos se ha optado por este por ser una de las formas más sencillas, intuitivas y extendidas de extraer el componente cíclico de una serie.

¹⁵ TRAMO-SEATS para Windows en su versión Beta 1.0.4, Rev. 136, de 16 de marzo de 2009, tras aplicar las opciones por defecto (con RSA=4) a las 33 series de ocupados recodificadas para el período 1999.I a 2008.IV.

¹⁶ Los filtros utilizados para extraer la señal cíclica no tienen capacidad para eliminar todos los puntos de giro espurios (máximos y mínimos inscritos en una clara trayectoria de aceleración o desaceleración) que se pueden detectar.

¹⁷ Véase, por ejemplo, Stekler (1991).

Figura 1. Ejemplo de comparaciones de evoluciones cíclicas (medidas en tasas de variación interanual sobre señales de ciclo-tendencia) de diversas series recodificadas utilizando las estrategias trasladar-agregar (T-A, INE) y agregar-trasladar (A-T).



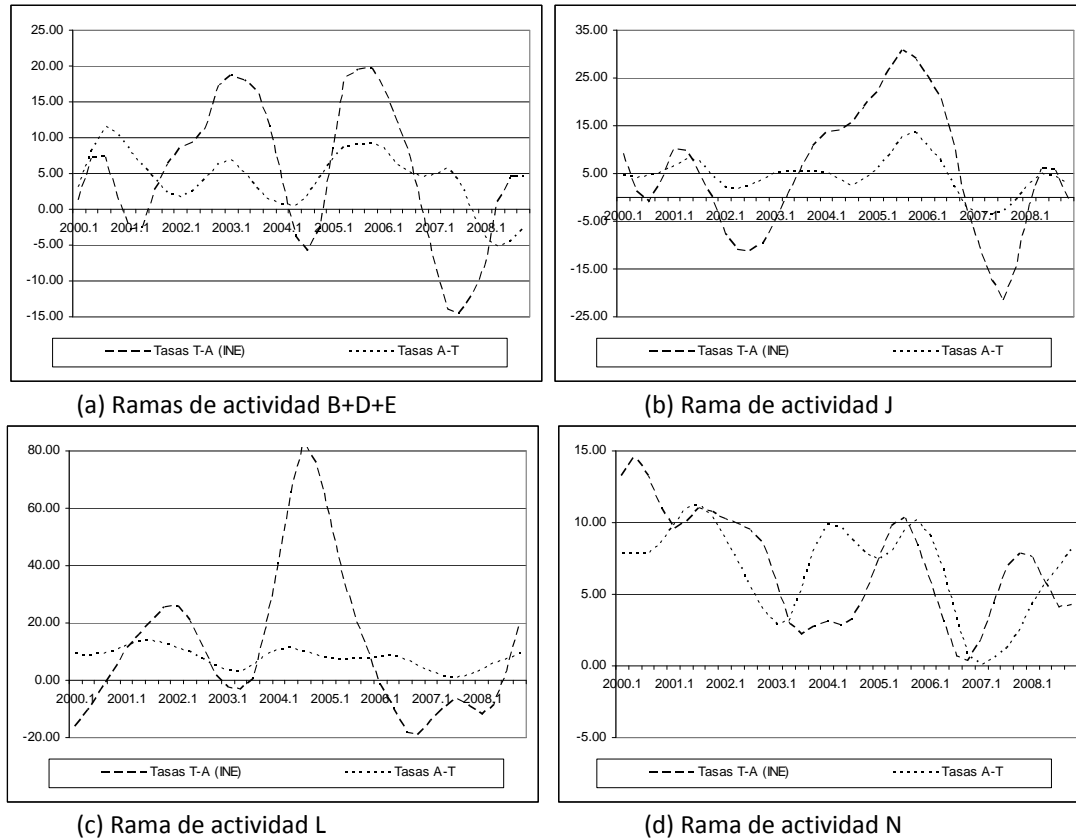
Las señales cíclicas de las series de ocupados sectoriales (agricultura, industria, construcción y servicios) obtenidas usando ambas estrategias (como ejemplo, en la Figura 1 se muestran las señales cíclicas de algunas de las series sectoriales recodificadas utilizando ambas estrategias) no presentan, en general, y como era de esperar tras las conclusiones alcanzadas en el epígrafe anterior, grandes diferencias. La excepción se encuentra en las evoluciones obtenidas para las series de construcción de Valencia, que divergen al final del período (ver la Figura 1b). Salvo para la serie anterior, todas las series¹⁸ muestran perfiles cíclicos muy similares y puntos de giro en los mismos instantes. Este resultado, sin duda, viene a confirmar y reforzar la conclusión de que a nivel sectorial ambas estrategias trasladar-agregar y agregar-trasladar son casi equivalentes.

Por otro lado, cuando el análisis se realiza sobre las señales cíclicas de las series por ramas de actividad las conclusiones son claramente diferentes. Si bien para la mayoría de ramas (A, C, F, G, I, K, O, P y Q) las señales cíclicas que se obtienen, con ambas estrategias, son indistinguibles y para otro grupo de ellas (H, M, R+S+U y T) son bastante similares, hay un conjunto de ramas (entre las que se encuentran las que mayores trasvases sufren en las clasificaciones de divisiones a dos dígitos) para las que

¹⁸ En el caso de las series de agricultura de Valencia también hay diferencias, pero estas son consecuencia del método de extracción empleado para obtener las componentes de ciclo-tendencia (que para la serie obtenida siguiendo la estrategia INE detecta un cambio de nivel). En todo caso, las señales obtenidas de las series de agricultura de Valencia logradas con ambas estrategias (agregar-trasladar y trasladar-agregar) son lineales, por lo que no tienen mayor interés.

las diferencias sí son destacables. Las mayores discrepancias se obtienen para las ramas B+D+E, L, J y N (ver la Figura 2). Llama la atención la inclusión de la rama N en este grupo, ya que las diferencias en niveles entre las series logradas con ambas estrategias no son especialmente grandes.

Figura 2. Ejemplo de comparaciones de evoluciones cíclicas (medidas en tasas de variación interanual sobre señales de ciclo-tendencia) de diversas ramas recodificadas utilizando las estrategias trasladar-agregar (T-A, INE) y agregar-trasladar (A-T).



A la vista de los resultados mostrados en la Figura 2, con carácter general se puede afirmar que la estrategia agregar-trasladar puede producir resultados sensiblemente diferentes a la estrategia recomendada por el INE y que, por tanto, no puede ser empleada como sustitutiva. No obstante lo anterior, para la mayoría de las ramas ambas estrategias (T-A y A-T) son perfectamente intercambiables e, incluso, para algunas ramas como la J (*Información y comunicaciones*) donde los perfiles cíclicos que se obtienen son claramente diferentes, se observa que ambas estrategias identificarían finalmente los puntos de giro en casi los mismos instantes temporales.

5. Comparación entre recodificación con matriz de España y de Comunidad Valenciana

El INE recomienda con carácter general utilizar la matriz de paso de España para trasladar de una a otra clasificación CNAE los valores observados. La matriz de transferencia de España, sin embargo, se calcula teniendo en cuenta las respuestas de las personas encuestadas para la elaboración de la EPA en toda España; por lo que, en buena medida, representará una media ponderada (con los pesos de las estructuras de los mercados laborales de cada región como ponderaciones) de las matrices de transferencia del conjunto de las Comunidades Autónomas (CCAA) y no reflejará, por

tanto, las características individuales que presente la estructura productiva de una región o provincia en particular. Consecuentemente, aunque la recomendación del INE presenta la ventaja de gozar de una mayor precisión en cuanto a los coeficientes de paso estimados (ya que están basados en mayores tamaños muestrales), presenta el inconveniente de que los valores que se obtienen tras su aplicación pueden estar sujetos a importantes desviaciones consecuencia de las obvias diferencias que pueden existir entre las estructuras productivas de España y del ámbito territorial al que se apliquen.

Ya hemos visto en el apartado cuarto (al comparar las estrategias trasladar-agregar y agregar-trasladar) que aplicada la recomendación del INE a la Comunidad Valenciana puede producir, para determinadas ramas de actividad, valores trasladados que se alejan significativamente de los valores que se obtienen directamente de la EPA. Las preguntas obvias son: ¿Es esto consecuencia de haber utilizado la matriz de paso de España?; ¿se hubiesen conseguido resultados más ajustados si se hubiese utilizado una matriz de paso propia de la Comunidad Valenciana?; más aún, dado que se puede razonar que teóricamente los resultados esperados deberían estar más próximos si se utilizase una matriz de transferencia propia, ¿serían las diferencias entre los valores obtenidos utilizando la matriz de España y de la Comunidad Valenciana importantes? Resumiendo, ¿deberían las series de la Comunidad Valenciana ser trasladadas con una matriz propia o la recomendación del INE produce en la Comunidad Valenciana series trasladadas que no difieren significativamente de las que se obtienen con la matriz de España? Este apartado trata de dar respuestas, entre otras, a estas preguntas.

Para ello, de igual modo que se construyeron las series de ocupados aplicando la estrategia trasladar-agregar con las matrices de paso de España, se han calculado esas mismas series empleando las matrices de transferencia de la Comunidad Valenciana obtenidas en el apartado tercero. La comparación en niveles y señales cíclicas entre estas series y las logradas utilizando la matriz de transferencia de España nos ayudará a buscar respuestas a las preguntas planteadas previamente. Así mismo, para completar los análisis, y al igual que se hizo utilizando la matriz de España en el apartado “¿Agregar y trasladar o trasladar y agregar?”, también se han recodificado las series a CNAE-2009 utilizando la estrategia agregar-trasladar, pero utilizando la matriz de transferencia de la Comunidad Valenciana.

Antes de abordar el análisis comparativo del conjunto de series generadas, no obstante, se aborda la comparación de las estructuras del mercado de trabajo que se derivan de las oleadas de la EPA de 2008 utilizadas para construir las matrices de transferencia de España y la Comunidad Valenciana; tal análisis se ofrece en el primer subapartado de esta sección.

5.1. Comparación de estructuras productivas y matrices de transferencia

5.1.1. Estructuras productivas

Obviamente las estructuras productivas y, por ende, los mercados de trabajo¹⁹ de España y la Comunidad Valenciana presentarán diferencias y, *a priori*, estas deberían ser tanto más importantes cuanto mayor sea el nivel de desagregación que se

¹⁹ En este estudio la estructura productiva se mide, de hecho, a través de la distribución del mercado de trabajo, por lo que, en este trabajo, son conceptos absolutamente equivalentes.

considere. En este apartado, el análisis se centrará en comparar, para las dos clasificaciones CNAE, las estructuras productivas (medias a partir de los valores medios correspondientes a las cifras observadas en la EPA durante los trimestres de 2008) en los niveles de: división, rama de actividad y sector²⁰. Con carácter general, para el análisis se implementarán un conjunto de coeficientes que servirán para medir el grado de similitud/disimilitud entre las distribuciones de los mercados de trabajo de cada uno de los ámbitos territoriales. Adicionalmente, el análisis se apoyará también en representaciones gráficas para los dos últimos niveles de agregación.

Para las comparaciones analíticas se han utilizado varios índices basados en distancias/normas entre vectores y coeficientes que miden la proximidad entre dos distribuciones. En concreto, se han empleado valores medios de las normas euclídeas –en términos absolutos ($\| \cdot \|_2$) y relativos ($\| \cdot \|_{2r}$)– e infinito –en términos absolutos ($\| \cdot \|_\infty$) y relativos ($\| \cdot \|_{\infty r}$)– y del coeficiente de entropía (basado en la distancia/divergencia de Kullback-Leibler, KL), cuyas expresiones matemáticas vienen dadas por:

$$\|y-x\|_2 = \sqrt{\frac{1}{k} \sum_{i=1}^k (y_i - x_i)^2} \quad \|y-x\|_{2r} = \sqrt{\frac{1}{k} \sum_{i=1}^k \left(\frac{y_i - x_i}{x_i} \right)^2}$$

$$\|y-x\|_\infty = \max_{1 \leq i \leq k} |y_i - x_i| \quad \|y-x\|_{\infty r} = \max_{1 \leq i \leq k} \left| \frac{y_i - x_i}{x_i} \right|$$

$$KL(x,y) = -\frac{1}{100} \sum_{i=1}^k x_i \log \left(1 - \left| \frac{y_i - x_i}{100} \right| \right)$$

donde x representa un vector cualquiera de España, y denota el correspondiente vector para la Comunidad Valenciana, x_i e y_i las i -ésimas componentes de los vectores x e y , respectivamente, y k el número de componentes; donde, como se observa por la definición de las normas en términos relativos y del coeficiente de entropía²¹, la distribución de España es utilizada como referencia.

Los valores obtenidos para los coeficientes anteriores, que miden el grado de similitud entre las estructuras productivas de España y la Comunidad Valenciana, se ofrecen en la Tabla VII, donde se observa que a medida que se trabaja con niveles de agregación superior las estructuras productivas van convergiendo en términos relativos. Obviamente, sin embargo, en términos absolutos las diferencias se van incrementando a medida que crece el nivel de agregación, ya que las cifras con las que se trabajan son cada vez de mayor magnitud. Con carácter general, no obstante, destaca la mayor similitud que existe entre las estructuras en la clasificación CNAE-1993 frente a las de la clasificación CNAE-2009.

²⁰ Las distribuciones de las estructuras productivas a nivel de división se pueden consultar en los Anexos II y III; mientras que desagregadas por ramas y sectores se ofrecen en las Tablas VIII y IX, respectivamente.

²¹ El coeficiente de entropía, tal y como está definido, mide la diferencia entre las distribuciones como una suma ponderada de las diferencias entre componentes (con la importancia del componente como ponderación). Es nulo cuando hay coincidencia perfecta y castiga más las diferencias correspondientes a las componentes de mayor tamaño. Este coeficiente representa una evolución del coeficiente clásico, definido como $-(100)^{-1} \sum x_i \log(x_i/y_i)$, que es más sensible a componentes próximos a cero en y . Comparado con el resto de coeficientes tiene la ventaja de ser menos sensible al número de sumandos.

Tabla VII. Grado de similaridad de las estructuras productivas de España y la Comunidad Valenciana.

Medida similaridad	División		Ramas		Sectores	
	CNAE-1993	CNAE-2009	CNAE-1993	CNAE-2009	CNAE-1993	CNAE-2009
Euclídea	0.53	0.45	1.15	1.18	1.97	2.19
Euclídea relativa	0.68	0.60	0.18	0.18	0.15	0.16
Infinito	2.10	2.10	2.10	2.87	2.61	2.97
Infinito relativa	3.98	3.88	0.33	0.43	0.23	0.25
Entropía	2.28	2.43	2.01	2.03	1.65	1.60

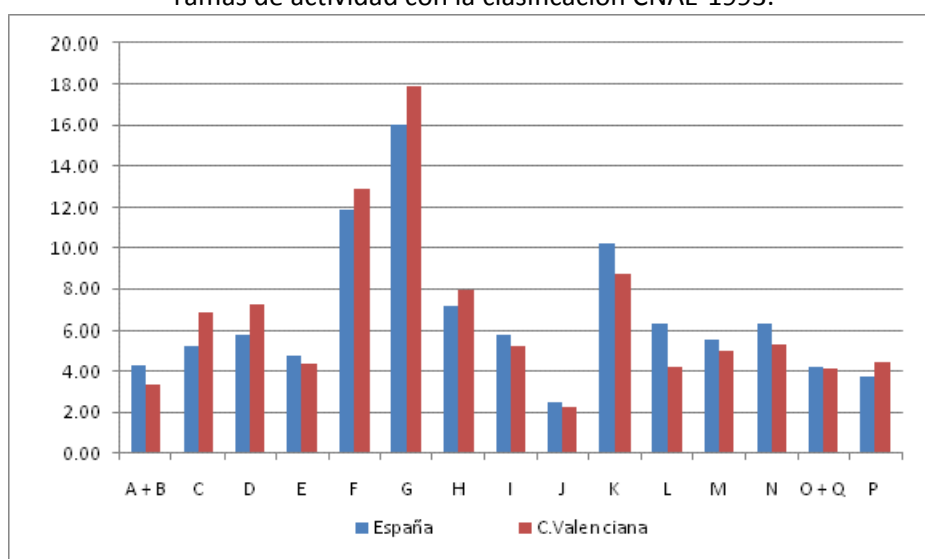
Fuente: elaboración propia.

Descendiendo en el nivel de detalle y comparando, en primer lugar, las estructuras en la clasificación CNAE-1993, las mayores diferencias, en términos absolutos, se observan en la división *Administración pública, defensa y seguridad social obligatoria* (75), que en España tiene, en media, una importancia más de dos puntos porcentuales superior a la que tiene en la Comunidad Valenciana. Junto a las diferencias para esta división también destacan (en valores absolutos) las que se registran para las divisiones 26, 51, 45 y 19 (ver el Anexo II), todas ellas con diferencias porcentuales en valor absoluto superiores al 1%, destacando por su importancia la Construcción cuyo peso en 2008 en la Comunidad Valenciana era superior al de España. En términos relativos, las mayores diferencias se localizan en la división 19, debido a su escasa importancia en el conjunto del país (con algo menos del 0.3% del empleo total) cuando en la Comunidad Valenciana representa algo más de 1.3% del empleo total. Además de la división anterior, las divisiones 26 y 17 también destacan por el mayor peso relativo que tienen en la Comunidad Valenciana.

En la comparación por divisiones en la clasificación CNAE-2009, es de nuevo la división *Administración Pública y defensa; Seguridad Social obligatoria* (84) la que registra las mayores diferencias en términos absolutos. De hecho, ambas divisiones miden exactamente lo mismo en cada una de las clasificaciones. Asimismo, y como no podía ser de otra manera, vuelven a destacar en CNAE-2009 las divisiones equivalentes a las identificadas en CNAE-1993. En concreto, en términos absolutos las mayores diferencias se concentran, en este caso, en las divisiones 23 (cuya división homóloga en CNAE-1993 es la 26), 56, 46 (que equivale a la división 51 en CNAE-1993) y 15 (heredera de la división 19 de CNAE-1993); ver el Anexo III para una descripción de las divisiones. Por otro lado, en términos relativos, se repite el mismo podio de la clasificación CNAE-1993, aunque hay un ligero cambio de orden. En particular, de nuevo la división 15 es donde se concentran las mayores diferencias relativas, seguido en esta ocasión de 13 (equivalente a la división 17 en CNAE-1993) y 23.

Estas diferencias a nivel de divisiones se trasladan lógicamente a nivel de ramas de actividad (obviamente atenuadas en términos relativos y algo incrementadas, por el efecto acumulación, en términos absolutos; ver la Tabla VII). Las estructuras productivas por ramas de actividad de España y la Comunidad Valenciana en la clasificación CNAE-1993 se ofrecen, de forma gráfica, en la Figura 3 y en mayor detalle en la Tabla VIII. Las mayores diferencias en términos absolutos vuelven a concentrarse en la rama L, que coincide con la división 75, seguido en esta ocasión de las ramas G, C, K, D, F y A+B (ver el Anexo I), todas ellas superando el 1% en términos absolutos.

Figura 3. Estructuras de los mercados de trabajo de España y la Comunidad Valenciana por ramas de actividad con la clasificación CNAE-1993.



En términos relativos las mayores diferencias vuelven a concentrarse en casi las mismas ramas que en términos absolutos. Con diferencias relativas que superan en valor absoluto, un 20% los correspondientes valores de las ramas de España L, C, D, A+B y P. Comparado con los términos absolutos destaca ahora la salida de las ramas G, K y F de la lista de los más destacados (sin duda debido a su enorme peso dentro del mercado de trabajo) y la entrada de la rama P (con un peso absoluto claramente menor).

Tabla VIII. Estructuras de los mercados de trabajo a nivel de rama de actividad de España y la Comunidad Valenciana.*

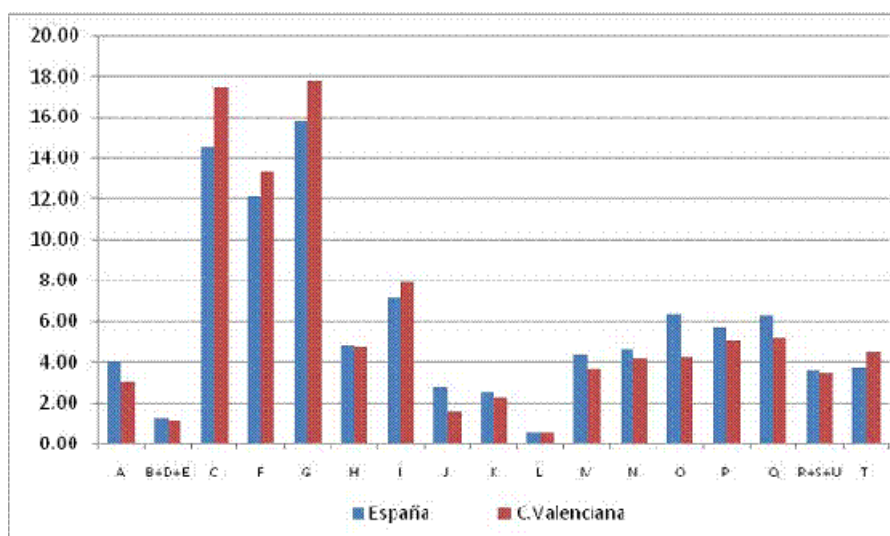
CNAE-1993			CNAE-2009		
Rama	España	C.Valenciana	Rama	España	C.Valenciana
A + B	4.34	3.33	A	4.04	3.01
C	5.26	6.84	B+D+E	1.22	1.14
D	5.85	7.22	C	14.57	17.44
E	4.81	4.40	F	12.11	13.32
F	11.88	12.95	G	15.81	17.81
G	15.99	17.92	H	4.77	4.72
H	7.17	7.91	I	7.17	7.91
I	5.83	5.28	J	2.76	1.57
J	2.50	2.24	K	2.51	2.25
K	10.24	8.74	L	0.58	0.57
L	6.31	4.21	M	4.33	3.66
M	5.59	5.02	N	4.57	4.17
N	6.30	5.35	O	6.30	4.21
O + Q	4.22	4.13	P	5.67	5.06
P	3.72	4.46	Q	6.23	5.24
			R+S+U	3.64	3.46
			T	3.72	4.46

* Los identificadores de las ramas de actividad utilizados en ambas CNAEs no tienen por qué guardar relación; ver las Tablas A.I y A.II.

Fuente: elaboración propia.

Para comparar en la clasificación CNAE-2009 las estructuras productivas por ramas de actividad de España y la Comunidad Valenciana, es preciso observar ahora los valores ofrecidos en la parte derecha de la Tabla VIII y la representación gráfica ofrecida en la Figura 4. Como es fácil deducir de la Figura 4, las mayores diferencias en términos absolutos se registran en las ramas C, O y G (ver el Anexo I); lógicamente tales ramas son herederas de aquellas que también fueron identificadas como las menos similares en la clasificación CNAE-1993. Además de para las ramas anteriores, también destacan las diferencias que se observan para F, J y A. En términos relativos las mayores diferencias también se registran en ramas similares a las identificadas previamente.

Figura 4. Estructuras de los mercados de trabajo de España y la Comunidad Valenciana por ramas de actividad con la clasificación CNAE-2009.



Aunque las ramas de actividad donde se han identificado las mayores diferencias son básicamente las mismas en ambas clasificaciones, no para todas ellas se producen diferencias significativas entre los valores trasladados y observados en la EPA (para 2008). Ello sin duda apunta a que probablemente influye, por encima de las diferencias globales, la manera en que se distribuyen dentro de cada rama las diferencias; es decir, las distribuciones de transferencia: la interacción de las divisiones que componen cada rama y las diferencias que existen entre las estructuras de España y la Comunidad Valenciana para esas ramas.

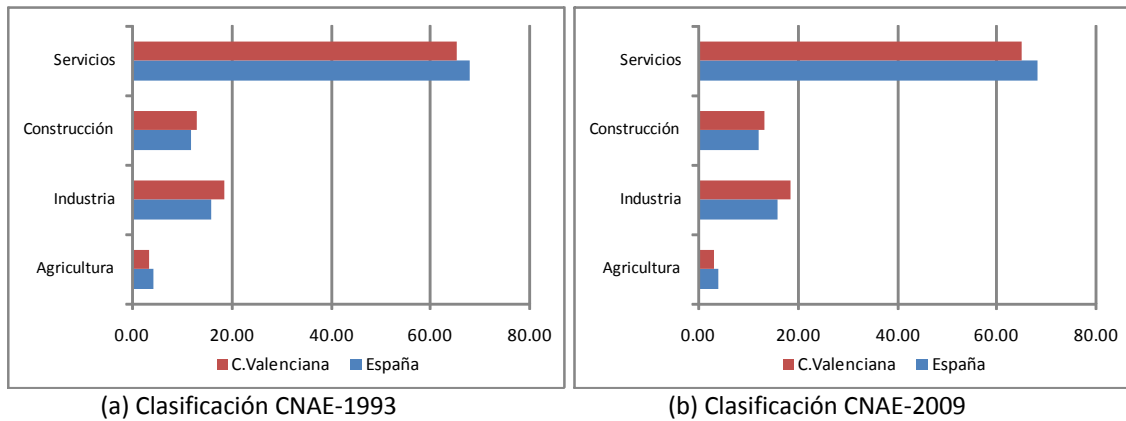
Tabla IX. Estructuras de los mercados de trabajo a nivel sectorial de España y la Comunidad Valenciana.

Sector	CNAE-1993		CNAE-2009	
	España	C.Valenciana	España	C.Valenciana
Agricultura	4.34	3.33	4.04	3.01
Industria	15.91	18.46	15.79	18.58
Construcción	11.88	12.95	12.11	13.32
Servicios	67.87	65.26	68.06	65.09

Fuente: elaboración propia.

Por último, para realizar las comparaciones sectoriales se ofrecen la Tabla IX y las representaciones de la Figura 5. En ellas se puede observar que las diferencias absolutas entre ambas clasificaciones prácticamente se mantienen y que, por tanto, en términos agregados sectoriales se produce aproximadamente el mismo trasvase porcentual de ocupados entre CNAEs. Obviamente, este resultado en términos absolutos, implica que la imagen para los trasvases entre clasificaciones en términos relativos sea completamente diferente. En la comparación entre estructuras productivas, las mayores diferencias, en términos relativos, se producen en Agricultura, seguido, en este orden, de Industria, Construcción y Servicios.

Figura 5. Estructuras de los mercados de trabajo de España y la Comunidad Valenciana por sectores.



En resumen, a todos los niveles (especialmente de división y ramas de actividad) se encuentran importantes diferencias entre las estructuras productivas de España y la Comunidad Valenciana, las cuales son lógicamente persistentes en ambas clasificaciones. Estas diferencias *per se*, sin embargo, no parecen suficientes para explicar las diferencias que se observan entre valores observados y trasladados, por lo que para tratar de comprender las diferencias observadas será necesario analizar, a través del estudio de las matrices de transferencia, cómo se distribuyen internamente estas diferencias.

5.1.2. Matrices de transferencia

Para valorar el grado de similitud de las matrices de transferencia se realizarán varios tipos de comparaciones. Por una parte, se compararán directamente las matrices de coeficientes (utilizando diversas métricas y realizando análisis detallados) para determinar en qué división/rama/sector se producen las mayores discrepancias en su reparto entre clasificaciones; mientras que, por otra, se compararán en cada ámbito las estructuras de transferencia de cada división/rama/sector sin tener en cuenta las diferencias de estructura productiva. Esto permitirá localizar dónde se concentran las mayores diferencias de paso no debidas a divergencias entre los tamaños de cada división, rama o sector.

5.1.2.1. Matrices de coeficientes

En primer término, para la comparación de las matrices de coeficientes de España y la Comunidad Valenciana se calculará para España una matriz media de coeficientes, $E_{n \times m}$. En efecto, dado que teóricamente se debería verificar la igualdad $C_{n \times m} = (D_{m \times n})^T$ (la

cual, como se expuso en el apartado segundo, no se verifica debido a problemas de redondeo) se han calculado las matrices:

$$E_{n \times m} = \frac{1}{2} (C_{n \times m} + (D_{m \times n})^T)$$

a partir de las cuales se han implementado varios test de bondad de ajuste y se han obtenido para los distintos órdenes (n x m) los coeficientes:

$$\|E - V\|_2 = \sqrt{\frac{1}{k} \sum_{j=1}^m \sum_{i=1}^n (e_{ij} - v_{ij})^2} \quad \|E - V\|_{2r} = \sqrt{\frac{1}{k} \sum_{e_{ij} \neq 0} \left(\frac{e_{ij} - v_{ij}}{e_{ij}} \right)^2}$$

$$\|E - V\|_\infty = \max |e_{ij} - v_{ij}| \quad \|E - V\|_{\infty r} = \max_{e_{ij} \neq 0} \left| \frac{e_{ij} - v_{ij}}{e_{ij}} \right|$$

donde E y V representan, respectivamente, las matrices de coeficientes de España y la Comunidad Valenciana, k es el número de coeficientes no nulos en la matriz de España²² y e_{ij} y v_{ij} sirven para notar un elemento genérico —el elemento (i, j) — de las matrices E y V , respectivamente.

En primer lugar y antes de analizar dónde se pueden concentrar las mayores diferencias entre las matrices de coeficientes de España y la Comunidad Valenciana estudiaremos si los valores observados para la Comunidad Valenciana en el conjunto de las oleadas de la EPA de 2008 podrían haber sido generados por la matriz de coeficientes de España. Para ello, se ha implementado un contraste de bondad de ajuste (comparando el coeficiente q de Pearson con la distribución χ^2) en el que se han considerado exclusivamente los coeficientes no nulos de la matriz de España.²³

En concreto, denotando por N al tamaño de la población ocupada en la Comunidad Valenciana en el acumulado de los cuatro trimestres de 2008, que se deriva de los factores de elevación de la EPA para la Comunidad Valenciana, se ha calculado para cada grupo de agregación (por divisiones, ramas y sectores) el estadístico:

$$q = \sum_{e_{ij} \neq 0} \frac{(Ne_{ij} - Nv_{ij})^2}{Ne_{ij}}$$

y se ha obtenido el p-valor correspondiente a una distribución χ^2 con tantos grados de libertad como coeficientes no nulos menos uno tiene la matriz correspondiente. El resumen de los cálculos implementados se ofrece en la Tabla X.

A la luz de los valores presentados en la Tabla X es obvio que a todos los niveles de agregación existen diferencias significativas entre las estructuras de los mercados de trabajo de España y la Comunidad Valenciana y que, por tanto, los valores observados para la EPA en la Comunidad Valenciana en 2008 no pueden considerarse una realización del conjunto de España. Este resultado, de hecho, queda corroborado observando los estadísticos de resumen presentados en la Tabla XI.

²² Se ha de tener en cuenta que de los 5280 coeficientes de la matriz de paso entre divisiones, solo 189 son no nulos; mientras que en las matrices de paso entre ramas los coeficientes no nulos *suben* hasta 51 de entre 255 y en las matrices entre sectores crecen hasta un 80%, 12 de los 16 coeficientes son no nulos.

²³ Al ser la Comunidad Valenciana un subconjunto de España, aquellos valores para los que el correspondiente coeficiente de España es nulo también lo será para la Comunidad Valenciana.

Tabla X. Análisis de bondad de ajuste entre las matrices de coeficientes de España y la Comunidad Valenciana.

Matriz	q de Pearson	Grados de libertad	p-valor
Divisiones	1.247.460	188	0.0000
Ramas actividad	352.431	50	0.0000
Sectores	173.710	11	0.0000

Fuente: elaboración propia.

Como se observa en la Tabla XI, las diferencias en términos absolutos entre los coeficientes de las distintas matrices crecen a medida que crece la agregación, algo lógico pues cada vez los valores son de mayor cuantía. Sin embargo, dado que los coeficientes de similitud se van reduciendo en términos relativos, podemos interpretar que las matrices van convergiendo a medida que se consideran mayores niveles de agregación. Llama la atención el valor obtenido para $\|E_{4 \times 4} - {}_{cv}C_{4 \times 4}\|_{cor}$, que es unitario, indicando que hay transferencia de trabajadores entre sectores productivos en España que, sin embargo, no se registran a nivel de la Comunidad Valenciana. De hecho, esto ocurre en tres ocasiones. En España hay algunos trabajadores clasificados en Agricultura en CNAE-1993 que son clasificados en Industria en CNAE-2009, trabajadores clasificados en Servicios en CNAE-1993 que son clasificados en Agricultura en CNAE-2009 y trabajadores clasificados en Construcción en CNAE-1993 que son clasificados en Servicios en CNAE-2009 y ninguno de tales trasvases se registran para el mercado de trabajo valenciano.

Tabla XI. Grado de similitud entre las matrices de coeficientes de España y la Comunidad Valenciana.

Medida	Divisiones	Ramas	Sectores
Euclídea	0.30	0.62	1.21
Euclídea relativa	1.56	0.70	0.55
Infinito	2.10	2.10	2.79
Infinito relativa	8.12	2.26	1.00

Fuente: elaboración propia.

A la vista de estos resultados, la cuestión inmediata que debería ser abordada sería ¿dónde se localizan las mayores diferencias? Para buscar respuestas a esta pregunta se ha realizado un rastreo pormenorizado de las entradas donde se registran mayores discrepancias.

En el análisis más detallado se observa que, a nivel de división, las mayores diferencias entre las matrices de coeficientes de España y la Comunidad Valenciana se localizan en los cruces (donde en primer término se nombra la división en CNAE-1993 y en segundo lugar en la clasificación CNAE-2009): 75 y 84; 55 y 56; 26 y 23; 51 y 46; 52 y 47; y, 1 y 1 (ver el Anexo II y III). Es decir, las diferencias más importantes en términos de división no son consecuencia de trasvases entre divisiones que se han comportado de manera diferente entre los ámbitos de España y la Comunidad Valenciana, sino consecuencia de las diferentes estructuras productivas de ambos territorios.

Por ramas de actividad, las mayores diferencias se concentran (donde de nuevo en primer lugar se nombra la rama en la clasificación CNAE-1993 y en segundo lugar la rama en la clasificación CNAE-2009) entre L y O, G y G, C y C, A+B y A, N y Q, H e I, y P y T (ver el Anexo I). Mientras que, a nivel sectorial, las mayores diferencias se vuelven a

concentrar en la diagonal (es decir, de transferencias entre sectores consecuencia de las diferentes estructuras productivas), con Servicios e Industria casi moviéndose en diferencias porcentuales del entorno del 2.75% y Agricultura y Construcción alrededor del 1.05%.

De nuevo, por tanto, las mayores discrepancias en las matrices de coeficientes son debidas a diferencias de estructura productiva y no trasvases entre sectores. Es decir, las discrepancias en las matrices de coeficientes entre divisiones, ramas o sectores detectadas en el subapartado anterior tienen su base en el elemento equivalente en la otra clasificación y no en los trasvases que se hayan podido producir al cambiar de clasificación.

Falta por analizar, no obstante, si de forma local (es decir, no teniendo en cuenta las diferencias de pesos de cada división, rama o sector) se producen diferencias significativas entre los trasvases que se registran. En otras palabras, estudiar si ponderando cada división, rama o sector en una misma base para ambos ámbitos territoriales la distribución de trasvases es similar o no. Esto es posible analizarlo comparando las filas de las matrices fila estandarizadas, ya que al ser todas las filas de suma unitaria (o 100 en porcentaje), las diferencias de tamaño de cada fila para cada ámbito desaparecen y solo se comparan las distribuciones de trasvases.

5.1.2.2. Matrices fila estandarizadas

De igual forma a como se hizo para comparar diferencias en estructuras productivas y matrices de coeficientes, se utilizarán un conjunto de estadísticos para valorar el grado de similitud existente entre las distribuciones de transferencia de cada división/rama/sector de España y la Comunidad Valenciana. Además de comparaciones de matrices a nivel global, se realizarán diversas pruebas de ajuste local (por división/rama/sector). A nivel global, se han utilizado las métricas introducidas en el subapartado anterior para medir proximidad entre matrices y, adicionalmente, se ha obtenido el estadístico CKL, basado en el concepto de entropía cruzada, que calcula la media ponderada de las entropías de cada grupo, con ponderaciones proporcionales al tamaño de cada grupo, es decir:

$$CKL(X, Y) = -\sum_{i=1}^n x_i \sum_{j=1}^m r_{ij} \log(1 - |s_{ij} - r_{ij}|)$$

donde n y m representan, respectivamente, el número de filas de la correspondiente matriz fila estandarizada, x_i el peso (en porcentaje) del grupo que representa la i -ésima fila en la correspondiente estructura productiva y r_{ij} y s_{ij} , respectivamente, los valores (en proporción) de los coeficientes (i, j) de las matrices de transferencia fila estandarizadas de España y la Comunidad Valenciana.

A nivel local, para cada fila de las correspondientes matrices fila estandarizadas (tomando las de España como referencia): (i) se han calculado los estadísticos introducidos en el apartado 5.1.1 para medir el grado de similitud entre estructuras productivas²⁴, donde obviamente al aplicar las métricas en términos relativos ($\| \cdot \|_{2r}$ y $\| \cdot \|_{\infty r}$) solo se han considerado los valores para los que la correspondiente entrada en la

²⁴ Al fin y al cabo, cada fila de una matriz fila estandarizada se puede ver como una distribución, por lo que lo que se hace es estudiar cómo se distribuyen los individuos clasificados (en un grupo) en una clasificación entre los grupos en que se divide la otra clasificación.

matriz de España es no nula; y (ii) se han aplicado pruebas χ^2 de bondad de ajuste en cada nivel de agregación. En concreto, para esto último, denotando por $r_i=(r_{i1}, \dots, r_{im})$ a la fila i -ésima de una matriz de transferencia fila estandarizada (en proporciones) de España y $s_i=(s_{i1}, \dots, s_{im})$ a la fila equivalente en la Comunidad Valenciana y por N_i al tamaño de la población ocupada del grupo i -ésimo de la Comunidad Valenciana (que se deriva de aplicar los factores de elevación de la EPA para la Comunidad Valenciana en el acumulado de los cuatro trimestres de 2008), se han calculado los estadísticos

$$q_i = \sum_{r_{ij} \neq 0} \frac{(N_i r_{ij} - N_i s_{ij})^2}{N_i r_{ij}}$$

que se han comparado con los correspondientes valores de la distribución χ^2 con tantos grados de libertad como el número sumandos en q_i menos uno.

Tabla XII. Grado de similaridad entre las matrices fila estandarizadas de transferencia de CNAE-1993 a CNAE-2009 de España y la Comunidad Valenciana.

Medida	Divisiones	Ramas	Sectores
Euclídea	11.70	2.10	1.21
Euclídea relativa	0.20	0.57	0.57
Infinito	98.0	7.17	2.59
Infinito relativa	5.28	2.81	1.00
CKL	3.32	1.47	0.56

Fuente: elaboración propia.

En la Tablas XII y XIII se ofrecen los coeficientes de similaridad globales calculados para los distintos niveles de agregación comparando, respectivamente, las matrices de transferencia fila estandarizadas de España y la Comunidad Valenciana entre las clasificaciones CNAE-1993 y CNAE-2009 (Tabla XII) y CNAE-2009 y CNAE-1993 (Tabla XIII). Comparando los resultados de las tablas para divisiones se observa que en términos absolutos las diferencias que presentan las matrices de transferencia son superiores a las que presentan las matrices de coeficientes, mientras que en términos relativos, en general, estas son inferiores. Estos resultados indican que cuando se elimina el efecto del tamaño de las divisiones las matrices son más parecidas, pero que, sin embargo, hay algunas pocas divisiones para las que las diferencias en las distribuciones de transferencia son extremadamente significativas.

A nivel de rama y sector se observan resultados parecidos, aunque los coeficientes en términos relativos son bastante similares, lo que señala que en este caso las matrices son bastante parecidas en términos globales incluso después de haber eliminado el efecto tamaño, empero lo anterior hay algunas ramas para las que las diferencias son notables.

Dejando para el último término los análisis estadísticos de bondad de ajuste, a continuación se identificarán las divisiones/ramas/sectores para las que las transferencias entre clasificaciones son más destacables. Analizando las medidas de similaridad calculadas comparando las matrices de paso entre las clasificaciones CNAE-1993 y CNAE-2009 por divisiones, se deduce que las mayores diferencias se concentran en las divisiones 10, 11, 2, 30, 18, 55, 64, 70 y 33 (ver el Anexo II). En efecto, como es fácil derivar atendiendo a los valores de la Tabla XIV, donde se ofrece el detalle de las distribuciones de transferencia para estas divisiones para España y la

Comunidad Valenciana, las diferencias de transferencia entre ambos ámbitos son evidentes y destacables para todas las divisiones identificadas, resaltando especialmente las diferencias entre las divisiones 10, 11, 2 y 30.

Tabla XIII. Grado de similaridad entre las matrices fila estandarizadas de transferencia de CNAE-2009 a CNAE-1993 de España y la Comunidad Valenciana.

Medida	Divisiones	Ramas	Sectores
Euclídea	13.75	2.27	0.34
Euclídea relativa	1.78	0.71	0.53
Infinito	100.00	8.66	0.81
Infinito relativa	10.66	2.62	1.00
CKL ⁽¹⁾	0.69	1.32	0.25

(1) Calculado excluyendo las divisiones 5, 6 y 7 para las que no se observa ningún ocupado en la Comunidad Valenciana en 2008.

Fuente: elaboración propia.

Tabla XIV. Principales diferencias a nivel de división de distribuciones porcentuales de transferencia entre CNAE-1993 y CNAE-2009 para España y la Comunidad Valenciana.

CNAE-1993	CNAE-2009	España	Comunidad Valenciana	CNAE-1993	CNAE-2009	España	Comunidad Valenciana
10	5	96.59	0.00	24	19	0.16	0.00
10	8	1.80	100.00	24	20	67.05	87.68
10	19	1.61	0.00	24	21	31.98	11.86
11	6	82.84	0.00	24	26	0.47	0.46
11	9	17.16	100.00	24	32	0.34	0.00
2	1	2.23	77.05	55	55	22.64	12.38
2	2	97.55	22.95	55	56	77.36	87.62
2	16	0.23	0.00	64	53	43.99	56.29
30	26	80.47	47.79	64	60	0.11	0.00
30	28	7.31	0.00	64	61	55.90	43.71
30	33	10.89	52.21	70	41	-23.54	35.51
30	62	1.33	0.00	70	68	60.39	56.80
18	13	0.14	0.00	70	81	16.06	7.68
18	14	94.47	75.20	33	26	32.01	20.88
18	15	5.39	24.80	33	28	0.43	0.68
				33	32	62.30	75.27
				33	33	5.26	3.17

Fuente: elaboración propia.

Comparando los resultados presentados en las Tablas XIV y XV, se observa que las diferencias a nivel de rama de actividad se amortiguan sensiblemente respecto a las de división, aunque estas todavía son evidentes para algunas ramas. De hecho, prácticamente todas las medidas de similaridad utilizadas apuntan a que las mayores diferencias entre las distribuciones de transferencia entre España y la Comunidad Valenciana se concentran en las ramas K, D, C, A+B y O+Q.

Tabla XV. Distribuciones porcentuales de transferencia entre CNAE-1993 y CNAE-2009 a nivel de rama de actividad para España y la Comunidad Valenciana.

CNAE		CNAE-2009																
1993		A	B+D+E	C	F	G	H	I	J	K	L	M	N	O	P	Q	R+S+U	T
Esp	A+B	93.16		0.01									6.81				0.02	
CV	A+B	90.58		-									9.42				-	
Esp	C		0.07	92.63					7.29								0.01	
CV	C		0.12	96.33					3.55								-	
Esp	D		14.01	85.83	0.10								0.06					
CV	D		9.21	90.68	0.11								-					
Esp	E		1.43	97.15	0.18				0.01								1.23	
CV	E		2.04	97.44	0.08				-								0.44	
Esp	F				99.94								0.06					
CV	F				100								-					
Esp	G			0.00		98.90	0.13										0.97	
CV	G			-		99.39	0.25										0.36	
Esp	H							100										
CV	H							100										
Esp	I						81.47		13.29			0.01	5.24					
CV	I						88.64		7.82			-	3.54					
Esp	J									100								
CV	J									100								
Esp	K	0.00		0.12	2.21				11.13	0.04	5.67	41.59	37.93		0.69		0.62	
CV	K	-		0.24	4.08				7.02	0.15	6.52	40.52	40.81		0.36		0.29	
Esp	L												0.05	99.93			0.02	
CV	L												-	100			-	
Esp	M														100			
CV	M														100			
Esp	N											1.20				98.80		
CV	N											2.15				97.85		
Esp	O+Q		7.78						10.81				1.69		0.25		79.46	
CV	O+Q		9.06						7.24				2.47		0.05		81.18	
Esp	P																	100
CV	P																	100

Fuente: elaboración propia.

Como se observa, las mayores discrepancias se concentran precisamente en algunas de las divisiones que originan las ramas para las que se obtuvieron mayores diferencias cuando en el apartado cuarto se realizó la comparación entre las estrategias trasladar-agregar y agregar-trasladar. Recuérdese que en aquel momento las mayores diferencias se observaron para las ramas origen B+D+E, J y L, siendo asimismo destacables las diferencias que se observan para las ramas C, M, N y R+S+U. De hecho, habría sido interesante realizar un análisis pormenorizado de las columnas de las matrices fila estandarizadas ya que se observa que, en esencia, los mayores cambios parecen producirse para las ramas para las que se obtuvieron en el apartado cuatro mayores diferencias. A la luz de estos resultados, además de las diferencias de

estructuras productivas, se puede afirmar que son las diferencias en las distribuciones de transferencia las que explicarían la mayoría de las discrepancias.

La comparación entre las estructuras de transferencia de España y la Comunidad Valenciana por sectores se ofrece en la Tabla XVI. Como se observa claramente a partir de los valores de la tabla (y de hecho así lo confirman todos los índices de similitud) las mayores diferencias se concentran en Agricultura e Industria, siendo más notables para el sector primario.

Tabla XVI. Distribuciones porcentuales de transferencia entre CNAE-1993 y CNAE-2009 a nivel sectorial para España y la Comunidad Valenciana.

CNAE-1993		CNAE-2009			
		Agricultura	Industria	Construcción	Servicios
España	Agricultura	93.16	0.01		6.83
C.Valenciana		90.58	-		9.42
España	Industria		97.10	0.09	2.81
C.Valenciana			98.52	0.06	1.42
España	Construcción			99.94	0.06
C.Valenciana				100	-
España	Servicios		0.50	0.33	99.16
C.Valenciana			0.61	0.55	98.85

Fuente: elaboración propia.

De igual forma que en el análisis del paso entre las clasificaciones CNAE-2009 y CNAE-1993, para comparar en detalle las distribuciones de transferencia entre las clasificaciones CNAE-2009 y CNAE-1993 se han calculado medidas resumen de similitud a nivel individual, los resultados de los cuales se ofrecen en el Anexo IV. Los valores de las distribuciones que presentan mayores diferencias a nivel de división se ofrecen en la Tabla A.3, mientras que para el conjunto de distribuciones de transferencia se puede observar la Tabla A.5 por ramas y la Tabla A.4 por sectores. A fin de no sobrecargar el texto, se omite un análisis detallado de estas tablas.

Después del extenso análisis anterior donde se han identificado los grupos (divisiones/ramas/sectores) donde se localizan las mayores diferencias en términos locales, se ha estudiado, mediante la implementación de pruebas de bondad de ajuste, hasta qué punto los valores de transferencia realmente observados en la Comunidad Valenciana podrían haber sido generados por las distribuciones de España. Este análisis complementará el anterior pues permitirá localizar para qué divisiones²⁵ las diferencias son estadísticamente significativas. En todo caso, es evidente que para aquellos casos en los que la estructura de transferencia se reduzca a un solo punto (es decir, donde el 100% de un grupo en una clasificación se asocia a un único grupo en la otra clasificación) no tendrá sentido calcular el ajuste.

²⁵ Debido a la sensibilidad que presenta el estadístico q al tamaño de la muestra (que lleva a rechazar *de facto* la hipótesis nula en casi todos los casos cuando el tamaño es grande) no se presentan estas pruebas para ramas y sectores, ya que en tales casos el resultado ha apuntado a que existen diferencias significativas.

En el caso del paso de CNAE-1993 a CNAE-2009 los resultados indican que a nivel de división para todas las divisiones donde tiene sentido realizar los cálculos existen diferencias significativas. Por otro lado, cuando el análisis se realiza de CNAE-2009 a CNAE-2009, los resultados son algo más variados (ver la Tabla XVII).

Tabla XVII. Análisis de bondad de ajuste por filas la matriz fila estandarizada por divisiones entre clasificaciones CNAE-2009 y CNAE-1993.

Rama	Estadíst.	g.l.	p-valor	Rama	Estadíst.	g.l.	p-valor	Rama	Estadíst.	g.l.	p-valor	Rama	Estadíst.	g.l.	p-valor
1	26.56	1	0.00	24	1.35	2	0.51	50	0.00	0	-	75	0.00	0	-
2	0.26	2	0.88	25	25.66	3	0.00	51	0.00	0	-	77	8.27	1	0.00
3	0.01	1	0.93	26	230.11	4	0.00	52	122.98	1	0.00	78	0.00	0	-
5	0.00	0	-	27	959.25	2	0.00	53	0.00	0	-	79	2.88	1	0.09
6	0.00	0	-	28	17.22	5	0.00	55	0.00	0	-	80	10.46	2	0.01
7	0.00	0	-	29	0.68	2	0.71	56	0.00	0	-	81	41.52	4	0.00
8	4.93	1	0.03	30	8.51	4	0.07	58	52.54	1	0.00	82	0.00	0	-
9	66.52	1	0.00	31	0.00	0	-	59	12.47	2	0.00	84	0.00	0	-
10	0.00	1	0.98	32	571.48	8	0.00	60	1.68	2	0.43	85	26.09	2	0.00
11	0.01	1	0.90	33	17.52	11	0.09	61	0.00	0	-	86	0.00	0	-
12	0.00	0	-	35	28.51	1	0.00	62	0.02	1	0.89	87	0.00	0	-
13	0.63	2	0.73	36	0.00	0	-	63	212.27	2	0.00	88	0.00	0	-
14	2.41	2	0.30	37	0.00	0	-	64	6.78	1	0.01	90	0.00	0	-
15	49.02	1	0.00	38	49.09	2	0.00	65	0.00	0	-	91	0.43	1	0.51
16	1.44	3	0.70	39	0.00	0	-	66	0.00	0	-	92	0.00	0	-
17	2.57	3	0.46	41	141.07	1	0.00	68	0.00	0	-	93	0.16	1	0.69
18	0.00	0	-	42	0.00	0	-	69	0.00	0	-	94	0.00	0	-
19	2.63	2	0.27	43	0.39	3	0.94	70	0.00	0	-	95	7.66	4	0.10
20	3.37	1	0.07	45	0.00	0	-	71	0.00	0	-	96	0.07	1	0.79
21	0.04	1	0.83	46	0.00	0	-	72	0.00	0	-	97	0.00	0	-
22	0.95	2	0.62	47	51.38	1	0.00	73	0.00	0	-	99	0.00	0	-
23	0.00	0	-	49	0.00	0	-	74	0.01	1	0.90				

Fuente: elaboración propia

5.2. Comparación en niveles

Una vez identificados los grupos (divisiones/ramas/sectores) donde se concentran las mayores diferencias en estructura productiva y en distribuciones de paso, es el momento de abordar el análisis de las diferencias que puedan existir entre las series trasladadas utilizando cada una de las matrices de transferencia. En primer lugar se analizará, empleando la estrategia agregar-trasladar, el impacto que tiene el uso de una u otra matriz sobre todas las series publicadas en la web del IVE (más de un millar), para a continuación realizar un estudio más detallado de las 33 series de ocupados analizadas en la sección anterior (epígrafe 5.2.2).

5.2.1. Comparación con valores observados en 2008

En 2008 el INE al realizar las encuestas recopiló la información básica utilizando ambas clasificaciones CNAE, por lo que para los cuatro trimestres de ese año se dispone de los valores (estimados) reales. Así, todas las series trasladadas a partir de las series publicadas en la página web del IVE, utilizando la estrategia agregar-trasladar y las matrices de transferencia de España y la Comunidad Valenciana, han sido comparadas con los valores reales para valorar que grupo de series presenta valores más próximos a los reales.

La comparación entre valores trasladados y valores observados en 2008 se ha realizado en términos agregados dado el elevado número de variables consideradas. A fin de buscar patrones generales. Lo que se ha hecho es medir, para cada grupo de variables, el porcentaje de veces en el que el valor trasladado utilizando la matriz de la Comunidad Valenciana genera mejores soluciones que el correspondiente valor trasladado empleando la matriz de España (los empates, que se corresponden con las variables totales que miden agregaciones por ramas o sectores, no han sido considerados).

En concreto, para el cálculo de los porcentajes (i) se ha obtenido para el valor de cada trimestre de cada serie la diferencia en valor absoluto entre el valor real, r , y los valores trasladados utilizando las matrices de España, e , y la Comunidad Valenciana, v ; (ii) se ha contabilizado el número de observaciones para las que la estimación con la matriz de España está más próxima al valor real, n_e , y también el número para las que la más próxima es la estimación obtenida tras trasladar utilizando la matriz de la Comunidad Valenciana, n_v , a partir de las cuales se ha calculado el estadístico:

$$100 \frac{n_v}{n_v + n_e}$$

El análisis de tales estadísticos ha permitido extraer varias conclusiones generales. Primero, utilizar la matriz de la Comunidad Valenciana para trasladar valores entre CNAEs produce, en general, valores más próximos a los reales que los que se obtienen con la matriz de España. Segundo, la ventaja de las estimaciones logradas utilizando la matriz de la Comunidad Valenciana es más clara para las series trasladadas a CNAE-2009 (precisamente las que despiertan un mayor interés práctico). Tercero, en general y como era de esperar, el uso de la matriz de la Comunidad Valenciana funciona mejor para las variables del conjunto de la Comunidad que para las variables provinciales. Cuarto, entre los grupos de variables donde se observa una mayor ventaja (media en porcentaje de veces que mejora resultados más próximos a los reales) al utilizar la matriz de la Comunidad Valenciana frente a la de España se encuentran las variables agrupadas por ramas de actividad. Este resultado es interesante, pues es precisamente para este conjunto de variables donde los resultados obtenidos con la estrategia agregar-trasladar utilizando la matriz de España presentaban mayores limitaciones, produciendo estimaciones menos precisas que desaconsejaban ese enfoque.

5.2.2. Trasladando con ambas matrices de transferencia

Ya se ha visto en el subepígrafe anterior que cuando se trabaja con valores trasladados a partir de variables agregadas por ramas y/o sectores, las estimaciones que se logran

utilizando la matriz de la Comunidad Valenciana son, en general, mejores que las que se obtienen con la matriz de España. Asimismo, cuando la estrategia recomendada (trasladar-agregar) fue utilizada para generar las series de número de ocupados en la clasificación CNAE-2009 se observó (ver la Tabla III) que esta no era una panacea. A pesar de que, en general, este enfoque mejoraba la estrategia agregar-trasladar, se obtuvo que para determinadas variables por ramas de actividad²⁶ producía importantes diferencias entre los valores reales (de 2008) y los trasladados. Casualmente, además, algunas de las ramas para las que se observaron las mayores desviaciones en la sección cuarta se corresponden precisamente con aquellas donde se concentran las mayores diferencias en las distribuciones de transferencia de España y la Comunidad Valenciana. Por lo que, en este contexto, cabría preguntarse si podrían obtenerse mejores ajustes o si estos variarían significativamente si en esta estrategia se emplease la matriz de la Comunidad Valenciana en lugar de la de España. Este subapartado tiene por objeto realizar tal comparación.

El análisis se realizará en dos etapas tanto por ramas de actividad como por sectores. En primer término se realizará una comparación para todo el período trasladado (1999.I a 2008.IV) utilizando como referente las series obtenidas utilizando la matriz de España (ya que es la recomendada por el INE) y, posteriormente, el análisis se completará con una comparación de los dos grupos de variables con los valores realmente observados en 2008. Al igual que en la sección cuarta, se ha utilizado como estadístico de comparación la diferencia absoluta relativa en porcentaje, D_r , entre el valor de referencia (los valores trasladados utilizando la matriz de España o los valores reales de 2008, según corresponda) y el valor alternativo o de comparación. Este estadístico se ha calculado para cada trimestre y variable, obteniéndose medias anuales y globales como resúmenes en que basar los análisis.

5.2.2.1. Ramas de actividad

En la Tabla XVIII se ofrecen los valores medios (anuales y globales) de las diferencias relativas en valor absoluto de los valores obtenidos, para cada una de las 17 series de ocupados de la Comunidad Valenciana por ramas de actividad, tras aplicar la estrategia agregar-trasladar con las matrices de España y Comunidad Valenciana. Como se observa para todas las ramas, las diferencias obtenidas son bastante estables a lo largo del tiempo, siendo no superiores al 5%, salvo las discrepancias observadas para las ramas J y L. Además de las ramas anteriores, las otras ramas donde se registran las mayores discrepancias relativas son H, N, M, A, R+S+U y B+D+E. Comparando estos resultados con los de la Tabla III, se observa que aparecen de nuevo las ramas J y L con diferencias notables. Un análisis de las distribuciones de transferencia entre CNAE-1993 y CNAE-2009 (Tabla XV) permite deducir fácilmente por qué en estas ramas se concentran las mayores diferencias. El análisis, sin embargo, hay que realizarlo por columnas²⁷, ya que es esto lo que permite descubrir el destino y, por tanto, la composición de las series en la clasificación CNAE-2009 identificadas con las mayores diferencias.

²⁶ Por sectores las diferencias eran poco significativas.

²⁷ La observación (por filas) de la matriz de transferencia fila estandarizada de CNAE-2009 a CNAE-1993 (Anexo IV) también permite observar (aunque algo distorsionado este fenómeno).

Tabla XVIII. Diferencias relativas en valor absoluto para las series de ocupados de la Comunidad Valenciana por ramas de actividad trasladadas con la matriz de la Comunidad Valenciana frente a las trasladadas con la matriz de España.

Rama	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	Global
A	2.45	2.53	2.41	2.40	2.46	2.52	2.50	2.41	2.47	2.52	2.47
B+D+E	2.75	1.37	3.50	1.20	2.33	1.39	1.79	1.40	1.52	1.45	1.87
C	0.81	0.85	0.76	0.74	0.72	0.76	0.81	0.81	0.85	0.79	0.79
F	0.63	0.71	0.64	0.70	0.65	1.04	1.29	0.99	0.84	0.97	0.85
G	0.31	0.35	0.34	0.33	0.38	0.38	0.39	0.37	0.41	0.44	0.37
H	4.80	4.92	4.91	4.45	3.98	4.56	5.96	5.31	4.01	3.78	4.67
I	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
J	18.63	18.25	18.10	17.61	17.63	16.44	18.47	16.40	18.11	16.73	17.64
K	0.35	0.29	0.45	0.47	0.47	0.45	0.39	0.40	0.41	0.44	0.41
L	5.94	5.94	5.94	5.94	5.94	5.94	5.94	5.94	5.94	5.94	5.94
M	2.76	2.69	3.06	2.97	3.07	2.97	2.89	2.79	2.85	2.83	2.89
N	5.57	5.67	4.91	4.40	4.70	3.65	2.57	2.87	3.30	3.43	4.11
O	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07	0.07
P	0.63	0.63	0.59	0.62	0.67	0.65	0.71	0.67	0.74	0.75	0.67
Q	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96	0.96
R+S+U	3.07	3.11	2.58	2.85	2.09	2.28	1.77	1.43	0.15	1.02	2.03
T	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Fuente: elaboración propia

Centrándose en primer lugar en la composición de la rama J, se observa que dependiendo de la matriz que se utilice su composición es notablemente diferente. Así, por ejemplo, destaca que utilizando la matriz de España a nivel de rama de actividad²⁸ el 7.29% de los ocupados clasificados en CNAE-1993 en el sector C serían asignados al sector J de CNAE-2009, cuando de acuerdo con la estructura del mercado de trabajo de la Comunidad Valenciana solo debería ser reasignados el 3.55% de tales trabajadores. Similares resultados ocurren con las reasignaciones que impone la matriz de España para las ramas I, K y O+Q que con la matriz de España pasarían, respectivamente, el 13.29%, 11.13% y 10.81% de sus trabajadores, cuando con la matriz de la Comunidad Valenciana solo correspondería el 7.82%, 7.02% y 7.24% respectivamente. Para la rama L, por su parte, la diferencia se centra en que con la matriz de España le correspondería el 5.67% de la rama K y con la de la Comunidad Valenciana el 6.52%.

²⁸ En realidad el análisis debería ser realizado a nivel de división; se ofrece a nivel de rama para simplificar la exposición. A nivel de división obviamente también se tienen diferencias notables. Por ejemplo, con la matriz de España el 41.75% de la división 22 es asignado a J, cuando con la matriz de la Comunidad Valenciana solo lo sería el 33.66%. Similares desviaciones se observan con las divisiones 64 y 92 que pasan a la rama J el 55.91% y 22.77%, respectivamente, de sus ocupados utilizando la matriz de España y el 43.71% y 15.50%, con la matriz de la Comunidad Valenciana.

Tabla XIX. Diferencias relativas en valor absoluto, para las series de ocupados por ramas de la Comunidad Valenciana, de los valores obtenidos tras trasladar con matrices de España y de la Comunidad Valenciana respecto a las cifras oficiales.

Rama	Error relativo de series obtenidas con matriz de España (INE)					Error relativo de series obtenidas con matriz Comunidad Valenciana				
	2008.I	2008.II	2008.III	2008.IV	2008	2008.I	2008.II	2008.III	2008.IV	2008
A	3.96	3.30	1.27	1.94	2.62	1.31	0.73	1.19	0.70	0.98
B+D+E	5.64	5.10	3.19	3.43	4.34	4.18	3.51	4.60	4.81	4.28
C	1.31	0.77	0.36	0.30	0.68	0.49	0.04	0.39	0.45	0.34
F	1.36	1.29	0.43	0.84	0.98	0.56	0.51	0.52	0.47	0.52
G	0.76	0.09	0.24	0.62	0.43	0.33	0.35	0.22	0.20	0.28
H	2.11	3.94	3.48	5.02	3.64	1.79	0.11	0.14	1.50	0.88
I	0.03	0.02	0.01	0.02	0.02	0.03	0.02	0.01	0.02	0.02
J	22.11	12.25	20.35	26.08	20.20	2.86	7.60	0.15	5.36	3.99
K	0.38	0.45	0.56	0.47	0.46	0.03	0.02	0.11	0.01	0.04
L	20.24	22.09	0.95	4.81	12.02	13.09	14.83	6.84	10.46	11.30
M	2.60	0.42	5.51	9.03	4.39	5.43	3.21	2.53	6.00	4.29
N	1.15	0.80	4.58	6.17	3.17	2.20	2.60	1.15	3.06	2.25
O	0.07	0.03	0.09	0.32	0.13	0.00	0.04	0.03	0.38	0.11
P	1.02	0.42	0.44	1.10	0.75	0.30	0.34	0.37	0.38	0.35
Q	1.09	1.32	0.94	0.58	0.98	0.12	0.35	0.02	0.38	0.22
R+S+U	3.56	0.74	1.28	0.28	1.46	2.07	0.89	3.01	1.00	1.74
T	0.03	0.05	0.02	0.01	0.03	0.03	0.05	0.02	0.01	0.03

Fuente: elaboración propia

En la Tabla XIX se ofrece el resumen de la comparación de los valores trasladados con cada una de las matrices (siguiendo la estrategia recomendada por el INE) respecto a los valores realmente publicados en 2008 con clasificación CNAE-2009. Como se observa, con carácter general el uso de la matriz de transferencia de la Comunidad Valenciana para pasar de una clasificación a otra genera, comparado con el empleo de la matriz de España, soluciones más próximas a los verdaderos valores. Destacan especialmente las reducciones de error que se observan para las ramas J, H y A. El uso de la matriz de la Comunidad Valenciana, sin embargo, tampoco resuelve todos los problemas, pues como se observa no es capaz de ajustarse suficientemente en la fluctuación intertrimestral para los valores trasladados de la rama L, que registra en las cifras oficiales un comportamiento bastante errático. En cualquier caso, y a pesar de lo anterior, el uso de una matriz propia de traslado presenta ventajas evidentes, pues además de la ventaja global que se obtiene hay que señalar que las medias anuales de los valores estimados también son claramente mejores.

5.2.2.2. Sectores

La comparación por ramas de actividad deja claro que el uso de la matriz de la Comunidad Valenciana produce soluciones de mayor calidad y, para determinadas ramas, sensiblemente diferentes a las que se obtienen con la matriz de España. Cuando la comparación se realiza a nivel de sector, sin embargo, se observa (ver la Tabla XX) que las soluciones son bastante parecidas excepto para el sector primario,

donde se obtienen diferencias sistemáticas por encima del 2%. Junto a este resultado general destacan las diferencias que se obtienen para el sector de la Construcción en la provincia de Alicante. En cualquier caso, se podría afirmar que salvo para la Agricultura el uso de una u otra matriz no implica grandes diferencias, aunque alcanzando como norma resultados ligeramente mejores cuando se utiliza la matriz de la Comunidad Valenciana. Esta conclusión viene, en cierta forma, a reforzar las conclusiones que se alcanzaron al analizar los datos de las variables provinciales de la Tabla XVIII obtenidos tras la estrategia agregar-trasladar con la matriz de la Comunidad Valenciana.

Tabla XX. Diferencias relativas en valor absoluto para las series de ocupados de la Comunidad Valenciana por sectores trasladadas con la matriz de la Comunidad Valenciana frente a las trasladadas con matriz de España.

Sector	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	Global
Alicante											
Agricultura	2.36	2.50	2.37	2.36	2.48	2.61	2.56	2.47	2.50	2.38	2.46
Industria	0.37	0.36	0.47	0.56	0.27	0.31	0.30	0.45	0.53	0.52	0.41
Construcción	0.80	1.19	0.73	0.65	0.77	1.58	2.22	1.37	1.06	1.40	1.18
Servicios	0.27	0.24	0.19	0.29	0.17	0.33	0.53	0.44	0.28	0.31	0.31
Castellón											
Agricultura	2.31	2.52	2.18	2.28	2.46	2.29	2.40	2.29	2.29	2.41	2.34
Industria	0.36	0.36	0.37	0.41	0.31	0.28	0.61	0.55	0.55	0.51	0.43
Construcción	0.58	0.39	0.75	0.88	0.56	0.52	0.43	0.73	0.58	0.76	0.62
Servicios	0.04	0.04	0.11	0.18	0.07	0.02	0.12	0.26	0.24	0.20	0.13
Valencia											
Agricultura	2.57	2.56	2.55	2.47	2.45	2.55	2.50	2.41	2.50	2.68	2.52
Industria	0.92	1.13	0.71	0.79	0.81	0.98	0.90	0.83	0.87	0.77	0.87
Construcción	0.54	0.45	0.57	0.70	0.58	0.68	0.83	0.77	0.75	0.76	0.66
Servicios	0.29	0.32	0.34	0.26	0.26	0.33	0.36	0.38	0.33	0.27	0.31
Comunidad Valenciana											
Agricultura	2.45	2.53	2.41	2.40	2.46	2.52	2.50	2.41	2.47	2.52	2.47
Industria	0.64	0.75	0.57	0.65	0.54	0.64	0.65	0.66	0.71	0.65	0.65
Construcción	0.63	0.71	0.64	0.70	0.65	1.04	1.29	0.99	0.84	0.97	0.85
Servicios	0.26	0.26	0.27	0.26	0.20	0.30	0.40	0.39	0.30	0.28	0.29

Fuente: elaboración propia

Por otro lado, comparando, a través de los resúmenes del estadístico D_r , los valores trasladados con los observados, se puede deducir de la Tabla XXI que en el ámbito provincial no parece existir dominancia de las estimaciones obtenidas con una u otra matriz. Las estructuras productivas y las distribuciones de transferencia provinciales no necesariamente están sistemáticamente más próximas a la de la Comunidad Valenciana que a la de España, sino que presentan un comportamiento mixto, aunque, en general, salvo para la provincia de Alicante, se obtienen mejores aproximaciones con la matriz propia. Conjuntamente, en el ámbito autonómico, se observa que el uso de la matriz de la Comunidad Valenciana sí produce una mejora significativa respecto al empleo de la matriz de España. En este caso, se puede afirmar sin género de dudas que a nivel de sector se producen valores trasladados más

próximos a los reales cuando se emplea la matriz de la Comunidad Valenciana que cuando se emplea la matriz de España.

Tabla XXI. Diferencias relativas en valor absoluto respecto a las cifras oficiales EPA, para las series de ocupados de la Comunidad Valenciana y sus tres provincias, de los valores obtenidos tras emplear las matrices de España y de la Comunidad Valenciana.

Sector	Error relativo de series obtenidas con matriz de España (INE)					Error relativo de series obtenidas con matriz Comunidad Valenciana				
	2008.I	2008.II	2008.III	2008.IV	2008	2008.I	2008.II	2008.III	2008.IV	2008
	Alicante					Alicante				
Agricultura	19.45	7.64	2.75	4.76	8.65	16.50	5.13	0.52	2.09	6.06
Industria	0.36	1.03	1.09	0.72	0.80	0.16	1.57	1.65	0.25	0.91
Construcción	0.60	0.63	2.05	2.22	1.38	0.23	0.24	3.62	4.63	2.18
Servicios	0.50	0.46	0.78	0.43	0.54	0.72	0.66	1.28	0.76	0.86
	Castellón					Castellón				
Agricultura	1.36	1.03	1.11	1.92	1.35	1.12	3.40	3.46	4.31	3.07
Industria	0.38	0.27	0.16	1.19	0.50	0.19	0.29	0.29	1.66	0.61
Construcción	1.15	0.17	2.56	2.63	1.63	0.64	0.58	1.81	1.65	1.17
Servicios	0.32	0.24	0.73	0.20	0.37	0.15	0.02	0.55	0.02	0.19
	Valencia					Valencia				
Agricultura	4.11	1.90	0.59	0.61	1.80	6.66	0.84	2.11	2.10	2.93
Industria	1.53	1.32	1.60	0.96	1.35	0.67	0.54	0.91	0.25	0.59
Construcción	1.86	1.91	1.44	2.18	1.85	1.01	1.18	0.82	1.43	1.11
Servicios	1.04	0.73	0.78	0.66	0.80	0.71	0.45	0.55	0.43	0.53
	Comunidad Valenciana					Comunidad Valenciana				
Agricultura	3.96	3.30	1.27	1.94	2.62	1.31	0.73	1.19	0.70	0.98
Industria	0.96	0.40	0.56	0.54	0.62	0.26	0.26	0.05	0.06	0.16
Construcción	1.36	1.29	0.43	0.84	0.98	0.56	0.51	0.52	0.47	0.52
Servicios	0.41	0.25	0.19	0.20	0.26	0.14	0.00	0.14	0.07	0.09

Fuente: elaboración propia

5.3. Comparación cíclica

En el apartado anterior se ha concluido que el traslado entre clasificaciones de los valores de la EPA podría ser mejorado utilizando la matriz de la Comunidad Valenciana. Comparando los valores de 2008 se observó que si se utiliza la matriz de la Comunidad Valenciana se obtienen valores más próximos a los publicados. Ahora bien, ¿qué impacto tiene el uso de una matriz sobre la evolución subyacente y los puntos de giro de las series?

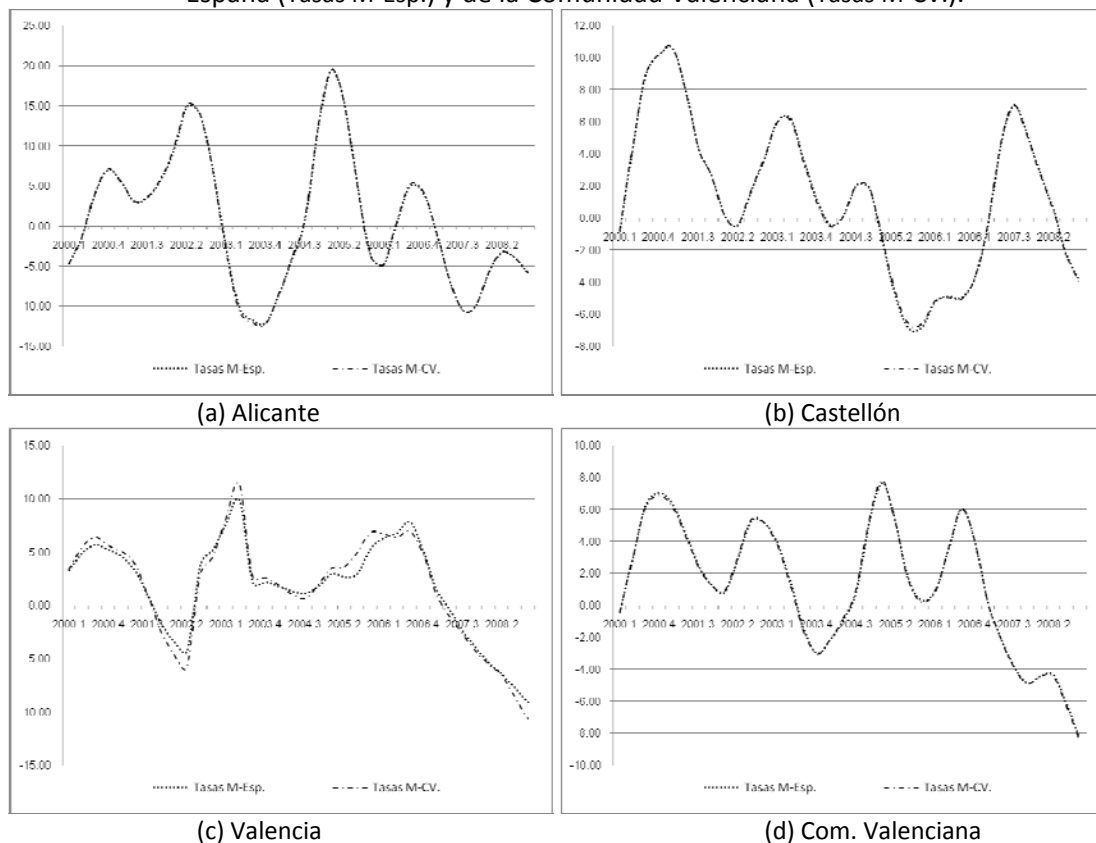
En esta subsección lo que se hace es comparar, para el período 1999.I a 2008.IV, los perfiles cíclicos de las 33 series de ocupados obtenidas tras trasladar con las matrices de España y la Comunidad Valenciana las series disponibles a dos dígitos y posteriormente agregarlas.

5.3.2.1. Sectores

Las señales cíclicas de las series de ocupados por sectores obtenidas no registran, en general, grandes diferencias, como se puede observar a partir de la Figura 6 donde a

modo de ejemplo se representan las señales cíclicas para cada ámbito territorial de industria. Las mayores diferencias se registran para las señales de construcción de Valencia y del agregado de la Comunidad Valenciana, donde al final del período analizado las series obtenidas con la matriz de España apuntan hacia una recuperación de la ocupación en el sector, mientras que las series obtenidas al aplicar la matriz de paso de la Comunidad Valenciana claramente muestran (al igual que ocurre para las series trasladadas para las provincias de Alicante y Castellón) que la ocupación en el sector está lejos de tocar fondo y continúa la caída que inició a mediados de 2007. Los resultados anteriores, sin duda, vienen a reforzar las conclusiones alcanzadas en el análisis en niveles. Si bien a nivel sectorial el uso de cualquiera de las dos matrices produce soluciones bastante parecidas, las soluciones que se obtienen empleando la matriz de la Comunidad Valenciana son claramente preferibles.

Figura 6. Comparaciones de evoluciones cíclicas (medidas en tasas de variación interanual sobre señales de ciclo-tendencia) de las series industriales trasladadas con las matrices de España (Tasas M-Esp.) y de la Comunidad Valenciana (Tasas M-CV.).

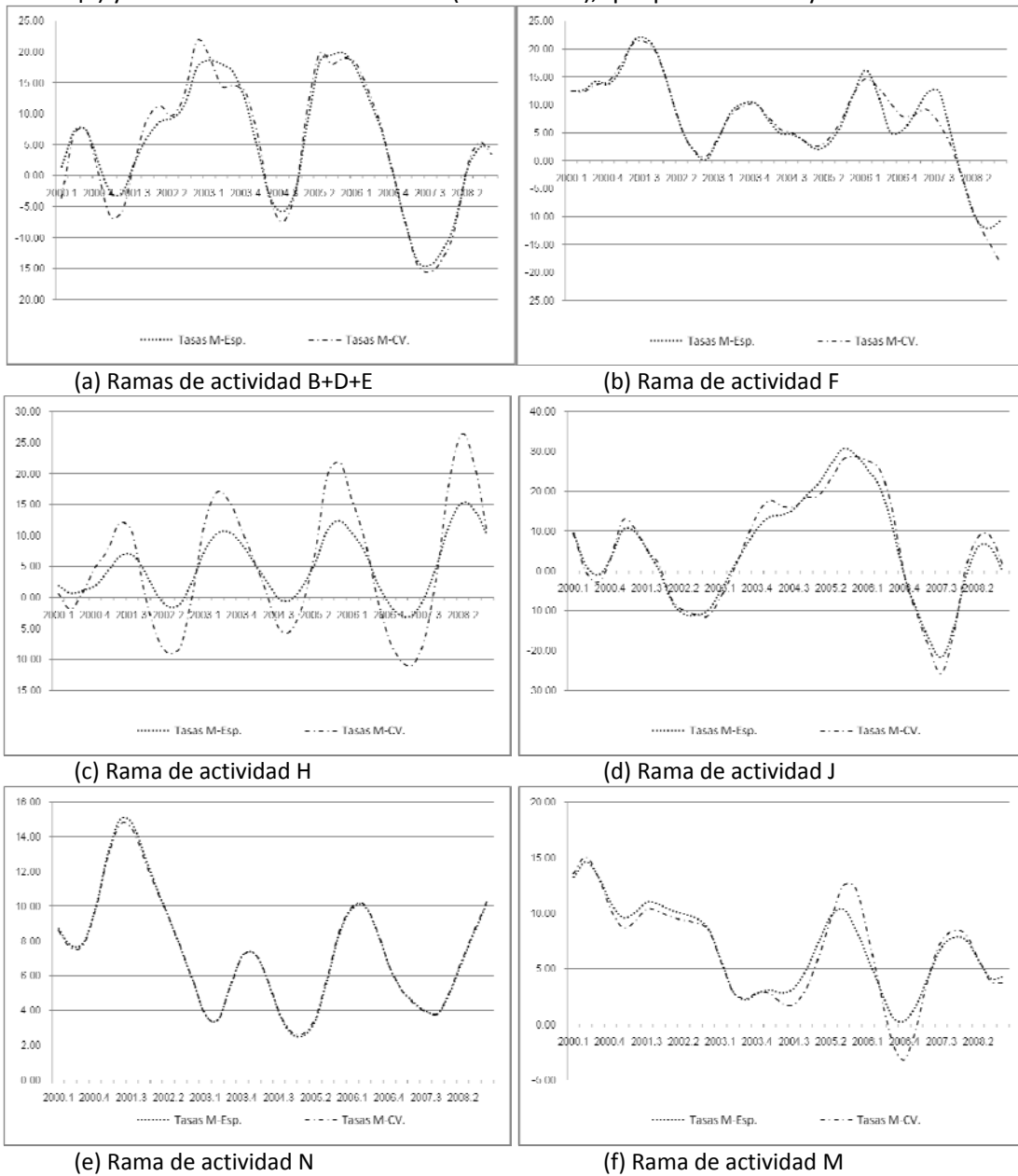


5.3.2.2. Ramas de actividad

Cuando el análisis se realiza sobre las señales cíclicas de las series por ramas de actividad las conclusiones son muy similares a las obtenidas para sectores. Las señales cíclicas para todas las ramas son prácticamente idénticas utilizando cualquiera de las dos matrices. Excepto para unas pocas ramas (B+D+E, F, H, J, N y M), cuyas representaciones de las señales cíclicas se ofrecen en la Figura 9, para el resto las

señales son casi absolutamente coincidentes. Es más, incluso para las ramas donde se observan algunas discrepancias, las diferencias son bastante escasas. Los puntos de giro de todas ellas, exceptuando la rama F para la que se presenta el problema ya reseñado en el apartado anterior al final del período de análisis, son coincidentes. Incluso en la rama H, para la que la serie trasladada con la matriz de la Comunidad Valenciana presenta ciclos con mayor amplitud que los de la serie trasladada con la matriz de España, se observan puntos de giro similares.

Figura 9. Comparaciones de evoluciones cíclicas (medidas en tasas de variación interanual sobre señales de ciclo-tendencia) de las ramas, trasladadas con las matrices de España (Tasas M-Esp.) y de la Comunidad Valenciana (Tasas M-CV.), que presentan mayores diferencias.



6. Conclusiones

Para evitar la pérdida de información histórica que supone la entrada en vigor en 2009 de la nueva Clasificación Nacional de Actividades Económicas en las series de EPA, el INE ha propuesto una metodología que permite el enlace de las series CNAE-1993 y CNAE-2009. El procedimiento de enlace se basa en trabajar con una matriz de paso nacional y el uso de los datos a nivel de división. En el ámbito regional el analista se encuentra, sin embargo, con una falta de disponibilidad de las series desagregadas regionales a nivel de división y una estructura del mercado de trabajo regional diferente a la del conjunto de España. Ante esta doble disyuntiva, este trabajo estudia cuál sería el impacto en las series enlazadas si (i) se trabajase en un nivel de agregación superior al de división y (ii) se utilizase una matriz de paso propia regional. El análisis se ha llevado a cabo para el caso de la Comunidad Valenciana y se centrado en las series de ocupados publicadas por el IVE.

A la luz de los resultados, se deduce que a nivel de rama de actividad el traslado de los valores agregados no parece generar, en general, resultados equivalentes a los que se obtienen cuando se trasladan los valores a nivel de división y posteriormente se agregan. Las causas de estas diferencias para aquellas ramas donde existen mayores diferencias se localizan en las diferentes estructuras productivas y, especialmente, en las diferentes composiciones de las distribuciones de transferencia que, entre divisiones, tienen las ramas de la Comunidad Valenciana y España. Empero lo anterior, es preciso notar que la estrategia recomendada por el INE tampoco es una panacea pues genera resultados claramente mejorables en aquellas agrupaciones donde más intensos han sido los cambios entre clasificaciones. Sectorialmente el panorama es muy diferente. De hecho, se puede afirmar que a nivel de sector y desde una perspectiva práctica la estrategia agregar-trasladar y la estrategia trasladar-agregar (propuesta por el INE) son básicamente equivalentes.

Por otro lado, la comparación de los resultados obtenidos mediante el uso de una matriz de transferencia propia de la Comunidad Valenciana frente a aquellos que se obtienen empleando la matriz de España, muestran que, en general, (i) con la matriz propia se obtienen mejores aproximaciones, que (ii) estas son relativamente de mayor calidad para las series trasladadas a CNAE-2009 y que (iii) por ramas de actividad se observa una mayor ventaja comparativa. A nivel sectorial, no obstante, el uso de cualquiera de las dos matrices produce soluciones bastante similares.

En resumen, se puede afirmar que, al menos para la Comunidad Valenciana, a nivel sectorial, el emplear una matriz propia o la matriz de España, y el trasladar las divisiones y luego agregar o trasladar directamente produce resultados prácticamente equivalentes; mientras que, por otra parte, a nivel de rama de actividad, se observa que, en general, es preferible trabajar con una matriz propia para no desvirtuar el significado de aquellas ramas donde existen más diferencias con España en cuanto a importancia y composición en su distribución de transferencia.

Finalmente, y teniendo en cuenta que los ficheros de microdatos anonimizados de la EPA están disponibles a nivel de grupo (3 dígitos) y el hecho de que como se ha obtenido en este trabajo incluso la estrategia recomendada por el INE genera soluciones claramente mejorables a nivel de sección, sería sumamente interesante abordar, en una futura investigación, el estudio del efecto que tendría en las series

trasladadas el trabajar con una matriz de paso a nivel de grupo en lugar de a nivel de rama. Siempre y cuando los tamaños muestrales lo permitiesen, las conclusiones logradas podrían ser de utilidad, además de para el conjunto de analistas económicos, para el propio INE y para las agencias estadísticas de las regiones de mayor tamaño.

Agradecimientos

La investigación reportada en este artículo describe parte de los trabajos realizados por los autores con motivo del contrato de prestación de servicios “*Recodificación de las Series EPA de la Comunidad Valenciana de acuerdo al último cambio en la Clasificación Nacional de Actividades Económicas, y Análisis de sus Efectos*” suscrito por el Instituto Valenciano de Estadística con la Universitat de Valencia. Los autores desean agradecer a dos evaluadores anónimos sus valiosos comentarios y a los responsables del Servicio de Estadísticas Económicas, Demográficas y Sociales del Instituto Valenciano de Estadística (IVE) y a Miguel Ángel García Martínez del Instituto Nacional de Estadística (INE) su inestimable ayuda, paciencia y respuestas a nuestros interrogantes. Por supuesto, cualquier inexactitud que persista es responsabilidad única de los autores.

Referencias

- INE (2008a) Correspondencia CNAE-93 Rev.1 con CNAE-2009, http://www.ine.es/daco/daco42/clasificaciones/cnae09/cnae93rev1_cnae2009.xls
- INE (2008b) Introducción a la CNAE-2009, http://www.ine.es/daco/daco42/clasificaciones/cnae09/int_cnae_2009.pdf
- INE (2009b) Enlace entre la CNAE 93 y la CNAE 09: matrices de paso de una a otra, http://www.ine.es/daco/daco42/daco4211/matriz_conversion_retro0009.xls
- INE (2009a) Matrices de doble codificación CNAE-93 – CNAE-2009, http://www.ine.es/daco/daco42/daco4211/nota_epa_retro0009.pdf
- INE (2009c) *Clasificación Nacional de Actividades Económicas CNAE-2009*. INE, Madrid.
- Kaiser, R. y Maravall, A. (1999) "Estimation of the business cycle: A modified Hodrick-Prescott filter", *Spanish Economic Review*, 1, 175-206.
- Niemira, M.P. y Klein, P.A. (1994) *Forecasting financial and economic cycles*. John Wiley and Sons Inc., New York.
- Stekler. H.O. (1991) "Turning Point Predictions, Errors and Forecasting Procedures", en Kajal Lahiri y Geoffrey H. Moore (eds.). *Leading Economic Indicators: New Approaches and Forecasting Records*. Nueva York: Cambridge University Press, pp. 169-179.

ANEXO I. Relación entre divisiones y ramas de actividad en series IVE

Tabla A.1. Correspondencia entre divisiones y ramas de actividad con clasificación CNAE-1993.

Rama	Descripción actividad económica	Divisiones
A + B	Agricultura, ganadería, caza y silvicultura; Pesca	01-02, 05
C	Industrias de alimentación, bebidas y tabaco; Textil y confección; Cuero y calzado; Madera y corcho; Papel, edición y artes gráficas	15 - 22
D	Industrias extractivas; Refino de petróleo; Industria química; Transformación de caucho, metalurgia; Energía y agua	10-14, 23-28, 40-41
E	Construcción de maquinaria, material eléctrico, material de transporte e industrias manufactureras diversas	29 - 37
F	Construcción	45
G	Comercio; Reparación de vehículos de motor, motocicletas y ciclomotores y artículos personales y de uso doméstico	50-52
H	Hostelería	55
I	Transporte, almacenamiento y comunicaciones	60-64
J	Intermediación financiera	65-67
K	Actividades inmobiliarias y de alquiler; Servicios empresariales	70-74
L	Administración pública, defensa y seguridad social obligatoria	75
M	Educación	80
N	Actividades sanitarias y veterinarias, servicio social	85
O + Q	Otras actividades sociales y de servicios prestados a la comunidad; Servicios personales; Organismos extraterritoriales (O+Q)	90-93, 99
P	Actividades de los hogares	95

Tabla A.2. Correspondencia entre divisiones y ramas de actividad con clasificación CNAE-2009.

Rama	Descripción actividad económica	Divisiones
A	Agricultura, ganadería, silvicultura y pesca	01-03
B+D+E	Industrias extractivas; Suministro de energía eléctrica, gas, vapor y aire acondicionado; Suministro de agua, actividades de saneamiento, gestión de residuos y descontaminación	05-09, 35-39
C	Industria manufacturera	10-33
F	Construcción	41-43
G	Comercio al por mayor y al por menor; Reparación de vehículos de motor y motocicletas	45-47
H	Transporte y almacenamiento	49-53
I	Hostelería	55-56
J	Información y comunicaciones	58-63
K	Actividades financieras y de seguros	64-66
L	Actividades inmobiliarias	68
M	Actividades profesionales, científicas y técnicas	69-75
N	Actividades administrativas y servicios auxiliares	77-82
O	Administración Pública y defensa; Seguridad Social obligatoria	84
P	Educación	85
Q	Actividades sanitarias y de servicios sociales	86-88
R+S+U	Actividades artísticas, recreativas y de entretenimiento; Otros servicios; Actividades de organizaciones y organismos extraterritoriales	90-93, 94-96, 99
T	Actividades de los hogares como empleadores de personal doméstico; Actividades de los hogares como productores de bienes y servicios para uso propio	97-98

ANEXO II. Códigos y ponderaciones de las divisiones de CNAE-1993 a dos dígitos

Cód.	Descripción	España	C.Val.
01	Agricultura, ganadería, caza y actividades de los servicios relacionados con las mismas	3.95	3.14
02	Selvicultura, explotación forestal y actividades de los servicios relacionados con las mismas	0.16	0.04
05	Pesca, acuicultura y actividades de los servicios relacionados con las mismas	0.24	0.15
10	Extracción y aglomeración de antracita, hulla, lignito y turba	0.03	0.01
11	Extracción de crudos de petróleo y gas natural; actividades de los servicios relacionados con las explotaciones petrolíferas y de gas, excepto actividades de prospección	0.02	0.01
12	extracción de minerales de uranio y torio	0.00	0.00
13	Extracción de minerales metálicos	0.01	0.00
14	Extracción de minerales no metálicos ni energéticos	0.20	0.32
15	Industria de productos alimenticios y bebidas	2.48	2.39
16	Industria del tabaco	0.03	0.02
17	Industria textil	0.42	0.95
18	Industria de la confección y de la peletería	0.42	0.34
19	Preparación curtido y acabado del cuero; fabricación de artículos de marroquinería y viaje; artículos de guarnicionería talabartería y zapatería	0.26	1.31
20	Industria de la madera y del corcho, excepto muebles; cestería y espartería	0.52	0.85
21	Industria del papel	0.21	0.26
22	Edición, artes gráficas y reproducción de soportes grabados	0.90	0.72
23	Coquerías, refino de petróleo y tratamiento de combustibles nucleares	0.09	0.05
24	Industria química	0.98	0.85
25	Fabricación de productos de caucho y materias plásticas	0.54	0.91
26	Fabricación de otros productos minerales no metálicos	1.04	2.45
27	Metalurgia	0.55	0.29
28	Fabricación de productos metálicos, excepto maquinaria y equipo	1.83	2.00
29	Industria de la construcción de maquinaria y equipo mecánico	1.28	1.05
30	Fabricación de máquinas de oficina y equipos informáticos	0.05	0.01
31	Fabricación de maquinaria y material eléctrico	0.44	0.44
32	Fabricación de material electrónico; fabricación de equipo y aparatos de radio, televisión y comunicaciones	0.18	0.19
33	Fabricación de equipo e instrumentos médico-quirúrgicos, de precisión, óptica y relojería	0.18	0.23
34	Fabricación de vehículos de motor, remolques y semirremolques	1.10	0.68
35	Fabricación de otro material de transporte	0.40	0.24
36	Fabricación de muebles; otras industrias manufactureras	1.10	1.49
37	Reciclaje	0.07	0.09
40	Producción y distribución de energía eléctrica, gas, vapor y agua caliente	0.36	0.13
41	Captación, depuración y distribución de agua	0.20	0.21
45	Construcción	11.88	12.95
50	Venta, mantenimiento y reparación de vehículos de motor, motocicletas y ciclomotores; venta al por menor de combustible para vehículos de motor	2.23	2.12
51	Comercio al por mayor e intermediarios del comercio, excepto de vehículos de motor y motocicletas	4.01	5.35
52	Comercio al por menor, excepto el comercio de vehículos de motor, motocicletas y ciclomotores; reparación de efectos personales y enseres domésticos	9.74	10.44
55	Hostelería	7.17	7.91
60	Transporte terrestre; transporte por tuberías	3.04	3.38
61	Transporte marítimo, de cabotaje y por vías de navegación interiores	0.11	0.07
62	Transporte aéreo y espacial	0.26	0.14
63	Actividades anexas a los transportes; actividades de agencias de viajes	1.05	0.75

64	Correos y telecomunicaciones	1.38	0.94
65	Intermediación financiera, excepto seguros y planes de pensiones	1.56	1.29
66	Seguros y planes de pensiones, excepto seguridad social obligatoria	0.66	0.54
67	Actividades auxiliares a la intermediación financiera	0.28	0.41
70	Actividades inmobiliarias	0.96	1.00
71	Alquiler de maquinaria y equipo sin operario, de efectos personales y enseres domésticos	0.23	0.27
72	Actividades informáticas	1.19	0.63
73	Investigación y desarrollo	0.28	0.23
74	Otras actividades empresariales	7.58	6.60
75	Administración pública, defensa y seguridad social obligatoria	6.31	4.21
80	Educación	5.59	5.02
85	Actividades sanitarias y veterinarias, servicio social	6.30	5.35
90	Actividades de saneamiento público	0.40	0.48
91	Actividades asociativas	0.42	0.28
92	Actividades recreativas, culturales y deportivas	2.01	1.93
93	Actividades diversas de servicios personales	1.39	1.44
95	Actividades de los hogares como empleadores de personal doméstico	3.72	4.46
99	Organismos extraterritoriales	0.01	0.01

ANEXO III. Códigos y ponderaciones de las divisiones de CNAE-2009 a dos dígitos

Cod.	Descripción	España	C.Val.
01	Agricultura, ganadería, caza y servicios relacionados con las mismas	3.65	2.85
02	Silvicultura y explotación forestal	0.16	0.01
03	Pesca y acuicultura	0.24	0.15
05	Extracción de antracita, hulla y lignito	0.03	0.00
06	Extracción de crudo de petróleo y gas natural	0.02	0.00
07	Extracción de minerales metálicos	0.01	0.00
08	Otras industrias extractivas	0.20	0.32
09	Actividades de apoyo a las industrias extractivas	0.01	0.01
10	Industria de la alimentación	2.20	2.26
11	Fabricación de bebidas	0.28	0.12
12	Industria del tabaco	0.03	0.02
13	Industria textil	0.39	0.93
14	Confección de prendas de vestir	0.43	0.27
15	Industria del cuero y del calzado	0.28	1.39
16	Industria de la madera y del corcho, excepto muebles; cestería y espartería	0.52	0.86
17	Industria del papel	0.22	0.27
18	Artes gráficas y reproducción de soportes grabados	0.52	0.48
19	Coquerías y refino de petróleo	0.09	0.05
20	Industria química	0.66	0.75
21	Fabricación de productos farmacéuticos	0.32	0.10
22	Fabricación de productos de caucho y plásticos	0.53	0.89
23	Fabricación de otros productos minerales no metálicos	1.04	2.45
24	Metalurgia; fabricación de productos de hierro, acero y ferroaleaciones	0.56	0.29
25	Fabricación de productos metálicos, excepto maquinaria y equipo	1.84	1.97
26	Fabricación de productos informáticos, electrónicos y ópticos	0.27	0.24
27	Fabricación de material y equipo eléctrico	0.48	0.43
28	Fabricación de maquinaria y equipo n.c.o.p.	0.78	0.69
29	Fabricación de vehículos de motor, remolques y semirremolques	1.15	0.71
30	Fabricación de otro material de transporte	0.34	0.18
31	Fabricación de muebles	0.85	1.31
32	Otras industrias manufactureras	0.27	0.32
33	Reparación e instalación de maquinaria y equipo	0.52	0.46
35	Suministro de energía eléctrica, gas, vapor y aire acondicionado	0.36	0.13
36	Captación, depuración y distribución de agua	0.20	0.21
37	Recogida y tratamiento de aguas residuales	0.03	0.05
38	Recogida, tratamiento y eliminación de residuos; valorización	0.34	0.37
39	Actividades de descontaminación y otros servicios de gestión de residuos	0.03	0.04
41	Construcción de edificios	6.36	7.05
42	Ingeniería civil	0.92	0.83
43	Actividades de construcción especializada	4.82	5.44
45	Venta y reparación de vehículos de motor y motocicletas	1.91	1.67
46	Comercio al por mayor e intermediarios del comercio, excepto de vehículos de motor y motocicletas	4.01	5.35
47	Comercio al por menor, excepto de vehículos de motor y motocicletas	9.89	10.79
49	Transporte terrestre y por tubería	3.04	3.38
50	Transporte marítimo y por vías navegables interiores	0.11	0.07

51	Transporte aéreo	0.26	0.14
52	Almacenamiento y actividades anexas al transporte	0.76	0.60
53	Actividades postales y de correos	0.61	0.53
55	Servicios de alojamiento	1.62	0.98
56	Servicios de comidas y bebidas	5.55	6.93
58	Edición	0.40	0.24
59	Actividades cinematográficas, de vídeo y de programas de televisión, grabación de sonido y edición musical	0.20	0.15
60	Actividades de programación y emisión de radio y televisión	0.24	0.14
61	Telecomunicaciones	0.77	0.41
62	Programación, consultoría y otras actividades relacionadas con la informática	1.06	0.57
63	Servicios de información	0.09	0.05
64	Servicios financieros, excepto seguros y fondos de pensiones	1.57	1.30
65	Seguros, reaseguros y fondos de pensiones, excepto Seguridad Social obligatoria	0.66	0.54
66	Actividades auxiliares a los servicios financieros y a los seguros	0.28	0.41
68	Actividades inmobiliarias	0.58	0.57
69	Actividades jurídicas y de contabilidad	1.58	1.46
70	Actividades de las sedes centrales; actividades de consultoría de gestión empresarial	0.32	0.22
71	Servicios técnicos de arquitectura e ingeniería; ensayos y análisis técnicos	1.10	0.92
72	Investigación y desarrollo	0.28	0.23
73	Publicidad y estudios de mercado	0.45	0.31
74	Otras actividades profesionales, científicas y técnicas	0.53	0.39
75	Actividades veterinarias	0.08	0.11
77	Actividades de alquiler	0.24	0.27
78	Actividades relacionadas con el empleo	0.15	0.17
79	Actividades de agencias de viajes, operadores turísticos, servicios de reservas y actividades relacionadas con los mismos	0.31	0.19
80	Actividades de seguridad e investigación	0.73	0.54
81	Servicios a edificios y actividades de jardinería	2.79	2.83
82	Actividades administrativas de oficina y otras actividades auxiliares a las empresas	0.35	0.18
84	Administración Pública y defensa; Seguridad Social obligatoria	6.30	4.21
85	Educación	5.67	5.06
86	Actividades sanitarias	4.40	3.94
87	Asistencia en establecimientos residenciales	1.06	0.95
88	Actividades de servicios sociales sin alojamiento	0.76	0.34
90	Actividades de creación, artísticas y espectáculos	0.28	0.39
91	Actividades de bibliotecas, archivos, museos y otras actividades culturales	0.19	0.16
92	Actividades de juegos de azar y apuestas	0.29	0.29
93	Actividades deportivas, recreativas y de entretenimiento	0.78	0.79
94	Actividades asociativas	0.42	0.28
95	Reparación de ordenadores, efectos personales y artículos de uso doméstico	0.28	0.11
96	Otros servicios personales	1.39	1.44
97	Actividades de los hogares como empleadores de personal doméstico	3.72	4.46
98	Actividades de los hogares como productores de bienes y servicios para uso propio	0.00	0.00
99	Actividades de organizaciones y organismos extraterritoriales	0.01	0.01

ANEXO IV. Comparación entre las diferencias en las distribuciones de transferencia de CNAE-2009 a CNAE-1993 entre España y la Comunidad Valenciana

Tabla A.3. Principales diferencias a nivel de división de distribuciones porcentuales de transferencia entre CNAE-2009 y CNAE-1993 para España y la Comunidad Valenciana.

CNAE-2009	CNAE-1993	España	Comunidad Valenciana	CNAE-2009	CNAE-1993	España	Comunidad Valenciana
5	10	100.00	0.00	32	19	0.10	0.00
6	11	100.00	0.00	32	20	0.61	0.00
7	13	100.00	0.00	32	24	1.23	0.00
9	11	64.57	100.00	32	25	1.59	3.98
9	14	35.43	0.00	32	26	0.13	0.00
63	72	40.25	21.84	32	28	0.68	0.00
63	74	30.11	62.45	32	29	0.22	3.48
63	92	29.64	15.72	32	33	41.26	54.16
27	29	22.55	7.28	32	36	54.18	38.38
27	31	75.69	91.79	58	22	95.07	100.00
27	32	1.76	0.93	58	72	4.93	0.00
26	24	1.72	1.64	35	15	1.00	5.88
26	30	14.26	1.26	35	40	99.00	94.12
26	31	3.91	6.56	52	50	2.66	7.45
26	32	58.89	70.78	52	63	97.34	92.55
26	33	21.22	19.76	38	23	0.09	0.00
				38	37	20.51	24.46
				38	90	79.40	75.54

Tabla A.4. Distribuciones porcentuales de transferencia entre CNAE-2009 y CNAE-1993 a nivel sectorial para España y la Comunidad Valenciana.

		CNAE-1993			
	CNAE-2009	Agricultura	Industria	Construcción	Servicios
España	Agricultura	99.99			0.01
C.Valenciana		100.00			-
España	Industria		97.84	0.00	2.16
C.Valenciana			97.87	0.00	2.13
España	Construcción		0.12	98.01	1.87
C.Valenciana			0.09	97.24	2.68
España	Servicios	0.44	0.66	0.01	98.90
C.Valenciana		0.48	0.40	-	99.12

Tabla A.5. Distribuciones porcentuales de transferencia entre CNAE-2009 y CNAE-1993 a nivel de rama de actividad para España y la Comunidad Valenciana.

CNAE		CNAE-1993														
2009		A+B	C	D	E	F	G	H	I	J	K	L	M	N	O+Q	P
Esp	A	99.99									0.01					
CV	A	100									-					
Esp	B+D+E		0.29	67.12	5.64										26.95	
CV	B+D+E		0.70	58.46	7.91										32.93	
Esp	C	0.00	33.41	34.44	32.06		0.00				0.08					
CV	C	-	37.80	37.51	24.57		-				0.12					
Esp	F			0.05	0.07	98.01					1.87					
CV	F			0.06	0.03	97.24					2.68					
Esp	G						100									
CV	G						100									
Esp	H						0.43		99.57							
CV	H						0.95		99.05							
Esp	I							100								
CV	I							100								
Esp	J		13.91		0.02				28.14		41.36				16.57	
CV	J		15.49		-				26.30		39.13				19.08	
Esp	K									99.84	0.16					
CV	K									99.41	0.59					
Esp	L										100					
CV	L										100					
Esp	M								0.01		98.24			1.75		
CV	M								-		96.86			3.14		
Esp	N	6.47		0.07		0.17			6.69		84.97	0.07			1.56	
CV	N	7.52		-		-			4.48		85.55				2.45	
Esp	O											100				
CV	O											100				
Esp	P										1.25		98.56		0.19	
CV	P										0.63		99.33		0.04	
Esp	Q													100		
CV	Q													100		
Esp	R+S+U	0.03	0.02		1.63		4.28				1.73	0.03			92.28	
CV	R+S+U	-	-		0.55		1.84				0.73	-			96.87	
Esp	T															100
CV	T															100