

TECNOLOGÍA: RESULTADOS DE INVESTIGACIÓN

Propuesta para medir los Sistemas Regionales de Innovación en Europa: uso de Variables Sintéticas

Proposal to measure Regional Innovation Systems in Europe: use of Synthetic Variables

Edición Nº 31 – Abril de 2018

Artículo Recibido: Marzo 13 de 2018

Aprobado: Marzo 28 de 2018

AUTOR

Cristián Gutiérrez Rojas

Doctor y Master en Economía, Universidad Complutense de Madrid. Licenciado en Cs. Económicas y de la Administración de Empresas de la Pontificia Universidad Católica de Chile. Profesional cualificado en economía de la innovación y el cambio tecnológico, economía industrial, economía ambiental, macroeconomía y econometría aplicada. Con conocimientos en políticas públicas y trámites legislativos por su experiencia en elaboración y tramitación de leyes en materias pesqueras, acuícolas y ambientales. Experiencia de trabajo con comunidades locales y pueblos originarios en proyectos relacionados a innovación social y conservación del medio ambiente.

Santiago, Chile.

Correo electrónico: cristian.gutierrez.ro@usach.cl

Resumen

Este trabajo presenta una metodología novedosa para configurar los sistemas regionales de innovación (SRI) en Europa. A diferencia de trabajos más tradicionales que utilizan muchas variables individuales altamente correlacionadas, a partir de 35 variables que reflejan distintos aspectos de la actividad innovadora regional, y usando la técnica estadística del análisis factorial, éstas son reducidas a siete variables compuestas o sintéticas, cada una representando aspectos claves o subsistemas de los sistemas de innovación. Esta técnica holística y objetiva, permite a su vez, una serie de aplicaciones que pueden entregar a los responsables de las políticas científicas y tecnológicas nacionales y regionales la posibilidad de visualizar la alta heterogeneidad

de los sistemas regionales en cuanto a la asignación de recursos para las actividades innovadoras, la organización de tales actividades y los productos científicos y tecnológicos resultantes de tales actividades y esfuerzos innovadores.

Palabras clave: Sistemas Regionales de Innovación, Análisis Factorial, Variables Sintéticas

Abstract

This paper presents a novel methodology for configuring regional innovation systems (RIS) in Europe. Unlike more traditional works that use many highly correlated individual variables, from 35 variables that reflect different aspects of the regional innovative activity, and using the statistical technique of factorial analysis, these are reduced to seven compound or synthetic variables, each one representing key aspects or subsystems of innovation systems. This holistic and objective technique, in turn, allows a series of applications that can provide national and regional scientific and technological policy makers the possibility of visualizing the high heterogeneity of regional systems in terms of the allocation of resources for innovative activities, the organization of such activities and the scientific and technological products resulting from such innovative activities and efforts.

Keywords: Regional Innovation Systems, Factor Analysis, Synthetic Variables

1.- Introducción

Medición de los sistemas de innovación: el uso de variables sintéticas.

Sistemas de innovación: conceptos teóricos

El sistema de innovación (SI) es uno de los conceptos que ha visto muy revalorizada su importancia lo que se refleja en las numerosas aportaciones académicas publicadas al respecto. El concepto de SI refleja el proceso de la división del trabajo en el campo de la innovación con la participación correspondiente de un amplio conjunto de agentes e instituciones interrelacionados entre sí, cuyas actividades deberían generar sinergias o ahorrar costes. De hecho, la innovación es una actividad cada vez más compleja e interdisciplinaria y su desarrollo exige la interacción de un elevado número de

instituciones, organismos y empresas. Las actividades de estos agentes del sistema de innovación son en muchas ocasiones complementarias, donde los grandes centros públicos de investigación (incluidas las universidades) se dedican a la investigación básica que, a menudo, no resulta económicamente explotable de forma directa, en tanto que las empresas innovadoras se dedican a desarrollar nuevos productos o procesos mediante la investigación aplicada. En el intermedio existe un amplio conjunto de organismos e instituciones que se ocupan de la transformación de los conocimientos científicos a productos comercializables y en la transferencia, difusión y adaptación de las nuevas tecnologías.

La rápida difusión del concepto de SI tanto a nivel académico como gubernamental se debe a las definiciones relativamente abiertas del mismo concepto, que resultó compatible con toda una serie de corrientes (teóricas) que, a pesar de partir de postulados distintos a los de la economía de la innovación, encajaban en ella y acabaron completando y ampliando el concepto original, entre otros, hacia un ámbito de análisis sub-nacional, no sólo regional, sino también local y sectorial. El concepto del sistema nacional y/o regional de innovación es resultado de la fusión de diversos enfoques teóricos. Se basa, inicialmente, en los conceptos de distrito industrial (Marshall, 1919)ⁱ, polos de crecimiento (Perroux, 1955)ⁱⁱ y en la teoría de los *clusters* (Porter, 1990)ⁱⁱⁱ. Estos enfoques tienen en común la importancia que adjudican a la proximidad espacial, las externalidades, la cultura e identidad regional y el proceso de aprendizaje colectivo o regional (Koschatzky, 2000)^{iv} y, por otro, en los resultados de la teoría del crecimiento que subrayan la importancia de la innovación para tales áreas geográficas.

La parte 'sistémica' del SI se revela debido a que muchos aspectos distintos en diversos ámbitos de la economía y la sociedad en general parecían comportarse de acuerdo a las necesidades de otras partes, como si muchos circuitos de retroalimentación positiva estuvieran operando de forma más o menos sincronizada. La OECD (1994: 4)^v afirma al respecto, que "los resultados innovadores globales de una economía no dependen tanto del desempeño específico llevado a cabo por parte de las instituciones formales (empresas, centros de investigación, universidades, etc.), sino de la forma en la que interactúan entre ellas, como elementos de un sistema colectivo de creación y uso de conocimiento, y del grado de interacción con las infraestructuras sociales (valores, normas

y el marco jurídico)”. El SI es un sistema heterogéneo, dinámico y abierto, caracterizado por la retroalimentación positiva y por la reproducción. Como afirma Lundvall (1992: 2)^{vi}: “con frecuencia, los elementos del sistema de innovación se refuerzan mutuamente en la promoción de procesos de aprendizaje e innovación o, a la inversa, se combinan en grupos, bloqueando dichos procesos. La causalidad acumulativa, y los círculos virtuosos o viciosos, son características de los sistemas y subsistemas de innovación.”

Elementos del sistema de innovación, su interacción y el uso de variables

sintéticas.

En este trabajo se propone dividir el SI en cuatro subsistemas:

- las empresas con sus relaciones inter-empresariales y las estructuras de mercado;
- las actuaciones públicas en relación con la innovación y el desarrollo tecnológico (incluido el marco legal e institucional y la política tecnológica);
- la infraestructura pública y privada de soporte a la innovación;
- el entorno regional.

Las diferencias entre las regiones en cuanto a la configuración de estos elementos son importantes y resultan decisivas para el funcionamiento del sistema regional en su conjunto. En realidad, la frontera entre estos subsistemas es a veces difusa y existe cierto solapamiento entre los distintos ámbitos; por ejemplo, la infraestructura pública de soporte a la innovación forma parte de la política tecnológica. Es decir, no siempre resulta fácil clasificar cada uno de los factores o actores exclusivamente en función de los cuatro subsistemas aquí utilizados; no obstante, tal clasificación —igual que el concepto del sistema nacional y regional de innovación— resulta muy útil como esquema analítico para estudiar un tema tan complejo como la innovación.

Como se ha argumentado en la sección anterior, los sistemas de innovación —y cada uno de sus subsistemas— son realidades complejas en las que participan múltiples agentes y cuya configuración institucional puede ser muy variada. Ello hace que, para la representación de esos sistemas, sea imprescindible recurrir al empleo de múltiples variables (muchas de ellas altamente correlacionadas). De esta manera surge la idea de generar variables sintéticas de carácter abstracto —a las que denominaremos factores—, entendidas como aquellas variables que agrupan a un

importante grupo de variables individuales originales resumiendo la información contenida en ellas y que son identificables con respecto a los elementos que conforman los subsistemas de innovación. Desde un punto de vista conceptual, las variables sintéticas son importantes porque existen dudas si ciertas variables individuales reflejan de forma correcta las características de un sistema de innovación y su potencial. Muchos indicadores individuales reflejan conceptos muy parecidos y pueden ser sustituidos entre ellos. En la literatura tradicional se han optado por variables individuales para medir la innovación, como las estadísticas de patentes (Schmookler, 1950)^{vii} o los gastos en I+D (Griliches, 1984)^{viii}, entre otras. La gran mayoría de estas variables están altamente correlacionadas, pero resulta que cada uno de los indicadores individuales que refleja aparentemente el mismo aspecto del sistema de innovación ofrece un panorama a veces muy distinto. Observando las diferencias en las conclusiones a partir de estos indicadores individuales sobre el nivel tecnológico de una región en particular, reflejado en cada uno de ellos, no cabe duda que su uso por separado marginalizaría la simultaneidad o el carácter holístico del comportamiento innovador. O, como indican Makkonen y Have (2013: 251)^{ix}: “un indicador individual es sólo una indicación parcial del esfuerzo innovador total realizado por un sujeto”, por lo que el uso de indicadores compuestos reflejaría mejor la realidad que cada uno de ellos de forma individual.

Por otro lado, la utilización del análisis factorial para crear indicadores compuestos, resuelven problemas econométricos (como, entre otros, la multicolinealidad y la falta de grados de libertad en los modelos de regresión) y parcialmente los problemas relacionados con errores de medición y procesamiento de los datos (como los errores en el momento de registrar los datos, etc.) o la existencia de ‘*outliers*’ en ciertas variables. Además, evitan en cierto modo el problema que ocasionan las fluctuaciones temporales atípicas de las variables individuales, debido a que cada factor se basa en una media ponderada de diversas variables, lo que amortigua el impacto de los ‘errores’ o valores atípicos de cada variable individual. En el caso que sólo una de las variables de un índice compuesto se haya visto afectada por tal cambio regulatorio o error de medición, su comportamiento atípico se verá parcialmente

corregido o ‘suavizado’ por las demás variables (altamente correlacionadas) incluidas en el indicador sintético del cual forma parte.

A pesar de las ventajas de usar indicadores compuestos existen también críticas respecto a su uso, su utilidad y calidad respecto a su elaboración o confección (Hollenstein, 1996^x; Buesa et al., 2006^{xi}; Makkonen y Have, 2013). El primer problema sería la forma de ponderación o el peso de cada variable individual en el proceso de su agregación al indicador compuesto (Grupp y Schubert, 2010^{xii}; Cerulli y Filipetty, 2012^{xiii}). Estos autores indican que especialmente en el caso de una correlación muy alta entre las variables, el uso de un indicador compuesto no siempre será superior al uso de variables individuales, ya que posiblemente llegarían a conclusiones muy parecidas. Aunque es verdad que en este caso los indicadores compuestos no solucionan un problema real, tampoco sería peor que las variables individuales. Además, aunque no sería superior en la mayoría de los casos, posiblemente corrige el problema de valores atípicos o extremos.

También Makkonen y Have (2013) señalan que las distintas formas de ponderación son una de las desventajas o problemas metodológicos más importantes a los que se afrontan los indicadores compuestos. Las decisiones subjetivas sobre la ponderación o el peso exacto de cada variable, influyen claramente sobre el resultado o puntuación final del indicador sintético compuesto. A pesar de ello, estos autores defienden el uso de indicadores compuestos subrayando que la innovación es un proceso interactivo donde interviene un amplio número de actores y que consiste en diversas fases como la creación de nuevas ideas y conocimientos científicos, el proceso de innovación que convierte las ideas en productos en combinación con las actividades de diseño de los productos y desarrollo de los innovaciones de proceso. También la industrialización, comercialización, transmisión y difusión son parte del proceso de innovación. Durante este proceso interactivo existen flujos continuos —interacciones y retroalimentaciones—, es decir, se trata de un modelo interactivo —no lineal o secuencial— en el que las fases están intensamente interrelacionadas (Kline y Rosenberg, 1986)^{xiv}. Además, existe una amplia heterogeneidad en la *expertise* y competencias de distintos agentes, regiones y sistemas de innovación. Por lo tanto,

sería un error utilizar indicadores (individuales) para medir la capacidad tecnológica o el nivel de eficiencia de sólo una de estas fases.

Los problemas mencionados en esta sección están lejos de resolverse de forma unánime y consensuada. La creación de indicadores compuestos en el campo de los sistemas de innovación es un fenómeno todavía novedoso o de reciente data y se requiere la construcción de un consenso y la estandarización del modelo metodológico para elaborar los índices sintéticos y la ponderación de las variables incluidas en ellos. En palabras de Grupp y Schubert (2010: 68): “Resumiendo un debate que viene prolongándose desde hace décadas, la medición de la ciencia y la tecnología requiere seguir una aproximación multidimensional. Hasta ahora no se ha desarrollado una variable ‘*catch-all*’ que recoja de forma satisfactoria todas las facetas de la ciencia e innovación”.

2.- Metodología. Las características de la base de datos y las variables empleadas.

Las unidades de análisis y la organización subestatal de la Unión Europea

EUROSTAT ha sistematizado una nomenclatura (NUTS) que divide todo el territorio de la Unión Europea en tres niveles regionales y tres niveles locales. En base a este sistema y siguiendo a Baumert (2006)^{xv}, se han escogido 132 regiones de acuerdo a los siguientes criterios:

- a) Que la unidad de análisis estadístico corresponda a un nivel político-administrativo del Estado.
- b) Que los niveles escogidos cuenten con ámbitos competenciales parecidos, sobre todo en materia económica y de I+D.
- c) Que reflejen, por tanto, el grado de descentralización política del país.
- d) Que sólo se combinen niveles de desagregación vecinos (por ejemplo NUTS 1 y NUTS 2, pero no NUTS 1 y NUTS 3).
- e) Que exista suficiente disponibilidad de datos a ese nivel de desagregación.

Dada la falta de mucha información relevante se ha decidido sacar a Grecia y a los países del este europeo del estudio, por lo tanto hablamos de la UE14.

Metodología, base de datos y variables utilizadas

En esta sección especificaremos el conjunto de variables acerca del esfuerzo y de los resultados innovadores de nuestra base de datos, apuntando también sus limitaciones. Posteriormente, en la siguiente sección se especificará la metodología utilizada en este estudio para crear las variables sintéticas o hipotéticas que reflejan las diferentes características del *input* o el esfuerzo de los sistemas regionales de innovación y sus *outputs*.

Medición del esfuerzo o ‘input’ de los sistemas de innovación

El insumo de mayor incidencia de acuerdo a distintos enfoques teóricos es el que representa el esfuerzo innovador que se mide tradicionalmente por el gasto en I+D y el personal en I+D. Por un lado, el gasto en I+D recoge todos los medios financieros destinados a esta actividad, e incluye tanto los gastos corrientes como los de capital y ha sido calculado como tanto por mil sobre el Producto Interior Bruto (en euros constantes de 2010). En este estudio se incorpora la variable *stock* del capital tecnológico, previo cálculo en base a la metodología de Soete y Patel (1985)^{xvi}. Por su parte, el personal en I+D, engloba a todas las personas dedicadas directamente a esta actividad así como a aquellos cuyo rendimiento contribuye de forma inmediata a la misma (directivos, personal administrativo, etc.). Disponemos de dos unidades de medida para el personal en I+D: el número de personas empleadas en términos absolutos (*head count*) y en equivalencia a dedicación plena (*full time equivalent*) calculadas como tanto por mil sobre el número de empleados de la región. Estos indicadores —aunque altamente correlacionados— son complementarios y miden conceptualmente aspectos diferentes.

Variables del contexto socioeconómico de los sistemas regionales de innovación

Como se ha indicado en previamente, la noción de entorno global regional incluye diversos aspectos que de forma indirecta influyen sobre la capacidad tecnológica de la región, como podría ser el sistema educativo, el nivel del capital humano, el sistema financiero (capital riesgo), el grado de exigencia de los demandantes de bienes y servicios, la cultura y el nivel de vida. Así se han introducido diversas variables que reflejan el contexto socioeconómico.

La primera de ellas —que se incluye de forma indirecta— es el tamaño. A la hora de trabajar con regiones muy heterogéneas debemos tomar en consideración el tamaño de las mismas. Respecto al tamaño de las regiones en Europa, nos encontramos con regiones de gran extensión pero muy poco pobladas (como Åland en Suecia) versus regiones pequeñas desde el punto de vista poblacional, pero con un fuerte desarrollo económico (como podría ser Navarra o el País Vasco en España). Por este motivo, es aconsejable corregir las distintas variables por el tamaño poblacional o económico, lo que se ha hecho oportunamente a través del número de habitantes medio anual o el Producto Interior Bruto (PIB). Además, se han incorporado variables que describen la realidad económica de las regiones. Para lo anterior se han agregado variables como el PIB per cápita y la productividad aparente del trabajo.

Otro aspecto importante del entorno sería el nivel de riqueza y la capacidad productiva de la región en términos relativos, la que se ha introducido mediante dos variables. El PIB per cápita que indicaría el nivel de vida y de forma indirecta el nivel tecnológico de la demanda de los consumidores. En el caso de un alto nivel de PIB per cápita, los consumidores exigirían productos de un mayor nivel de calidad y prestaciones, lo que a su vez induciría a las empresas a aumentar su esfuerzo innovador (*demand pull*). Por otro lado, un mayor nivel de vida y salarios altos sirven para atraer a nuevos talentos y a los mejores investigadores y/o inventores. La otra variable, correlacionada de forma directa al PIB per cápita y ligada al nivel innovador de una región o industria es la productividad aparente. Estas magnitudes suelen incrementarse según se acrecienta el nivel tecnológico de la región o de una industria en concreto, siendo mucho mayor en industrias de media y alta tecnología que en las industrias más tradicionales (*technology push*).

Como último aspecto del entorno socioeconómico se incluyó el esfuerzo a innovar por parte de las regiones, y especialmente por parte de sus empresas y la especialización y propensión a patentar a través del empleo en manufacturas de alta y media tecnología como porcentaje del empleo.

Otras variables del entorno con cierta importancia no se han podido incluir por falta de datos estadísticos disponibles públicamente como, la calidad de las universidades, el nivel de cooperación, etc. Por otro lado, algunos aspectos señalados

como importantes para analizar y comparar los sistemas de innovación, pierden relevancia si se comparan regiones de un solo país, ya que estas variables se hacen más homogéneas como es el caso del sistema bancario o la calidad institucional.

Indicadores del capital humano

Otro aspecto muy importante para la innovación es el capital humano. Son los investigadores, tecnólogos, ingenieros y los profesionales dedicados a la I+D —con su talento, experiencia y calidad— los que lideran el proceso de innovación y determinan en gran parte su nivel de éxito y eficiencia. La medición del capital humano no es fácil y los datos son más bien aproximaciones aunque, por otro lado, los indicadores disponibles son generalmente aceptados y se pueden considerar bastante acertados. El personal científico y técnico contribuye igualmente al avance tecnológico a través de su participación en la producción, el control de calidad, la dirección o la educación. La medición de estos recursos humanos dedicados a la ciencia y tecnología es objeto del Manual de Canberra, cuyas recomendaciones y definiciones son las aplicadas por EUROSTAT.

De esta manera, en nuestro modelo factorial las variables que se incorporan adecuadamente son:

- Recursos Humanos en Ciencia y Tecnología (C&T)- Educación (miles de personas), primer criterio.
- Recursos Humanos en C&T- Ocupación (miles de personas), segundo criterio.
- Recursos Humanos en C&T - *Core* (miles de personas), tercer criterio.
- Número de alumnos que están llevando a cabo estudios de postgrado. Esta variable si bien no mide directamente los recursos humanos en ciencia y tecnología, sí permite reflejar el *output* en la formación de investigadores de las universidades. El número de alumnos de tercer ciclo, además del *output* en recursos humanos de las universidades, nos reflejaría la oferta de investigadores altamente cualificados. La variable ha sido medida con respecto a la población para cada región.

Variables del output tecnológico de los procesos de innovación

Las variables utilizadas como *output* fueron las patentes solicitadas ante la Oficina de Patentes Europea (EPO, por sus siglas en inglés). Obviamente las patentes

no son el indicador perfecto del desempeño innovador entre otras razones porque ellas varían enormemente en su valor e importancia (Hu y Mathews, 2005)^{xvii}. Sin embargo, al menos ellas garantizan un mínimo nivel ‘objetivo’ de novedad internacional y hay una significativa, creciente y sofisticada literatura que usa las patentes como una medida común del *output* innovador (Krammer, 2009)^{xviii}. Además de lo anterior, para un análisis que comienza en el año 2000, las patentes representan el mejor indicador disponible, como ha sido repetidamente confirmado por diferentes autores, y un reciente chequeo de compatibilidad entre múltiples indicadores confirma la utilidad de las patentes como medida del *output* de la innovación en el contexto de los estudios de innovación regional (Li, 2009)^{xix}.

La variable de patentes ha sido incorporada al estudio de dos maneras: por un lado, las solicitudes de patentes totales, y por otro, las solicitudes de patentes de alta tecnología a nivel europeo vía EPO a través de las estadísticas de EUROSTAT, ambas corregidas por millón de habitantes y por empleados activos, de tal manera de expresarlas en términos per cápita. EUROSTAT regionaliza los datos de patentes teniendo en cuenta la localización del domicilio del inventor (o grupo de investigación que obtiene el conocimiento patentado) y no el domicilio del propietario de los derechos protegidos por esas patentes. Ello hace que el uso de esta estadística sea el más adecuado para la investigación que aquí se presenta.

Para terminar de configurar el *output* de los sistemas regionales de innovación se ha agregado el crecimiento del PIB y del PIB per cápita como indicadores del desempeño económico de las regiones.

3.- Representación de los sistemas regionales de innovación en Europa: Aplicación del análisis factorial

Los 35 indicadores utilizados, a su vez, pueden ser sintetizados, por medio del análisis factorial de componentes principales, en un número menor de variables sintéticas —a las que denominaremos factores— de carácter abstracto, aunque identificables con respecto a los elementos que conforman el sistema de innovación, que tienen la capacidad de resumir la mayor parte de la información contenida en las variables originales.

En cuanto a su viabilidad, se puede decir que, en el análisis factorial, las variables no se asignan a priori a un factor, sino que es el propio procesamiento estadístico el que las agrupa. En este sentido, un análisis factorial solo es útil si los resultados son interpretables, de manera inequívoca, a partir del marco conceptual que proporciona la teoría. Esta interpretación será posible si de forma simultánea se cumple que:

- Las variables incluidas en un factor pertenecen al mismo componente o subsistema del sistema regional de innovación.
- Las variables pertenecientes a un cierto subsistema se agrupan en un solo factor.
- Se puede asignar a cada factor o variable hipotética no observable un «nombre» que, sin ninguna ambigüedad, exprese claramente un concepto ajustado a la teoría.
- Los test estadísticos y las medidas de adecuación validen el modelo factorial obtenido (véase la siguiente subsección).

Desarrollo y validación estadística del análisis factorial

Respecto a los tests estadísticos y las medidas de adecuación que validen el modelo factorial obtenido, los cuatro aspectos fundamentales que debe cumplir el modelo factorial son los siguientes (Hair et al., 2004)^{xx}:

- La medida de adecuación muestral de Kaiser-Meyer-Olkin (KMO), que se basa en el estudio de los coeficientes de correlación parcial, debe adoptar un valor entre 0,6 y 0,8.
- El test de esfericidad de Barlett, que contrasta la hipótesis nula que identifica la matriz de correlaciones con la matriz identidad, debe rechazar esa hipótesis nula.
- La varianza total explicada por los factores, que refleja el porcentaje de la varianza inicial (anterior al análisis factorial) explicada por los factores, debe ser superior al 75%.
- Las comunalidades, que son las variables encargadas de medir la variabilidad de cada uno de los indicadores reales utilizados que se conserva en los factores, deben estar por encima del 50%.

Por otra parte, interesa que las variables se saturan en los distintos factores de manera que estos puedan interpretarse sencilla y claramente. Esta es la finalidad que

persigue la rotación Varimax, que además maximiza la ortogonalidad de los factores — o minimiza su correlación—, con lo que se evitan los problemas de multicolinealidad cuando se utilizan en la estimación de modelos econométricos.

El modelo factorial resultante de la aplicación de esta técnica multivariante a la batería de los indicadores disponibles para describir los SRI europeos, es el que se refleja en el cuadro 1. Los aspectos estadísticos relevantes que validan ese modelo se señalan a continuación:

- La medida KMO es igual a 0,8.
- Se rechaza la hipótesis nula del test de esfericidad de Barlett con un nivel de confianza del 99%.
- Se conserva un porcentaje del 90,78% de la varianza total de la muestra.
- Todas las comunalidades son superiores al 85%, excepto cinco.

El cuadro 1 muestra el resultado extrayendo siete factores de acuerdo con el método de componentes principales. El método de componentes principales es un procedimiento para la extracción del espacio factorial cuyo objetivo es obtener proyecciones de las nubes de puntos sobre un número de ejes de manera tal que los factores resultantes sean perpendiculares entre sí. Es decir, se trata de pasar de un conjunto de variables correlacionadas entre sí, a un nuevo conjunto de variables, combinaciones lineales de las originales, que estén incorrelacionadas. Las comunalidades (correlación de cada variable con respecto al conjunto de las demás variables que forman ese factor) de las variables son relativamente altas, la mayoría de ellas superiores a 0,75, lo que garantiza la fiabilidad de los resultados e indica el alto grado de conservación de la varianza de las mismas. Consideramos, por tanto, que el modelo con siete factores viene avalado por dos hechos: en primer lugar, resulta de un procesamiento objetivo (el análisis de componentes principales). A esto se suma, como veremos a continuación, que el modelo permite una fácil interpretación (al no estar las variables saturadas más que en un factor), los factores obtenidos encajan con la teoría de los sistemas regionales de innovación, y que el modelo resulta sumamente robusto, además de mantener un elevado porcentaje de la varianza original, como se puede apreciar en el cuadro 1. Este nos muestra la varianza total explicada, en tres apartados: el primero señala los autovalores iniciales, el segundo indica la suma de las

saturaciones al cuadrado de la extracción, y el tercero presenta la suma de las saturaciones al cuadrado tras rotar los factores.

Los autovalores iniciales reflejan el porcentaje de la varianza explicado por cada variable, y es por este valor por el que el sistema se rige al incorporar variables en el modelo. Obviamente, al incluir todas las variables (cada variable sería un factor) se explica el 100% de la varianza, pero con ello no habríamos alcanzado nuestro objetivo de reducir el número de datos con el que trabajamos. El segundo apartado, nos muestra el porcentaje de la varianza explicado por cada uno de los siete factores extraídos con acuerdo a las especificaciones anteriores, así como el porcentaje acumulado, antes de la rotación. Como podemos apreciar, con siete factores el modelo mantiene el 90,78% de la varianza, es decir, que al pasar de 35 variables a siete factores, sólo perdemos menos del 10% de la información.

Sin embargo, a nuestros efectos resultan más interesantes los porcentajes de la varianza explicados por los factores tras la rotación. Como podemos apreciar, el porcentaje de la varianza acumulada por el conjunto de factores, resulta el mismo en uno u otro caso. Sin embargo, lo que se ve alterado es la contribución específica de cada factor al total. La rotación consiste en girar los ejes en el origen hasta alcanzar una determinada posición, de manera que se maximice la carga o saturación de las variables en un factor, minimizándolas simultáneamente en los restantes, permitiendo así una solución más interpretable. Existen distintos procedimientos de rotación —los de rotación ortogonal y los de rotación oblicua— aunque para nuestro caso solo tomamos en consideración los primeros, ya que mantienen un ángulo de 90 grados entre los ejes, garantizando así la ortogonalidad entre los factores. En concreto se ha llevado a cabo una rotación de tipo Varimax, ya que el patrón factorial obtenido por este procedimiento tiende a ser más robusto al obtenido por métodos alternativos.

Cuadro 1 Varianza total explicada

Variable	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	14.955	42.729	42.729	14.955	42.729	42.729	11.857	33.877	33.877
2	6.255	17.871	60.600	6.255	17.871	60.600	4.272	12.206	46.084
3	3.551	10.145	70.745	3.551	10.145	70.745	3.881	11.088	57.172
4	2.468	7.051	77.796	2.468	7.051	77.796	3.817	10.906	68.078

5	1.981	5.661	83.456	1.981	5.661	83.456	3.745	10.700	78.778
6	1.509	4.310	87.767	1.509	4.310	87.767	2.233	6.379	85.158
7	1.053	3.009	90.775	1.053	3.009	90.775	1.966	5.618	90.775
8	.678	1.936	92.712						
9	.528	1.509	94.221						
10	.383	1.095	95.315						
11	.312	.892	96.207						
12	.215	.614	96.820						
13	.184	.527	97.347						
14	.175	.501	97.848						
15	.133	.380	98.228						
16	.113	.322	98.550						
17	.101	.289	98.839						
18	.073	.208	99.048						
19	.067	.191	99.238						
20	.057	.162	99.400						
21	.044	.124	99.525						
22	.042	.120	99.645						
23	.035	.101	99.746						
24	.022	.063	99.809						
25	.020	.056	99.866						
26	.014	.041	99.907						
27	.011	.031	99.938						
28	.007	.019	99.957						
29	.005	.015	99.972						
30	.003	.010	99.982						
31	.003	.007	99.989						
32	.002	.006	99.995						
33	.001	.004	99.998						
34	.001	.002	100.000						
35	4.969E-05	.000	100.000						

Fuente: Elaboración propia. Método de extracción: Análisis de Componentes principales. Uso de software SPSS.

De acuerdo a los resultados la adscripción de cada variable a un único factor resulta ahora inequívoca, pudiendo distinguirse siete factores claramente interpretables, y que equivalen al entorno regional (33,8%), las instituciones de enseñanza superior (12,2%) —que reflejan la generación específica de conocimiento científico y corresponden al subsistema de soporte a la innovación—, el *output* tecnológico a través de las patentes y patentes de alta tecnología (11,08%) —como parte del subsistema empresarial—, a las empresas (10,96%) —que recoge la actividad específica de creación de conocimiento tecnológico—, la administración pública (10,7%) —tanto en su

dimensión de soporte como de acción directa a la innovación—, y el grado de sofisticación de la demanda en sentido tecnológico (6,34%), y el *output* económico a través del crecimiento económico regional (5,62%) como elementos del subsistema del entorno. Los resultados obtenidos a través del análisis factorial coinciden por tanto básicamente con los determinantes apuntados por la teoría. Nuestro modelo queda, por tanto, definido.

En resumen, el modelo factorial que hemos estimado proporciona una representación adecuada de los SRI en Europa (UE14), al cumplirse todos los requisitos estadísticos y conceptuales que son exigibles para ello. Por tanto, se pueden emplear los factores resultantes en ese modelo —expresivos de los recursos, organización e interrelaciones que describen a los sistemas de innovación— para abordar análisis posteriores acerca de cómo se desarrollan las actividades de creación y difusión del conocimiento científico y tecnológico en las regiones europeas.

4.- Conclusiones

En este trabajo se ha presentado una metodología para configurar cuantitativamente los sistemas regionales de innovación en Europa. Esta metodología está basada en la construcción de índices compuestos a partir de la técnica estadística del análisis factorial para así reducir un gran número de variables individuales altamente correlacionadas y que reflejan aspectos parciales de las actividades innovadoras de las regiones, en una pocas variables hipotéticas o factores. El modelo resultante además de satisfacer adecuadamente las exigencias estadísticas, es coherente con la teoría de los sistemas de innovación, representando adecuadamente aspectos o subsistemas centrales del sistema regional. La relevancia de estos aspectos para los responsables de las políticas científicas y tecnológicas es la posibilidad de contar con una herramienta que permita reflejar la alta heterogeneidad de los sistemas regionales en cuanto a la asignación de recursos para las actividades innovadoras, la organización de tales actividades y los productos científicos y tecnológicos resultantes de tales actividades y esfuerzos innovadores. Para el caso de América Latina esta metodología es perfectamente aplicable en la medida que la información estadística disponible respecto a los esfuerzos y resultados de la I+D sea compatible y homogénea por lo

menos a nivel de países (sistemas nacionales de innovación). Esto implica un importante desafío para las oficinas estadísticas oficiales latinoamericanas, pero es imprescindible para mejorar el análisis, diseño y evaluación de las políticas nacionales en I+D+i en el continente.

Referencias Bibliográficas

-
- ⁱ Marshall, A. (1919). *Industry and Trade*. London: McMillan.
- ⁱⁱ Perroux, F. (1955). Note sur la notion de pôle de croissance. *Économie Appliquée*, **7**.
- ⁱⁱⁱ Porter, M. (1990). *The Comparative Advantage of Nations*. London: Free Press and McMillan.
- ^{iv} Koschatzky, K. (2000). The regionalisation of innovation policy in Germany –theoretical Foundations and recent experience. *Arbeitspapiere Unternehmen und Región* Nº.1 2000.
- ^v OECD (1994). *Accessing and Expanding the Science and Technology Base*. Paris.
- ^{vi} Lundvall, B. (1992). *National Systems of Innovation: Towards a Theory of Innovation and Interactive Learning*. 1 edn. London: Pinter Publishers.
- ^{vii} Schmookler, J. (1950). *The Interpretation of Patent Statistics*. *Journal of the Patent Office Society*, **32** (2), pp. 123-146.
- ^{viii} Griliches, Z. Ed. (1984). *R&D, Patents and Productivity*. Chicago, IL: The University of Chicago Press.
- ^{ix} Makkonen, T. y Van Der Have, R. (2013). Benchmarking regional innovative performance: composite measures and direct innovation counts. *Scientometrics*, **94** (1), 247-262.
- ^x Hollenstein, H. (1996). A composite indicator of a firm's innovativeness. An empirical analysis based on survey data for Swiss manufacturing. *Research Policy*, **25** (4), 633-645.
- ^{xi} Buesa, M., Heijs, J., Martínez Pellitero, M. y Baumert, T. (2006). Regional systems of innovation and the knowledge production function: the Spanish case. *Technovation*, **26** (4), 436-472.
- ^{xii} Grupp, H. y Schubert, T. (2010). Review and new evidence on composite innovation indicators for evaluating national performance. *Research policy*, **39** (1), 67–78.
- ^{xiii} Cerulli, G. y Filippetti A. (2012). The complementary nature of technological capabilities: Measurement and robustness issues. *Technological Forecasting & Social Change*, **79**, 875–887.
- ^{xiv} Kline, S. y Rosenberg, N. (1986). An Overview of innovation. In: R. Landau & N. Rosenberg, eds, *The Positive Sum Strategy: Harnessing Technology for Economic Growth*. First edn. Washington, DC: National Academy of Sciences, 275-306.
- ^{xv} Baumert, T. (2006). *Los determinantes de la innovación. Un estudio aplicado sobre las regiones de la Unión Europea*. Tesis Doctoral, Universidad Complutense de Madrid, Madrid.
- ^{xvi} Soete, L. y Patel, P. (1985). Recherche-Développement, importations de technologie et croissance économique. Une tentative de comparaison internationales. *Revue Economique*, **36** (5).
- ^{xvii} Hu, M.C. y Mathews, J.A. (2005). National Innovative Capacity in East Asia. *Research Policy*, **34** (9), 1322-1349.
- ^{xviii} Krammer, S.M.S. (2009). Drivers of National Innovation in Transition: Evidences from a Panel of Eastern European Countries. *Research Policy*, **38**, 845-860.
- ^{xix} Li, X. (2009). China's Regional Innovation Capacity in Transition: An Empirical Approach. *Research Policy*, **38** (2), 338-357.
- ^{xx} Hair, J.F., Anderson, R., Tatham, R. y Black, R. (2004). *Análisis multivariante de datos*. Madrid.