

Iglesias, C.L.; López, M.E.; Sánchez, P.

Dimensionalidad de la capacidad...

DIMENSIONALIDAD DE LA CAPACIDAD ECONÓMICA EN LAS COMARCAS GALLEGAS

CARLOS L. IGLESIAS PATIÑO* / MARÍA ESTHER LÓPEZ VIZCAÍNO**

PATRICIO SÁNCHEZ FERNÁNDEZ***

*Departamento de Estadística e Investigación Operativa

Facultade de Matemáticas

Universidade de Santiago de Compostela

** Instituto Galego de Estatística (IGE)

***Consello Económico e Social (CES)

Recibido: 24 febrero 2000

Aceptado: 10 octubre 2000

Resumen: En este artículo el objetivo general que se establece es realizar un análisis exploratorio de un conjunto suficientemente amplio de variables que, en un primer momento, están relacionadas con la capacidad económica a nivel comarcal. Dentro de este marco general, en primer lugar, se crea un indicador sintético de la capacidad económica mediante un análisis de componentes principales que permitirá obtener una clasificación según diversas características combinadas. A continuación, se realiza un análisis factorial para resumir y simplificar la información de las variables de caracterización de las comarcas. Finalmente, se lleva a cabo un análisis cluster para conseguir agrupaciones de las comarcas teniendo en cuenta criterios de homogeneidad.

Palabras clave: Análisis de componentes principales / Análisis factorial / Análisis cluster / Análisis económico regional.

DIMENSIONALITY OF THE ECONOMIC CAPACITY WITHIN THE GALICIAN COUNTIES

Summary: The main objective of this paper is to carry out an exploratory analysis of a broad enough group of variables related to the economic capacity at a local level. Within this frame, the first step is to create a synthetic index of the economic capacity. This index will allow to achieve a ranking according to several mixed characteristics. Afterwards, a factorial analysis is carried out. Such an analysis summarizes and simplifies the information about the different variables which characterize the counties (comarcas). The final task consists of a cluster analysis whose aim is to obtain groups of counties according to criteria of homogeneity.

Keywords: Principal components analysis / Factorial analysis / Cluster analysis / Regional economic analysis.

1. INTRODUCCIÓN

El presente artículo tiene por objeto ahondar en el conocimiento de las comarcas gallegas en lo que se refiere a su realidad socioeconómica. Con este objetivo lo que se pretende es dar a conocer un análisis exploratorio de un conjunto amplio de variables que se pensaba, a priori, que podrían estar relacionadas con la “capacidad económica” de las distintas comarcas. También, se trata su distribución geográfica, su localización en Galicia y las dimensiones que la explican.

Si a los problemas que crean la complejidad y la ambigüedad de algunos aspectos del comportamiento humano para las llamadas ciencias sociales se les añaden las dificultades de medida de las variables que intervienen en los fenómenos so-

cioeconómicos, se explica por qué se emplea, a veces, el análisis estadístico para explorar relaciones entre los datos. Así, el análisis exploratorio suele emplearse en nuevas áreas de investigación, como es el ámbito en el que se desenvuelve el presente artículo, e intenta reconocer algún patrón no aleatorio o una estructura explicativa y, además, puede servir para generar hipótesis interesantes para estudios posteriores. En este tipo de análisis no se requieren modelos formales ni definiciones rígidas. También es necesario mencionar el hecho de que las técnicas que se utilizan en este estadio ponen su énfasis en las representaciones claras (gráficos, “rankings”, agrupaciones...) y no en el posible modelo estocástico latente.

Una aproximación descriptiva, desde el punto de vista tradicional, fue la *Base estadística para a obtención de indicadores socioeconómicos das comarcas galegas* (IGE, 1999). Ésta constituye un ejemplo, a la vez, de estudio metodológico previo y una recopilación de tablas sobre la materia. Ahora bien, es conocido que la tabulación, a pesar de ser el método más empleado para presentar los resultados en la estadística oficial, tiene limitaciones, como por ejemplo:

- Las de tipo paradoja de Simpson.
- La visión parcial que ofrece en las situaciones multidimensionales (la tabla es una proyección y, por lo tanto, puede dejar de mostrar variables relevantes del fenómeno de estudio porque simplemente no se introducen en la tabla).
- La dificultad en la lectura y la comprensión de los cuadros con muchas variables (las tablas de tres niveles son difíciles de analizar plenamente). Aunque esto hay que matizarlo por los esfuerzos que están realizando los desarrolladores de software estadístico (p. ej.: tablas pivotales).
- Las limitaciones para poder obtener juicios sintéticos a la vista de los datos presentados.

Por ello, deben emplearse técnicas complementarias a las tradicionales, entre las que se encuentran las de análisis multivariante. Para ahondar sobre este tema, en relación a la explotación de una encuesta, se puede consultar a Teijeiro (1991).

En definitiva, cuando las dimensiones de la matriz de datos resultan un obstáculo para su lectura y asimilación directa, será útil disponer de alguna técnica de reducción de la dimensionalidad original que, mediante su aplicación a los datos originales, permita establecer las relaciones existentes dentro de cada conjunto, así como las interrelaciones entre el conjunto de individuos y el de variables. De la misma manera, resultaría útil evaluar la pérdida de información original que supondría la reducción de la dimensionalidad.

Hacer un análisis multivariante sobre estos datos lleva, primeramente, a una simplificación estructural que posteriormente permite, entre otras cosas, hacer una clasificación de las comarcas. Esta simplificación estructural trae consigo la reducción de la dimensionalidad y el análisis de la interdependencia, con las que se

consigue una descripción abreviada y, por lo tanto, más manejable de las variables iniciales y permite mostrar patrones o relaciones no atribuibles a la casualidad.

El concepto empleado de “capacidad económica” intenta reflejar la potencialidad o el poder de cada zona a la hora del empleo de su renta, bien para el consumo o bien para el ahorro. No se debe confundir, por tanto, con otros términos que aparecen matizados al intervenir la capacidad adquisitiva de sus miembros o su renta disponible. También es importante hacer mención al hecho de que el análisis se está haciendo en términos per cápita, por lo que los datos están relativizados en función del número de habitantes.

El verdadero embrión y marco de referencia de este trabajo es la publicación del Instituto Nacional de Estadística de Portugal (INE), en un trabajo hecho por su Direcção Regional Centro (DRC), conocido como *Estudo sobre o poder de compra concelhio*. A partir de aquí surgió la idea de la realización de un proyecto conjunto entre el Instituto Galego de Estatística (IGE) y la Direcção Regional Norte (DRN) del INE portugués en la búsqueda de este indicador o de alguno relacionado. De esta manera, este proyecto conjunto se configura como el principal referente de nuestro trabajo y, así, explica en gran medida su objetivo, las variables empleadas, el ámbito de referencia y sus unidades territoriales, así como otros aspectos complementarios.

Sin embargo, es importante destacar que el *Estudo sobre o poder de compra concelhio* constituye sólo el origen del presente artículo, ya que a partir de él las modificaciones sufridas en el proyecto en la mayoría de los campos son significativas. Así, en primer lugar, hay que mencionar que las variables utilizadas para la obtención del indicador se alejan de las empleadas por el estudio portugués, en algunos casos por la falta de disponibilidad y en otros por contar con variables más útiles para la medición que se pretendía llevar a cabo. Otra diferencia importante es la unidad territorial de referencia, que en el caso portugués se trata del municipio y que en este caso es la comarca, procurando una mayor homogeneidad entre las dos regiones (IGE, 1999). Asimismo, en lo que respecta al tratamiento estadístico de los datos, en este trabajo se emplea la técnica de rotación *varimax* frente al trabajo portugués que emplea la *quartimax*. Más allá de todo lo anterior, y por constituir un complemento importante al objeto del estudio, se llevó a cabo el análisis de componentes principales previo y el análisis cluster posterior, este último amplía la información que dieron los distintos indicadores.

La obtención de un indicador que esté relacionado de algún modo con nuestra “capacidad económica” a nivel territorial desagregado no es algo nuevo en la literatura económica. De esta manera, son muchos los ejemplos que se podrían citar, destacando como pretéritos, el *Índice agregativo del bienestar a nivel provincial*, recogido en la monografía sobre desarrollo regional del III Plan de Desarrollo de 1972 y la elaboración de un *Indicador complejo del nivel de bienestar de cada provincia*, llevada a cabo por el Servicio Sindical de Estadística de 1976. Desde entonces son muchos los estudios realizados con este o con parecidos objetivos,

como son los casos del *Atlas Comercial de España*, del Banco Central Hispano, el *Anuario del Mercado Español de España*, de Banesto, el *Anuario Comercial*, de la Caixa y, más recientemente, el *Atlas Económico-Comercial de Galicia*, del Correo Galego.

Una vez presentado e introducido el artículo, se refleja brevemente cuál es la estructura que se sigue en él. Así, en primer lugar se centra en las variables empleadas en la fase de recopilación de los datos dando una explicación sobre su consideración. Posteriormente, se describen los métodos estadísticos utilizados en el estudio, que consisten fundamentalmente en un análisis de componentes principales, en un análisis factorial y en un análisis cluster de los datos. La siguiente parte está formada por los distintos resultados del estudio llevado a cabo, tanto en lo que se refiere a la ordenación como a la clasificación de las comarcas según su capacidad económica. También son tratados aquí los factores explicativos como parte fundamental del trabajo. A continuación estos resultados son interpretados llegando a una serie de conclusiones en función de la información disponible y sirviendo de complemento para los resultados reflejados en el apartado anterior. Para terminar el presente estudio, las conclusiones dan paso a la bibliografía y a las referencias, que constituyen un apartado importante del trabajo dada la variedad y diversidad de antecedentes, fuentes y material de trabajo empleados.

2. VARIABLES Y MÉTODOS EMPLEADOS

Como dice Herrero en Junta de Castilla y León (1998), los fenómenos económicos en dimensión espacial tienen una estructura típica multivariante, ya que no se pueden medir directamente mediante una variable sino que son el resultado de un conjunto de características variadas, sobre todo si lo que se quiere estudiar no tiene una definición estándar ampliamente admitida por la comunidad científica, como es este caso.

El presente estudio partió de la publicación del IGE (1999) que contiene un amplio conjunto de variables relacionadas, en la mayoría de los casos, con la “capacidad económica” de las distintas comarcas. Así, la primera etapa fue realizar una revisión de esta base con el fin de recopilar aquellas variables relacionadas estrechamente con el propósito del estudio, aunque mediatizada en parte por el intento de armonización con el trabajo precedente portugués. Desde el punto de vista de la descripción de las variables, éstas se pueden considerar agrupadas de la siguiente manera:

- *Demografía*: incluye las variables referidas a las características de la población, midiendo aspectos como su dinamismo, su distribución por edades y su carácter urbano o rural a través de la “medida de urbanización”.

- *Empleo*: engloba las principales magnitudes referidas al mercado laboral, con una especial atención sobre las actividades de comercio, alojamiento y restauración y las actividades marineras dada su importancia en el caso gallego.
- *Estructura*: agrupa el conjunto de variables económicas que reflejan la estructura funcional, concentración de equipamientos y especialización de cada comarca.
- *Fiscalidad*: recoge información sobre el ámbito tributario y se centra en tres impuestos: el impuesto sobre vehículos de tracción mecánica (IVTM), el impuesto sobre bienes inmuebles (IBI) y el impuesto sobre la renta de las personas físicas (IRPF).

En una segunda fase, la selección de las variables, se realizaron una serie de ensayos factoriales previos teniendo en cuenta criterios habituales en la literatura especializada; de esta manera, las variables finalmente retenidas fueron un total de 21, siendo las siguientes (junto con el año de referencia y con la abreviatura empleada en el análisis):

Demográficas

- Crecimiento de la población, 1991-1996 - *crec*
- Porcentaje de población menor de 25 años, 1996 - *pob_25*
- Porcentaje de población mayor de 65 años, 1996 - *pob_65*
- Población en los núcleos de más de 1000 habitantes / población total. Esta variable se denominó “medida de urbanización”, 1996 - *urbaniz*

Empleo

- Tasa de actividad, 1996 – *taxa_act*
- Número de trabajadores en las empresas por habitante, 1994 - *traballa*
- Número de trabajadores en las empresas de comercio, alojamiento y restauración por habitante, 1994 – *trab_com*
- Número de tripulantes en barcos por habitante, 1994 - *tripulac*

Estructurales

- Viviendas no principales por habitante, 1991 – *viv_nonpr*
- Número de empresas por habitante, 1994 - *empresas*
- Número de empresas de comercio, alojamiento y restauración por habitante, 1994 – *emp_com*
- Número de establecimientos hoteleros y casas rurales por habitante, agosto 1995 - *hoteles*
- Número de habitaciones en los establecimientos hoteleros y casas rurales por habitante, 1995 – *habit_ho*
- Número de buques por habitante, 1994 - *buques*
- Número de restaurantes por habitante, abril 1998 - *resta*

- Número de teléfonos por habitante, 1996 - *telefono*
- Número de oficinas bancarias por habitante, 1996 - *bancos*
- Número de cajeros por habitante, 1998 - *caixeiro*

Fiscales

- Impuesto sobre vehículos de tracción mecánica por habitante, 1993 - *ivtm_h*
- Impuesto sobre bienes inmuebles por habitante, 1993 - *ibi_h*
- Rendimiento medio del IRPF por habitante 1995 - *rendimi*

Con respecto a las variables retenidas, lo primero que hay que mencionar es que como criterio general se procuró utilizar el mayor número de variables relativizadas (algunas ya estaban relativizadas de por sí –tasa de actividad– y otras se relativizaron *a posteriori*) porque ponen de manifiesto la estructura comarcal respecto de un concepto determinado y sustituyen, en cierta manera, la información de la variable expresada en valores absolutos. En segundo lugar, hay que considerar las unidades de las variables, si las variables emplean escalas diferentes los resultados pueden ser equívocos o de difícil interpretación, de ahí que en este caso se llevase a cabo la estandarización de las variables.

Los años de referencia de las variables empleadas pertenecen al período 1993-1996, exceptuando las variables número de restaurantes por habitante y número de cajeros por habitante, para las cuales el año de referencia es 1998, y el número de viviendas no principales por habitante que tiene como año de referencia 1991. No se dispone de datos para años precedentes para las dos primeras, ni de datos para años posteriores en el caso de la última; por lo tanto, estas variables se pueden considerar como *proxys* a la hora de realizar el estudio. El hecho de utilizar estas variables *proxy* está justificado por la importancia que tienen en el objetivo de este artículo.

Una vez seleccionadas las variables, el siguiente paso fue reducir la dimensión del conjunto de variables observables llegando a un número menor que, por su relevancia conceptual, pudiera sustituirlas. Con este objetivo se utilizaron algunas técnicas multivariantes que constituyen uno de los ejemplos más claros de utilización de las técnicas estadísticas en el campo de los estudios regionales. El análisis multivariante se puede definir como un conjunto de métodos estadísticos que analizan, de forma simultánea, dos o más variables observadas, permitiendo obtener una visión en conjunto de los problemas al desentrañar la compleja interacción que siempre existe en los fenómenos regionales.

En el contexto del análisis multivariante se puede entender por modelo una descripción simplificada de la estructura de las observaciones. El propósito de construir un modelo es dotar de una descripción sencilla a la población estudiada que sea consistente con los datos. Desde luego, si el modelo se complica lo suficiente se puede conseguir un buen ajuste, pero un modelo complicado tiene menos

poder explicativo que un modelo sencillo, que siempre resultará más elegante y fácil de interpretar. Por lo tanto, un modelo en esta situación se considerará una manera de simplificar y representar la estructura de un conjunto de datos, y no implicará necesariamente ningún mecanismo causal.

La tendencia lógica del investigador es recoger datos del mayor número posible de variables relacionadas con el tema con el fin de garantizar que ningún elemento esencial de la información se pierda en el proceso de selección de las variables. Pero el manejo de un número muy grande de variables puede resultar confuso y crear dificultades considerables en el proceso del análisis, y también puede ocurrir que parte de esas variables sean redundantes porque recogen parcialmente la misma información. De esta forma, algunos métodos multivariantes consiguen combinar variables y eliminar información redundante obteniendo como resultado unos indicadores nuevos. Dentro de las técnicas multivariantes están las que algunos autores denominan técnicas factoriales y, entre ellas, el análisis de componentes principales (ACP) y el análisis factorial (AF). Las dos buscan la mayor reducción de la dimensionalidad de la matriz de datos, siempre que se mantenga una porción considerable de información inicial, aunque ésta se valore atendiendo a diferentes criterios en las dos técnicas.

El ACP tiene como objetivo sintetizar la información contenida en los datos, substituyendo las variables originales por otras nuevas, de tal manera que, si se considera un número menor de estas últimas, expliquen una porción suficientemente grande de varianza inicial. Estas variables nuevas se llaman componentes principales y no contienen información redundante, en el sentido de estar incorreladas entre sí. Se procede, iterativamente, seleccionando una primera combinación lineal de las variables iniciales (primera componente) que resuma el mayor porcentaje de varianza total y que constituya la mejor síntesis unidimensional de los datos. Posteriormente, se elige una segunda combinación lineal (segunda componente), incorrelada con la anterior y con el siguiente valor más alto de la varianza. De esta forma se van seleccionando las sucesivas componentes, que explican cada vez porciones más pequeñas de la varianza, todas ellas incorreladas entre sí. A menudo, un número de componentes principales reducido explica un porcentaje elevado, manteniendo casi toda la información inicial.

Este método es muy frecuente en la construcción de indicadores regionales, la primera componente es la mejor explicación unidimensional de la información y constituye, en general, un excelente indicador. Sin embargo, como ocurre en este caso, cuando la variabilidad explicada por ella no es suficiente, hay que incorporar más información añadiendo las componentes sucesivas, previamente ponderadas en función de su capacidad explicativa hasta obtener un indicador sintético con un nivel de explicación adecuado (Butler y Peters, 1970; Martín Guzmán, 1988). El criterio que se utilizará para la formación de este indicador es que la varianza acumulada por las componentes empleadas en su formación sea igual a la de este nuevo indicador. Con este objetivo, entre otros, se realizó el ACP en este artículo.

Por otro lado, la finalidad del AF es expresar, si es posible, un conjunto de variables en función de unas pocas características subyacentes, pero no observables, llamadas factores. Por lo tanto, el AF no es sólo una técnica de reducción de la dimensionalidad sino que se basa en un modelo explícito de relaciones entre las variables –en este caso el modelo tradicional: el lineal–, que interpreta los factores como características que las variables miden en común. Además, establece que son los niveles o valores que las comarcas poseen de estos factores los que influyen en mayor o menor medida sobre los observados en las variables. De esta manera se puede motivar el modelo factorial del siguiente modo: supóngase que las variables originales puedan ser agrupadas por sus correlaciones lineales, es decir, que las variables correspondientes a un determinado grupo están muy correlacionadas entre ellas mismas, pero tienen relativamente pequeñas correlaciones con variables de otro grupo diferente. Entonces se puede pensar que cada grupo de variables representa un constructo subyacente, el factor, que es responsable de las correlaciones observadas.

La extracción de factores, que son los que van a permitir interpretar la información contenida en las variables originales, se puede realizar mediante diversos métodos. En este estudio se empleó el método de las componentes principales, lo que continúa o amplía el estudio inicial realizado mediante dichas componentes. El AF parte de la matriz de correlaciones, la misma que se utiliza en el ACP en este trabajo, dado que se trabaja con las variables estandarizadas, y de la conjetura de que dichas correlaciones no son fortuitas sino que se deben a que las variables comparten causas comunes llamadas factores. Este tipo de AF se puede considerar una extensión del ACP. Los dos pueden ser vistos como un intento de aproximar la matriz de correlaciones. Sin embargo, la aproximación basada en el modelo del AF es más elaborada.

Para tener una visión más inteligible de la situación, lo ideal es separar las variables en conjuntos disjuntos, cada uno de los cuales se asocia a un factor. Para esto hay que transformar la matriz de carga obtenida, mediante una matriz ortogonal, que mantendrá la misma proporción de variabilidad explicada. La transformación es equivalente a una rotación de ejes del espacio original. En este sentido, se hizo una rotación por el método varimax. Este método se basa en la obtención de factores, de tal manera que cada uno de ellos tenga cargas grandes para pocas variables y muchas cargas próximas a cero para el resto de las variables.

El modelo utilizado es el siguiente: $X=FA+U$, donde X es una matriz que tiene por columnas las variables y por filas los valores de las variables en cada una de las comarcas. La matriz F está constituida por los valores desconocidos que los factores toman en cada una de las comarcas, siendo la matriz A la constituida por los coeficientes, también desconocidos, que cuantifican la contribución marginal de cada factor para la explicación de cada variable. Finalmente, la matriz U está constituida por los factores específicos, es decir, comprende la parte de cada variable no explicada por los factores comunes.

El AF suele realizarse en cuatro etapas fundamentales:

- Cálculo de la matriz de correlaciones entre todas las variables.
- Extracción de los factores que van a representar la totalidad de los casos.
- Rotación, que consiste en la transformación de los factores para hacerlos más interpretables.
- Cálculo de los valores de los factores para cada caso.

Una vez conseguido el objetivo de reducción de la dimensión y simplificación estructural, es de interés conocer la puntuación de las comarcas en los factores, pues puede emplearse para análisis posteriores como la agrupación homogénea de comarcas mediante el análisis cluster, que es el que se realizará a continuación. El cálculo de estas puntuaciones se realizará a partir de la matriz factorial rotada y basándose en el modelo de regresión lineal múltiple. Las puntuaciones permiten determinar la importancia de los factores en cada una de las comarcas.

La pretensión última del análisis cluster es el establecimiento de grupos entre observaciones, de manera que las agrupaciones resultantes presenten la mayor homogeneidad posible dentro de cada grupo y la mayor heterogeneidad entre grupos diferentes, medidas ambas de acuerdo con criterios establecidos en función de los objetivos del estudio. Los elementos a agrupar vendrán caracterizados por los valores que en ellos toman ciertas variables, que pueden ser muy distintos de unos elementos a otros.

Calculando una matriz de distancias de las observaciones, el análisis cluster proporciona una clasificación de las comarcas con base en la información disponible, es decir, en función de la similitud de todas las variables consideradas simultáneamente.

Para clasificar las comarcas se siguieron dos etapas:

- Medición de la similitud entre ellas.
- Clasificación de las comarcas en cada cluster.

La matriz de proximidades, que constituye el cálculo inicial para medir el grado de homogeneidad entre las comarcas estudiadas, se construye a partir del concepto de distancia euclídea, y el algoritmo de agrupación empleado fue el de las K -medias, según el cual las comarcas se van agrupando al centro más próximo y éste se va actualizando a medida que se añade una comarca nueva. Este procedimiento comienza usando los valores de k comarcas tomadas aleatoriamente como estimaciones temporales de las medias de los k clusters, donde k es el número de clusters especificado. Los centros iniciales de los clusters se forman asignando cada comarca al cluster que tenga el centro más próximo a ella. A continuación comienza un proceso iterativo hasta encontrar los centros de los clusters finales. A cada paso, las comarcas se agrupan en el cluster que tiene el centro más próximo, y los centros de los clusters se vuelven a calcular. Este proceso continua hasta que no se

producen cambios muy apreciables en los centros o hasta que se alcanza el máximo número de iteraciones. Esta última situación no se presenta en este estudio.

3. RESULTADOS

Antes de dar comienzo al ACP y al AF se hicieron una serie de ensayos factoriales que tuvieron bastante incidencia en las variables seleccionadas finalmente. Como consecuencia de éstos se puede mencionar:

- Eliminación de variables de baja comunalidad.
- Eliminación de variables que contribuían a crear factores con poca varianza explicada por lo que hacían crecer el número de factores y eran muy pobres en contenido.
- Exclusión de variables con poca carga factorial, pues no contribuían mucho a la explicación total.

En primer lugar, es necesario resaltar que las técnicas de componentes principales no necesitan la hipótesis previa de normalidad sobre los datos, hipótesis que en la mayoría de los casos no se cumple y, por lo tanto, es arriesgado asumirla (Martín Guzmán, 1988).

La matriz de datos de la que se parte incluye 21 variables originales medidas en 53 comarcas convenientemente estandarizadas. Del análisis de esta matriz de datos se obtienen los coeficientes que expresan la transformación lineal de las 21 variables originales en las 21 componentes principales. Cuanto mayor sea el coeficiente más importancia tiene la variable en la componente particular. A partir de estos coeficientes se puede calcular la puntuación de cada comarca en cada componente.

La importancia de cada componente se mide en función de la varianza explicada por ella poniéndola en relación con la varianza total.

Una vez obtenidas las componentes es útil obtener una ordenación de las comarcas teniendo en cuenta estos nuevos ejes. La ordenación de las comarcas con respecto a la primera y a la segunda componente se muestra en la tabla 2.

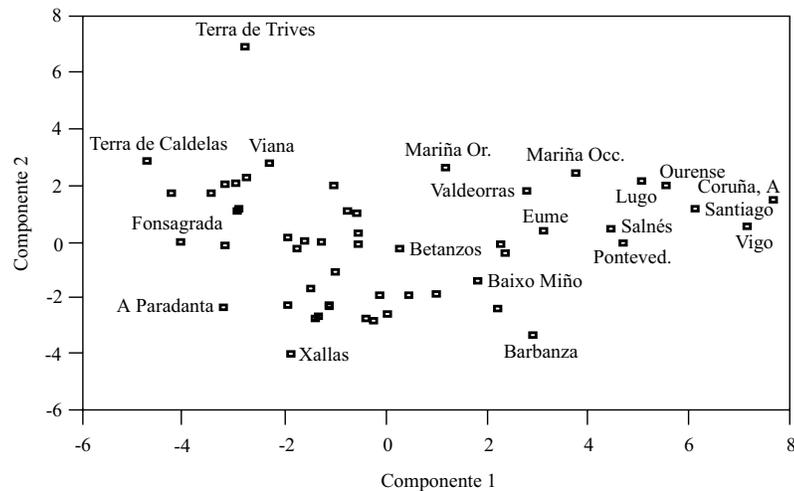
Tabla 1.- Autovalores y varianza explicada por cada componente

COMPONENTE	AUTOVALOR	% VARIANZA	% ACUMULADO
1	9,29	44,28	44,28
2	4,42	21,04	65,32
3	2,40	11,45	76,76
4	1,39	6,62	83,38
5	0,58	4,74	86,12
6	0,50	2,39	88,52

Es de interés, también, ver la posición en el plano de las componentes de las distintas comarcas, así se podrá valorar la importancia de las comarcas en cada una de las componentes principales.

Tabla 2

	1ª COMPONENTE	2ª COMPONENTE
1	Coruña, A	Terra de Trives
2	Vigo	Terra de Caldelas
3	Santiago	Viana
4	Ourense	Mariña Oriental, A
5	Lugo	Mariña Occidental, A
6	Pontevedra	Quiroga
7	Salnés, O	Lugo
8	Mariña Occidental, A	Ribeiro, O
9	Eume	Allariz-Maceda
10	Barbanza	Terra de Lemos
11	Valdeorras	Ourense
12	Ferrol	Valdeorras
13	Mariña Central, A	Baixa Limia
14	Morrazo, O	Ancares, Os
15	Baixo Miño, O	Coruña, A
16	Mariña Oriental, A	Terra de Celanova
17	Sar, O	Santiago
18	Condado, O	Carballiño, O
19	Betanzos	Limia, A
20	Noia	Verín
21	Bergantiños	Vigo
22	Fisterra	Salnés, O
23	Caldas	Eume
24	Sarria	Sarria
25	Tabeirós-Terra de Montes	Meira
26	Verín	Ortegal
27	Carballiño, O	Terra Chá
28	Muros	Fonsagrada, A
29	Terra de Lemos	Pontevedra
30	Ordes	Mariña Central, A
31	Deza	Tabeirós-Terra de Montes
32	Terra Chá	Ulloa, A
33	Barcala, A	Betanzos
34	Terra de Soneira	Chantada
35	Terra de Melide	Ferrol
36	Ortegal	Muros
37	Chantada	Baixo Miño, O
38	Xallas	Terra de Melide
39	Meira	Sar, O
40	Arzúa	Bergantiños
41	Viana	Condado, O
42	Quiroga	Arzúa
43	Terra de Trives	Deza
44	Terra de Celanova	Ordes
45	Limia, A	Paradanta, A
46	Ribeiro, O	Morrazo, O
47	Allariz-Maceda	Noia
48	Ulloa, A	Barcala, A
49	Paradanta, A	Caldas
50	Ancares, Os	Terra de Soneira
51	Fonsagrada, A	Fisterra
52	Baixa Limia	Barbanza
53	Terra de Caldelas	Xallas

Gráfico 1.- Gráfico de dispersión de la 1ª y de la 2ª componente

La primera componente explica el 42,73% de la varianza total, constituye un buen indicador pero su variabilidad explicada no es suficiente, de ahí que se necesite la segunda componente para crear un indicador con un nivel de explicación suficiente. Para la creación de este indicador sintético se calculó la suma ponderada de las dos primeras componentes (previamente estandarizadas) empleándose como ponderación de las componentes la raíz cuadrada de cada varianza. De esta forma, el indicador así obtenido, que se llamará indicador sintético de la capacidad económica (ISCE), ofrecerá una ordenación de las comarcas teniendo en cuenta las dos componentes combinadas. Así pues, la expresión será:

$$\text{ISCE} = \text{componente_st } 1 * \sqrt{9,29} + \text{componente_st } 2 * \sqrt{4,42}$$

En este sentido, las comarcas ordenadas de acuerdo con este nuevo indicador (ISCE) aparecen en la tabla 3.

Una vez determinadas, se busca la interpretación de cada una de las componentes. La matriz de saturaciones no permite tomar posturas nominalistas, e incluso aceptar como buena esta primera solución puede llevar a confusiones. El criterio de maximizar la varianza explicada por las componentes puede llevar a obtener componentes de difícil interpretación que tienen pesos importantes en casi todas las variables, como ocurre en este caso. Todas estas consideraciones llevarán a realizar el AF con fines interpretativos.

El AF parte del cálculo de la matriz de correlaciones, ya que precisamente uno de los objetivos es la obtención de factores que ayuden a explicar las correlaciones entre las variables. Uno de los requisitos que se debe cumplir para realizar el AF es

que las variables estén altamente correlacionadas. Haciendo un examen de esta matriz, se observa que la mayoría de las correlaciones son altas. Para complementar el estudio de la adecuación de la muestra, a la hora de realizar este tipo de análisis se hicieron una serie de pruebas, entre las que destacan el test de esfericidad de Bartlett ($p=0,000$), el cálculo del determinante de la matriz de correlaciones que nos ofrece un resultado muy bajo, que significa que las variables tienen intercorrelaciones muy altas, el indicador de adecuación simple Kaiser-Meyer-Olkin, que ofrece un valor considerado mediano (0,79) aunque próximo a meritorio, el cual significa que la suma de los coeficientes de correlación parcial entre todos los pares de variables es pequeño, y la anti-imagen de la matriz de correlación, que presenta unos valores en la diagonal bastante altos (todos superiores a 0,5). Todas estas pruebas constituyen un síntoma de que el AF puede considerarse aceptable.

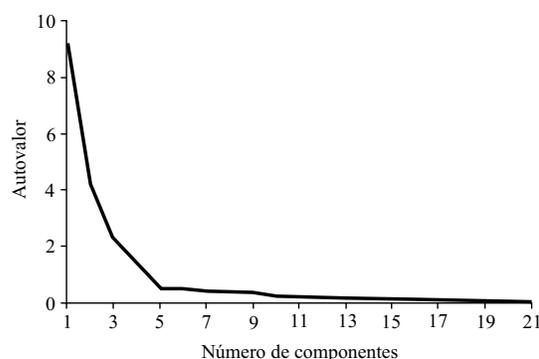
Tabla 3

1	Coruña, A	28	Allariz-Maceda
2	Vigo	29	Terra Chá
3	Ourense	30	Condado, O
4	Santiago	31	Ortegal
5	Lugo	32	Terra de Celanova
6	Mariña Occidental, A	33	Ancares, Os
7	Salnés, O	34	Meira
8	Pontevedra	35	Terra de Caldelas
9	Valdeorras	36	Limia, A
10	Terra de Trives	37	Bergantiños
11	Mariña Oriental, A	38	Chantada
12	Eume	39	Muros
13	Mariña Central, A	40	Baixa Limia
14	Ferrol	41	Noia
15	Terra de Lemos	42	Fisterra
16	Viana	43	Caldas
17	Verín	44	Terra de Melide
18	Baixo Miño, O	45	Ulloa, A
19	Carballiño, O	46	Deza
20	Betanzos	47	Ordes
21	Morrazo, O	48	Barcala, A
22	Sarria	49	Fonsagrada, A
23	Barbanza	50	Terra de Soneira
24	Quiroga	51	Arzúa
25	Tabeirós-Terra de Montes	52	Paradanta, A
26	Sar, O	53	Xallas
27	Ribeiro, O		

La determinación del número de factores fue una cuestión extensamente tratada y se han propuesto diversos criterios basados muchas veces en la experiencia y en la intuición. Teniendo en cuenta el porcentaje de varianza explicada por cada factor, en este caso se decidió optar por los criterios de tomar las componentes en las que los valores propios son superiores a 1 y el de utilizar el gráfico de sedimentación en el cual se escogen los autovalores correspondientes hasta que la pendiente

de la gráfica deja de decrecer bruscamente. Empleando ambos criterios se obtienen 4 factores, que explican un total del 83,4% de la varianza total, con lo cual el objetivo de síntesis parece conseguido.

Gráfico 2.- Gráfico de sedimentación



Otro instrumento que ayuda a evaluar la bondad de ajuste de la AF es la comunalidad de cada variable, es decir, el porcentaje de variación total de una variable explicada por los factores que, por tanto, indica hasta qué punto los factores ayudan a explicar la variable. Las comunalidades son todas superiores a 0,67, es más, están todas entre el 67% y el 94%, como se muestra en la tabla 4.

Tabla 4.- Comunalidades

VARIABLES	COMUNALIDADES
viv nonpr	0,77
emp com	0,84
empresas	0,85
trab com	0,89
traballa	0,84
buques	0,94
tripulac	0,93
hoteis	0,90
habit ho	0,87
resta	0,74
telefono	0,78
bancos	0,77
caixeiro	0,81
pob 65	0,92
pob 25	0,92
urbaniz	0,86
ibi h	0,80
ivtm h	0,67
rendimi	0,83
crec	0,79
tasa act	0,75

Una vez evaluada la calidad del AF, es necesario realizar la interpretación de los resultados. Para esto es necesario obtener los coeficientes que relacionan las variables originales con cada uno de los factores una vez realizada la rotación varimax. Estas relaciones constituyen la clave de su interpretación (tabla 5).

Tabla 5.- Matriz de componentes rotadas

VARIABLES	FACTOR 1	FACTOR 2	FACTOR 3	FACTOR 4
viv nonpr	0.00	-0.88	0.07	-0.04
emp com	0.83	0.36	0.15	0.02
empresas	0.83	0.39	0.09	0.00
trab com	0.86	0.38	0.07	-0.05
traballa	0.82	0.40	0.05	0.02
buques	-0.03	0.21	0.07	0.94
tripulac	0.05	0.18	0.11	0.94
hoteis	0.00	-0.30	0.89	0.09
habit ho	0.23	0.03	0.89	0.14
resta	0.18	0.05	0.84	-0.02
telefono	0.80	-0.36	0.14	0.07
bancos	0.09	-0.85	0.20	-0.02
caixeiro	0.87	0.14	0.14	0.10
pob 65	-0.33	-0.84	-0.01	-0.33
Pob 25	0.25	0.85	0.03	0.35
Urbaniz	0.88	0.23	-0.03	0.18
lbi h	0.88	-0.13	0.10	-0.03
lvtm h	0.74	-0.26	0.23	-0.09
Rendimi	0.84	0.34	0.04	0.01
Crec	0.53	0.70	0.02	0.16
Tasa act	0.25	0.83	0.00	0.06

Teniendo en cuenta las variables que tienen más peso en cada factor, la interpretación de los resultados es la siguiente.

Primer factor: aparece definido por las correlaciones positivas con el rendimiento medio, número de empresas, número de empresas de comercio, alojamiento y restauración, número de trabajadores, número de trabajadores en las empresas de comercio, alojamiento y restauración, número de cajeros, número de teléfonos, impuesto sobre vehículos de tracción mecánica, impuesto sobre bienes inmuebles, todas ellas en términos per cápita, y la “medida de urbanización”. Este factor se denominará *dinamismo económico* dada su importancia para la evolución de las distintas comarcas al estar correlacionado positivamente con las variables demográficas, laborales, estructurales y fiscales con mayor incidencia para una evolución económica favorable.

Segundo factor: bipolar, relacionado, por un lado, con las variables número de viviendas no principales por habitante, número de oficinas bancarias por habitante y porcentaje de población mayor de 65 años frente al crecimiento de la población, a la tasa de actividad y al porcentaje de la población menor de 25 años. Este factor enmarca la *estabilidad socioeconómica* de las comarcas en clara contraposición al primero, puesto que las variables con incidencia positiva se refieren al asentamiento y consolidación en los campos social y financiero, mientras que las variables

con incidencia negativa recogerían el aspecto contrario a esta estabilidad socioeconómica.

Así pues, estos dos primeros factores aparecen de alguna manera como complementarios entre ellos. Este hecho es debido a que, mientras que el primer factor estaría asociado con el concepto de renta como flujo corriente de bienes y servicios, el segundo llevaría a un concepto más estático, o de riqueza, que se refiere al stock de capital que permite crear ese flujo de renta, es decir, la consecuencia o el efecto de las variables del primer factor (Muñoz Ciudad, 1992).

Tercer factor: delimita claramente las comarcas *turísticas*. Aparece relacionado positivamente con el número de restaurantes, número de establecimientos hoteleros y casas rurales y el número de plazas en estos establecimientos, todas ellas per cápita.

Cuarto factor: aparece altamente relacionado con las variables relativas a la *pesca*. Delimita las comarcas donde las actividades pesqueras y actividades conexas son una de las principales fuentes de ingresos.

Una vez resumida la información proporcionada por las 21 variables en cuatro factores mediante el AF, se aplicó el análisis cluster a la realidad comarcal de Galicia basándose en los valores que tomaban los factores para cada una de las comarcas, ya que los factores resumen y simplifican la mayor parte de la información inicial además de eliminar redundancia y, por tanto, pueden sustituir a los datos originales.

Se realizaron una multitud de ensayos de agrupación imponiendo, *a priori*, sucesivos números de conglomerados, de forma que llegamos, finalmente, a una clasificación en seis clusters considerada como la más adecuada para los objetivos de la investigación. Los valores de los cuatro factores en los centros de los seis clusters son los que se recogen en la tabla 6.

Tabla 6.- Valores de los factores en los centros de los clusters

	CLUSTER 1	CLUSTER 2	CLUSTER 3	CLUSTER 4	CLUSTER 5	CLUSTER 6
Factor 1	0.50	1.79	-0.21	-0.65	0.35	-0.36
Factor 2	0.31	0.46	-1.33	0.51	0.36	0.29
Factor 3	2.78	-0.28	0.15	-0.31	-1.33	-0.08
Factor 4	0.52	-0.42	-0.24	-0.59	4.16	1.53

El número de comarcas que se incluyen en cada cluster se muestran en la tabla 7.

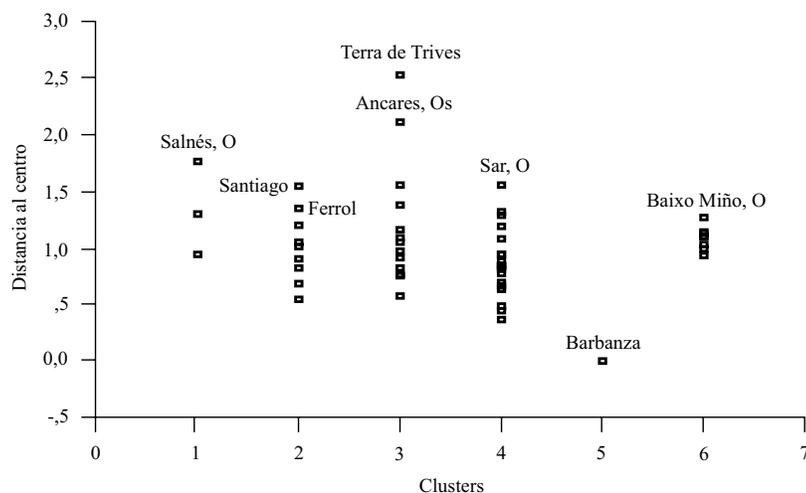
Tabla 7.- Número de comarcas en cada cluster

CLUSTER	Nº DE COMARCAS
1	3
2	9
3	13
4	19
5	1
6	8

La interpretación de las características de las comarcas incluidas en cada cluster se hace a partir de los valores de los factores en el centro del cluster (tabla 6). Mediante esto se pueden intuir los valores de las variables originales en cada uno de los clusters, y son estas variables las que, finalmente, caracterizarán a cada uno de ellos. A continuación se analizan las comarcas incluidas en los clusters y sus características:

- Cluster 1 - *Turístico*: A Mariña Oriental, A Mariña Occidental y O Salnés. Todas ellas son comarcas caracterizadas por un fuerte peso en el factor del turismo y con bastante dinamismo económico, menor, eso sí, que el del cluster de las *Grandes ciudades*. Tienen también cierta importancia pesquera, son todas comarcas costeras donde se llevan a cabo actividades relacionadas con el mar. La comarca que se encuentra más lejana del centro del cluster es O Salnés.
- Cluster 2 - *Grandes ciudades*: A Coruña, Ferrol, Eume, Lugo, Valdeorras, Ourense, Pontevedra, Vigo y Santiago, es decir, las que tienen un mayor peso en el primer factor. Son todas ellas comarcas que tienen un alto rendimiento medio del IRPF y que cuentan con un número bastante alto de empresas. Estas comarcas son las de mayor población de Galicia, a excepción de las comarcas del Eume y de Valdeorras, y todas ellas cuentan con una alta “medida de urbanización”. Las comarcas que se encuentran más lejanas del centro son Santiago y Ferrol. La comarca de Ferrol se caracteriza por tener en términos per cápita un número de empresas y trabajadores de comercio, alojamiento y restauración considerablemente menor al resto de las comarcas incluidas en este cluster.
- Cluster 3 - *Galicia interior*: O Carballiño, O Ribeiro, Terra de Celanova, Allariz-Maceda, Terra de Caldelas, Terra de Lemos, Os Ancares, Quiroga, Terra de Trives, Viana, Verín, A Limia y Baixa Limia. Son todas comarcas de la provincia de Ourense y del sur de Lugo, con un peso muy importante en el segundo factor, que destacan por tener tasas de actividad muy bajas y crecimientos de la población bajos debido, en parte, a su población envejecida. La comarca que se encuentra más lejos del centro es la de Terra de Trives, caracterizada por tener el número de bancos y viviendas no principales por habitante más elevado de toda Galicia, es decir, tiene un peso muy elevado en el segundo factor.
- Cluster 4 - *Grandes vilas*: A Fonsagrada, Meira, Terra Chá, Betanzos, Terra de Melide, Arzúa, Ordes, Bergantiños, Xallas, A Barcala, O Sar, Caldas, Tabeirós-Terra de Montes, Deza, A Ulloa, Chantada, Sarria, O Condado y A Paradanta. Caracterizadas por la poca importancia del primer factor en todas ellas y por contar, en términos relativos y comparándolas con las comarcas de la *Galicia interior*, con población más joven y con tasas de actividad más altas. Las comarcas que se sitúan más lejos del centro del cluster son O Sar y Xallas, debido a que son comarcas con una tasa de actividad muy superior a las del resto del cluster.

- Cluster 5 - *Barbanza*: esta comarca única destaca del cluster *Marinero* porque tiene una importancia muy superior como comarca pesquera y, además, tiene un peso muy superior en el primer factor, aunque más bajo en el factor del turismo.
- Cluster 6 - *Mariñeiro*: A Mariña Central, Ortegal, Terra de Soneira, Fisterra, Muros, Noia, O Morrazo y Baixo Miño. Son comarcas costeras con un porcentaje bastante alto de habitantes dedicados a la pesca. Las comarcas que se encuentran más lejos del centro son el Baixo Miño y la Terra de Soneira, que se caracterizan por tener una importancia menor en cuanto a número de habitantes dedicados a las actividades pesqueras en relación con las otras comarcas del cluster.

Gráfico 3.- Distancias a los centros de los clusters

Para comparar los distintos clusters y ver en qué medida son distintos los unos de los otros, la matriz de distancias entre los centros de los clusters indica los que son más parecidos (tabla 8).

Tabla 8.- Distancias entre los centros de los clusters

CLUSTER	1	2	3	4	5	6
1						
2	3,45					
3	3,27	2,72				
4	3,49	2,45	1,98			
5	5,49	4,92	4,97	4,96		
6	3,15	2,90	2,41	2,15	3,01	

Observando la matriz de distancias se puede comprobar que el cluster del *Barbanza* presenta los mayores valores, confirmando que este cluster tiene unas caracte-

terísticas totalmente diferentes a los demás. Esta característica justifica el hecho de que Barbanza, como comarca única, forme un cluster independiente. La menor distancia corresponde al cluster de la *Galicia interior* y al cluster de las *Grandes villas*, que se caracterizan por tener poco dinamismo económico, aunque diferentes tasas de actividad e índices de envejecimiento de la población. El cluster de las *Grandes villas* cuenta con comarcas con un grado de actividad mayor y con una población menos envejecida. El cluster de la *Galicia interior* tiene un número de bancos y un número de viviendas no principales por habitante muy superior al cluster de las *Grandes villas*.

Por tanto, al ser distintas las características de los clusters con menores distancias, no parece necesario reducir ni aumentar su número.

4. INTERPRETACIÓN Y CONCLUSIONES

A la hora de establecer las conclusiones sobre el estudio aquí realizado, primeramente es necesario destacar las diferencias que surgen frente al embrión de este trabajo que es el *Estudo sobre o poder de compra concelhio*. En este sentido, hay que mencionar que la primera componente no explica tanta varianza como en el estudio portugués y, por lo tanto, se necesita la segunda componente para alcanzar un nivel de explicación similar para el caso gallego. Después de las rotaciones, el tercer factor, en el caso gallego, y el segundo factor, en el caso portugués, coinciden en su interpretación –el turismo–; sin embargo, frente a un factor en el caso portugués (interpretado como poder de compra) hay dos en este estudio –dinamismo económico y estabilidad socioeconómica–. Esto es, se necesitan dos factores para explicar la dimensionalidad del fenómeno en cuestión, dando muestra de mayor complejidad que el fenómeno estudiado en el país vecino. En parte como intento de mitigar la situación anteriormente descrita se construye el ISCE, aunque no es absolutamente comparable con el poder de compra ni por su construcción ni por lo que intenta cuantificar, como ya se comentó en el comienzo del artículo.

Una vez realizada esta comparación con el *Estudo* portugués, es necesario centrarse en lo que es la “capacidad económica” de las comarcas gallegas. Conviene insistir en el hecho de que el ISCE, aunque tiene un objetivo parecido a alguno de los existentes en la literatura económica, se configura por sus características como distinto y, por lo tanto, no comparable en un sentido estricto con los otros. Una de las principales características de la unicidad de este indicador viene dada por el conjunto de variables empleadas que constituyen su base. Dentro de ellas es necesario destacar la importancia de tener información sobre la recaudación del IRPF para cada comarca.

Sobre los resultados obtenidos, éstos pueden ser fruto de muchas interpretaciones tanto en la vertiente de su clasificación, según el ISCE, como en la vertiente del análisis cluster realizado. Incluso el hecho de constituir una novedad en mu-

chos aspectos, como es la utilización de las comarcas como unidad territorial, hace que la extracción de los principales resultados no sea una tarea fácil. Por lo tanto, siendo conscientes de que habría más aspectos que podrían ser comentados, se piensa que los principales puntos a destacar son los que siguen a continuación.

En primer lugar, hay que mencionar el efecto de suavización de los resultados que se produce al emplear como unidad territorial la comarca, siendo esto muy importante en casos como Betanzos, incluido en el cluster de las *Grandes villas*, ya que si fuese considerado solo el propio municipio presentaría un carácter marcadamente diferente a los restantes municipios de esa comarca, así como a las restantes comarcas del citado cluster, carácter que se ve diluido al considerar el conjunto de la comarca.

Si se tiene en cuenta el plano de las dos primeras componentes (gráfico 1), se observa como las comarcas que gozan de mayor ISCE aparecen marcadamente diferenciadas del resto de las comarcas en la medida en que se salen de la nube de puntos.

Con respecto a la clasificación de las comarcas según su ISCE, probablemente el aspecto más destacable sea la presencia de las comarcas de Valdeorras, Eume, O Salnés y las Mariñas Occidental y Oriental dentro del grupo de cabeza. Este hecho se confirma en el análisis cluster al estar incluidos dentro del cluster de las *Grandes ciudades* y en el cluster *Turístico* (clusters compuestos por las comarcas de mayor dinamismo económico). Sin embargo, dadas las características de estas comarcas y, sobre todo, de sus cabeceras, no extraña su clasificación en los primeros lugares y en los citados clusters.

Un aspecto que no llama mucho la atención es la situación de las comarcas de A Paradanta y Xallas como las de menor ISCE, a gran distancia de las primeras. Son éstas unas comarcas en las que ciertamente todas y cada una de las variables asociadas con el ISCE y, concretamente, con su dinamismo (IRPF, empresas y trabajadores, cajeros automáticos, teléfonos e impuestos municipales) presentan los valores más bajos.

También resulta interesante observar la diferente clasificación de las comarcas en función de si se considera el primer factor o bien el segundo. Así, como comarcas con mayor dinamismo económico (es decir, con mayor importancia en el primer factor) aparecen A Coruña, Ourense, Vigo, Lugo y Santiago, que son también las que están clasificadas en los primeros lugares en el listado, según la primera componente en el ACP. Por otro lado, con respecto al segundo factor, las comarcas mejor clasificadas son Terra de Trives, Terra de Caldelas y O Ribeiro, respondiendo a la idea de estabilidad socioeconómica que recoge este segundo factor y no tanto a su dinamismo económico, que es el que recoge el primer factor.

Sobre los clusters resultantes de este estudio, lo primero que es necesario mencionar es la clasificación que se puede hacer de ellos siguiendo un criterio geográfico. Así, con las excepciones de las *Grandes ciudades*, *Barbanza* y el *Turístico*, el resto de los clusters siguen una expansión espacial en torno a focos concretos que

nos llevan a hablar de: *Grandes villas*, *Galicia interior* y *Marinero*. Esta lógica geográfica hace pensar que, de no existir las características particulares de las *Grandes ciudades*, la geografía sería el condicionante principal de esta clasificación.

La presencia de *Barbanza* como comarca única dentro de un cluster es un hecho significativo que da idea de la singularidad de esta área por la influencia tanto de los factores pesqueros y de los turísticos como por el despegue económico que viene experimentando últimamente esta zona. Además, la configuración del cluster *Turístico* sugiere que las tres comarcas que lo forman –la Mariña Occidental, la Mariña Oriental y O Salnés– presentan características distintas de las que tienen en su entorno geográfico, bien por su dinamismo económico o bien por la importancia del turismo en dichas zonas.

Con respecto a los clusters establecidos con criterios geográficos, hay que mencionar que la *Galicia interior* está constituida por comarcas de las provincias del sur de Lugo y de Ourense, siendo en muchos casos comarcas fronterizas y correspondientes al sudeste de la Comunidad, con la clara excepción de Valdeorras. Estas comarcas constituyen un grupo que se diferencia del cluster más amplio de nuestra clasificación (el de las *Grandes villas*) fundamentalmente por la estabilidad socioeconómica.

Por otro lado, las *Grandes villas* se configuran como un cluster más heterogéneo que puede corresponder con la franja sudoeste-nordeste de la Comunidad, exceptuando, lógicamente, a las *Grandes ciudades* que tienen un peso significativamente más alto en el factor del dinamismo económico que las *Grandes villas*. Es importante destacar las diferencias que presentan las comarcas clasificadas en este cluster en relación al número de parados. Si se considerase la tasa de paro en el análisis (lo que crearía un nuevo factor puro con muy poca varianza explicada y, por lo tanto, muy pobre en contenido) este cluster aparecería dividido en dos, que corresponderían al área de influencia de Lugo y a la de Santiago (Iglesias *et al.*, 1999). En relación con el cluster de las *Grandes villas*, y debido a la influencia geográfica de las *Grandes ciudades*, no es descartable pensar que en un futuro se produzca una ruptura por su centro. Más concretamente se podría asistir a una brecha coincidiendo con el foco de Santiago, que se configura como uno de los más pujantes y, por lo tanto, con clara influencia en su entorno.

Sobre el cluster *Marinero* es necesario destacar que sólo cuenta con las comarcas donde el predominio y la importancia de las actividades relacionadas con la pesca son determinantes. Así, ya fueron comentados los casos de O Salnés y las Mariñas Occidental y Oriental, donde la importancia del turismo lleva a hablar del cluster *Turístico*, así como de Barbanza que, por sus características especiales, se configura como un cluster único. Además, otra comarca costera como Bergantiños no cuenta con significación suficiente de la pesca, de manera que pertenece al cluster de las *Grandes villas* y no al *Marinero*.

Es necesario destacar también el hecho de que este estudio cuenta con clusters muy estables como la *Galicia interior* o las *Grandes ciudades*, es decir, clusters que sufren pequeñas o nulas modificaciones tanto ante la introducción de algunas variables nuevas en el análisis como ante el cambio en la ordenación aleatoria de las comarcas al inicio del análisis cluster.

Este análisis de la clasificación y agrupamiento podría no concluir si se hiciese un repaso a la situación particular de cada comarca y a su comparación con otras. Así pues, los aquí reflejados pretenden ser los datos más significativos. Resta, eso sí, dejar constancia de la potencialidad de este trabajo hacia venideros estudios en términos dinámicos, y no estáticos, de la capacidad económica de las comarcas gallegas. De esta manera, una vez sentadas las bases de este indicador, se proyecta ciertamente interesante y atractivo buscar información de un período temporal posterior al aquí empleado para así poder realizar comparaciones temporales y ver cómo se modifican los resultados obtenidos.

BIBLIOGRAFÍA

- ALUJA, T.; MORINEAU, A. (1999): *Aprender de los datos: el análisis de componentes principales. Una aproximación desde el Data Mining*. Barcelona: EUB.
- BANCO CENTRAL HISPANO (1994): *Atlas Comercial de España*. Madrid: Banco Central Hispano.
- BANESTO (1993): *Anuario del Mercado Español*. Madrid: Banco Español de Crédito.
- BATISTA, J.M.; MARTÍNEZ, M.R. (1989): *Análisis multivariante. Análisis de componentes principales*. (Col. ESADE). Barcelona.
- BISQUERRA, R. (1989): *Introducción conceptual al análisis multivariante*. Barcelona: PPU.
- CABRER, B.; MAS, M.; SANCHO, A. (1991): *Necesidades, dotación y déficits en las Comunidades Autónomas*. Valencia: Instituto Valenciano de Investigación Económica.
- CAIXA, LA (1999): *Anuario Comercial de España*. Barcelona: Caja de Ahorros y Pensiones de Barcelona.
- CORREO GALLEGO, EL (1999): *Atlas Económico-Comercial de Galicia*. Santiago de Compostela.
- FERRANDO, P.J.; LORENZO, U. (1998): "Análisis factorial", en J. Renom [coord.]: *Tratamiento informático de datos*, pp. 101-125. Barcelona: Masson.
- GARCÍA SUÁREZ, X. (1998): *Introducción á análise multivariante. Análise factorial*. (Memoria mimeografiada no publicada). Santiago de Compostela.
- HARMAN, H.H. (1980): *Análisis factorial moderno*. Madrid: Saltés.
- IGLESIAS, C. ET AL. (1999): "Capacidade económica das comarcas galegas", *Actas do IV Congreso Galego de Estatística e Investigación Operativa*. Santiago de Compostela.
- INSTITUTO GALEGO DE ESTATÍSTICA (1999): *Base estatística para a obtención de indicadores socioeconómicos das comarcas galegas*. Santiago de Compostela: Xunta de Galicia, Consellería de Economía e Facenda.
- INSTITUTO NACIONAL DE ESTATÍSTICA (1997): *Estudo sobre o poder de compra concelhio*. Coimbra: Direcção Regional do Centro.

- JAUMOTE, CH.; PAELINCK, J.H.P. (1971): "The Differential Economic Structures of the Belgian Provinces: A Time Varying Factor Analysis", *Regional and Urban Economics*, vol. 1, núm. 1, pp. 41-75.
- JOHNSON, R.A.; WICHERN, D.W. (1998): *Applied Mutivariate Statistical Analysis*. New Jersey: Prentice Hall
- JUNTA DE CASTILLA Y LEÓN (1998): *Perspectivas de desarrollo territorial: renta municipal y desarrollo económico en las comarcas de Castilla y León*. Valladolid: Consejería de Economía y Hacienda.
- MALLO, F. (1985): *Análisis de componentes principales y técnicas factoriales relacionadas*. León: Universidad de León.
- MARTÍN GUZMÁN, M.P. (1988): "Métodos estadísticos en el análisis regional", *Estudios Regionales*, núm. 22, pp. 149-170.
- MUÑOZ CIDAD, C. (1992): "La estadística económica en España", en J.L. García Delgado [coord]: *España Economía*, anexo. Madrid: Espasa Calpe.
- PETERS, W.S.; BUTLER, J.K. (1970): "The Construction of Regional Economic Indicators by Principal Components", *Annals of Regional Science*, vol. IV, núm. 1, pp. 1-14.
- SÁNCHEZ CARRIÓN, J.J. (1984): *Introducción a las técnicas de análisis multivariante aplicadas a las ciencias sociales*. Madrid: Centro de Investigaciones Sociológicas.
- SAS INSTITUTE INC. (1989): *SAS/STAT User's Guide: Version 6*, vol. 1, Cary. 4ª ed. NC: SAS Institute Inc.
- SPSS INC. (1999): *SPSS 9.0. Manual del usuario*. Chicago: SPSS Inc.
- SPSS INC. (1999): *SPSS Base 9.0. Applications Guide*. Chicago: SPSS Inc.
- STATA CORP. (1999): *Stata Reference Manual: Release 6.0*. Texas: Stata Press, College Station.
- TEJEIRO, E. (1991): "Algunas técnicas multivariantes útiles para la presentación de resultados de una encuesta", *Estadística Española*, vol. 33, núm. 127, pp. 305-324.