

CAPACIDAD TECNOLÓGICA Y HETEROGENEIDAD DE LOS SISTEMAS REGIONALES DE INNOVACIÓN EN EUROPA

APPLICATION OF THE FACTORIAL ANALYSIS FOR THE MEASUREMENT OF REGIONAL INNOVATION SYSTEMS IN EUROPE

Cristian Gutiérrez Rojas^a · Thomas Baumert^b · Mikel Buesa^c · Joost Heijts^d

Clasificación: empírico – investigación
Recibido: 24 – octubre - 2017 / Aceptado: 29 – junio - 2018

Resumen

Este artículo presenta una metodología para configurar los sistemas regionales de innovación (SRI) en Europa. A partir de la técnica estadística del análisis factorial, 35 variables altamente correlacionadas, pero que reflejan distintos aspectos de la actividad innovadora regional, son reducidas a siete variables compuestas o factores, 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, entre ellas, el cálculo de un índice de capacidad tecnológica, la configuración de una tipología de los SRI y la estimación de una función de producción de conocimiento identificando los determinantes de la innovación.

Palabras clave: sistemas regionales de innovación, análisis factorial, índices compuestos, función de producción de conocimiento.

Abstract

This paper presents a methodology for setting up regional innovation systems (RIS) in Europe. From the statistical technique of factorial analysis, 35 highly correlated variables that reflect different aspects of regional innovation activity are reduced to seven composite variables or factors, each one representing key aspects or subsystems of the innovation systems. This holistic and objective technique, in turn, allows a series of applications, among them, the calculation of an index of technological capacity and the configuration of a typology of RIS and the estimation of a knowledge production function identifying the determinants of innovation.

Keywords: Regional innovation systems, factorial analysis, composite indexes, knowledge production function.

^a Facultad Tecnológica, Universidad de Santiago de Chile, Santiago, Chile. Correo electrónico: cristian.gutierrez.ro@usach.cl.

^b ESIC Business & Marketing School, Madrid, España. Correo electrónico: thomas.baumert@esic.edu.

^c Instituto de Análisis Industrial y Financiero, Universidad Complutense de Madrid, España. Correo electrónico: mbuesa@ccee.ucm.es.

^d Instituto de Análisis Industrial y Financiero, Universidad Complutense de Madrid, España Correo electrónico: joost@ccee.ucm.es.

Los autores agradecen los comentarios y sugerencias de los correctores anónimos, los que mejoraron sustancialmente la versión original del artículo.

Introducción

La innovación es un elemento esencial para el crecimiento económico y el bienestar social de un país o de una región, pues de ella se derivan los impulsos que propician la aparición y difusión de las nuevas actividades de producción, así como los nuevos procesos que posibilitan las combinaciones de factores productivos de manera más eficiente (Aghion, Blundell, Griffith, Howitt y Prantl, 2009; Fagerberg, Mowery y Nelson, 2005; Griliches, 1990; Mansfield, 1986; Schumpeter, 1935). Reconocida la importancia de la innovación, en la sociedad existe la convicción de la pertinencia del apoyo público a la ciencia y la tecnología; y en los gobiernos de la gran mayoría de los países existe la preocupación sobre la necesidad de sostener un cierto abanico de instrumentos de política económica destinados a favorecer las actividades de I+D. En este sentido, el concepto de “sistema de innovación” ha visto muy revalorizada su importancia lo que se refleja en las numerosas aportaciones académicas publicadas al respecto. Tal sistema se puede definir como

“[...] el conjunto de instituciones distintivas que de forma conjunta e individual contribuyen al desarrollo y difusión de nuevas tecnologías y que proveen un marco en el que los gobiernos formulan e implementan políticas con el propósito de influir en el proceso de innovación. Se trata, por tanto, de un sistema de instituciones interconectadas destinadas a crear, guardar y transferir conocimientos, aptitudes y artefactos que definen nuevas tecnologías” (Metcalf, 1995).

En este trabajo aportamos una metodología cuantitativa que permita configurar estos sistemas de innovación a escala regional, manteniendo sus postulados teóricos y que permita, a su vez, una serie de aplicaciones, entre ellas, el cálculo de un índice de capacidad tecnológica, la configuración de una tipología de los sistemas regionales de innovación (SRI) y la estimación de una función de producción de conocimiento identificando los determinantes de la innovación. Lo anterior enmarcado en el enfoque evolucionista y sistémico de la economía de la innovación. Los sistemas de innovación que pertenecen a la familia de modelos que forman la economía evolucionista son sistemas abiertos y heterogéneos, sus elementos coevolucionan y se refuerzan unos con otros promoviendo procesos de aprendizaje e innovación, o inversamente, combinándose en constelaciones que bloquean tales procesos. La causación acumulada así como los círculos virtuosos y viciosos son característicos de los sistemas o subsistemas de innovación (Lundvall, 1992). Todo ello implica que un sistema de innovación es un concepto abstracto difícil de medir de forma directa a

base de variables individuales. De esta manera, en el presente artículo se propone una metodología de configuración de los SRI, profundizando en los aportes pioneros de Buesa, Heijs, Martínez Pellitero y Baumert (2006), Grupp y Schubert (2010) y Makkonen y Van Der Have (2013), respecto al uso de índices compuestos para medir la actividad innovadora.

El artículo se organiza de la siguiente manera. En las secciones segunda y tercera se exponen las raíces conceptuales del concepto de “sistema de innovación”; partiendo por una revisión histórica del tratamiento del cambio tecnológico y la innovación en la literatura económica; y terminando con la explicación acerca del uso de indicadores compuestos o sintéticos como una aproximación holística para medir los sistemas de innovación. Después de la cuarta sección, donde se describen las características de la base de datos y las variables empleadas, se presentan los resultados del análisis factorial y la configuración de los SRI en Europa (sección cinco), para luego presentar tres aplicaciones. El cálculo de un índice de capacidad tecnológica más una tipología de los SRI en Europa en la sección seis, y la estimación de una función de conocimiento en la siete. Finalmente, se presentan las conclusiones y algunas recomendaciones.

Teoría del cambio tecnológico y del sistema de innovación

Hasta la segunda mitad de la década de los setenta, la teoría económica —el denominado “enfoque neoclásico”— consideraba la tecnología básicamente como información, cuyo proceso de producción resultaba de la acción secuencial de las instituciones de investigación —exógenas al sistema económico— y de las empresas innovadoras. Este modelo, denominado “modelo lineal del cambio tecnológico”, fue el fundamento teórico sobre el que se sostuvo, hasta mediados de los años ochenta, la política tecnológica de la mayoría de los países desarrollados. La teoría lineal de la innovación considera que el producto o resultado (*output*) está relacionado de forma lineal con el factor de entrada o insumo (*input*), hecho este que implica una visión de la I+D como actividad aislada, llevada a cabo en centros de investigación, y que resulta ajena a los incentivos del mercado y a los requerimientos de otras unidades de la empresa. La transferencia tecnológica es concebida como un proceso automático, sin costos significativos ni retrasos en el tiempo, basado en el mecanismo de la “mano invisible” smithiana, y en el que la tecnología resultaría una información fácil de copiar (Caniëls, 1996). En consecuencia, el modelo lineal niega de manera virtual factores como la influencia institucional, las estrategias y actitudes competitivas de otras

empresas o naciones, así como los factores relacionados con la demanda, la educación u otros aspectos vinculados al ámbito geográfico en cuestión. Por consiguiente, las políticas basadas en el modelo lineal están dirigidas hacia la generación de innovaciones mediante la creación de centros de investigación, el apoyo a la I+D básica para tecnologías claves o la financiación directa de las actividades de investigación empresariales.

No obstante, el modelo neoclásico básico, adolece de una serie de limitaciones, entre las que destaca su incapacidad de explicar la ausencia de convergencia entre las economías de las distintas naciones —que se derivaría de sus postulados— y que tiene su origen en el modelo de crecimiento de Solow (1956). Este modelo, que estuvo vigente más allá de la década de los cincuenta, consideraba el progreso tecnológico como uno de los factores que permitían explicar el crecimiento económico a largo plazo. Sin embargo, no se estimaba que fuera este un factor que explicara las diferencias existentes entre naciones, dado que la tecnología se concebía como un bien público del que no solo se beneficia el mercado en el que se había generado la innovación, sino que también participaban de ella todos los demás (Baumert, 2006). Por consiguiente, esta visión llevaba a la economía neoclásica a pronosticar una convergencia de la tasa de crecimiento del PIB per cápita de todos los países en el largo plazo. Sin embargo, el hecho de que en la práctica este supuesto no se haya confirmado, indujo una crisis de la teoría neoclásica hacia finales de los años setenta, obligando a una revisión de sus premisas, que dio pie a la aparición de enfoques alternativos.

Así fue como a lo largo de la década de los ochenta los economistas siguieron dos sendas de trabajo. Por un lado, los teóricos neoclásicos revisaron y ampliaron el modelo básico de Solow, incluyendo en él la tecnología como variable endógena. Entre estas aportaciones podemos destacar el modelo de crecimiento endógeno de Romer (1990). Por otro, experimentó un importante desarrollo la teoría evolucionista o neo-schumpeteriana, cuya visión del proceso de innovación se diferencia radicalmente de la escuela neoclásica (Nelson y Winter, 1982).

De acuerdo con Schumpeter¹, es el proceso de “destrucción creativa” el que rige la evolución histórica del capitalismo, pudiendo diferenciarse entre cinco tipos de innovaciones: la introducción de un nuevo bien, la introducción de un nuevo método de producción, la apertura

de un nuevo mercado, la conquista de una nueva fuente de aprovisionamiento de materias primas o bienes semifabricados y la creación de una nueva organización de cualquier industria.

No obstante, el enfoque schumpeteriano no es monolítico, sino que presenta una evolución, por lo que podemos distinguir entre dos tipos de modelos básicos, que resultan complementarios entre sí (Baumert, 2006): el modelo denominado *mark I* corresponde a una visión de la innovación como proceso que tiene lugar en un entorno competitivo de empresarios capitalistas, caracterizado por invenciones —económicamente no determinadas— y descubrimientos científicos exógenos. La actividad innovadora del empresario, consiste en identificar, entre las invenciones y nuevos conocimientos disponibles, aquellos que entrañan un potencial económico, e implementarlos para transformarlos en innovaciones. Al actuar así, deja obsoletas las viejas tecnologías, proceso este que Schumpeter denomina de “destrucción creativa”.

Esta visión inicial de Schumpeter, se complementa con el posterior modelo denominado *mark II*, que se caracteriza por el hecho de que las innovaciones son endógenas y porque la investigación y desarrollo es llevada a cabo principalmente en los departamentos de I+D de las grandes empresas, en un proceso denominado “acumulación creativa”. Este modelo implicaría, pues, el paso de una concepción inicial enfocada hacia el papel del empresario individual, hacia una visión que resalta la importancia de la innovación colectiva desempeñada en el seno de las (grandes) empresas. En palabras de Heijs (2001, p. 29): “Schumpeter reconoció tanto la importancia de las pequeñas empresas e inventores individuales como el importante papel del nivel de concentración para el desarrollo innovador”. En definitiva, Schumpeter allanó el camino para el desarrollo del enfoque evolucionista de la innovación —dando así la posibilidad de soslayar las limitaciones de la teoría neoclásica—, poniendo de relieve la importancia de la destrucción y acumulación creativa, y resaltando el papel del empresario. A su vez, del enfoque evolucionista se deriva el modelo interactivo del cambio tecnológico, desarrollado en la década de los ochenta, y que se opone al modelo lineal imperante hasta entonces.

La aparición del modelo interactivo del cambio tecnológico, ha implicado cambios radicales no solo en la gestión tecnológica de las empresas, sino también en el diseño de las políticas tecnológicas por parte de la administración pública. Este modelo parte de una interacción continua entre los distintos agentes y elementos del mismo a lo largo de todo el proceso de innovación y de la posterior

1 Para una visión amplia sobre la “inflexión schumpeteriana”, véase Vence (1995), pp. 106-143 y para el pos-schumpeterianismo, pp. 144-178. Véanse también los distintos artículos recogidos en Scherer (1984).

comercialización de los productos. Incluso, cuando estos se encuentran plenamente introducidos en el mercado, el proceso prosigue mediante el perfeccionamiento y la diversificación de los productos, de los procesos de producción y de las tecnologías empleadas.

Así, mientras que en el modelo lineal inciden únicamente las actividades tecnológicas del departamento de I+D, en el modelo interactivo destaca la capacidad tecnológica de la empresa en general, considerándose la gestión de la innovación un proceso estratégico y corporativo en el que debe estar implicada toda la empresa, además de sus proveedores, distribuidores y clientes.

De acuerdo con este modelo, la capacidad tecnológica de una empresa se basa en el “saber-hacer” y tiene una dimensión tácita y acumulativa. La transferencia tecnológica resulta difícil y costosa, y la absorción de nuevas tecnologías exige un fuerte gasto en tiempo y en dinero. El modelo interactivo considera la innovación como un proceso dinámico, evolutivo, acumulativo, sistémico e interrelacionado (Vence, 1995), con efectos de realimentación continuos entre las distintas etapas que, además, se desarrolla en un ambiente cambiante (Malerba y Orsenigo, 1995), donde los actores y competidores reaccionan, a su vez, ante cada uno de los cambios.

Cada uno de estos dos modelos refleja un concepto opuesto del bien tecnológico, aunque, en la práctica, la mayoría de los conocimientos se pueden clasificar como una forma mixta de ambos. La aparición del modelo interactivo conllevó el desarrollo del concepto de “sistema nacional de innovación”, planteamiento que, con posterioridad, ha sido aplicado con éxito al plano subnacional o regional (Baumert, 2006).

El enfoque de los “sistemas de innovación”

Origen y desarrollo histórico del término

Una de las aportaciones más fructíferas y que en mayor medida ha marcado el desarrollo de la teoría económica de la innovación, es la que ha centrado su atención en el estudio de los sistemas nacionales y regionales de innovación. El término *sistema (nacional) de innovación* aparece por primera vez en el libro de Freeman² acerca de la tecnología y el desempeño económico en Japón, aunque recibió un importante impulso gracias a la aparición de dos publicaciones casi simultáneas de Lundvall (1992) y Nelson (1993), a las que se les sumó más adelante otra obra de referencia, en esta ocasión de Edquist (1997).

² Freeman (1987).

Sin embargo, más allá de las definiciones estrictas de los sistemas de innovación, a efectos del presente trabajo es relevante analizar aquellos elementos que los diversos autores han señalado como fundamentales y determinantes en la generación de estos sistemas y que han sido considerados en el diseño de modelos y aplicaciones empíricas. Por ejemplo, Lundvall (1992) identifica cinco elementos diferenciadores entre los sistemas: la organización empresarial, las relaciones entre empresas, los gastos en I+D del sector público, la estructura del sector financiero y la organización e intensidad del I+D empresarial. En ese contexto, Nelson (1993) destaca como elementos relevantes las estructuras nacionales, los incentivos a la innovación, la capacidad creativa de los agentes económicos y las singularidades culturales. De la misma forma, Patel y Pavitt (1994) enumeran cuatro elementos que identifican como centrales de la innovación: las empresas, los centros de educación superior e investigación aplicada, y la administración pública. A su vez, Edquist y Jhonson (1997) identifican como componentes relevantes de un sistema de innovación al sector privado, concretamente las empresas, y el sector público en su conjunto (Edquist, 2005). De acuerdo con Edquist (2005), estas organizaciones serían estructuras formales con un objetivo explícito, que han sido creadas conscientemente; es decir, se han establecido con un fin determinado por los agentes o actores del sistema. Finalmente, podemos concluir que el concepto de sistema de innovación puede analizarse desde dos perspectivas: una, más específica y otra, más amplia (Asheim y Gertler, 2005). Una definición estrictamente teórica incluiría a la I+D de las empresas, de las universidades y de la administración pública y reflejaría un modelo lineal *top-down* similar al enfoque de “triple-hélice” (Etzkowitz y Leydesdorff, 2000). En una definición más amplia del término se incluirían todos los elementos y aspectos de la estructura económica y de las redes institucionales que afectan tanto al aprendizaje como a la investigación y a la exploración. Este enfoque resulta más interactivo y reflejaría un modelo *bottom-up* (Baumert, 2006).

Elementos del sistema de innovación y su interacción

Siguiendo con lo expuesto en la subsección anterior, en cuanto a los elementos de los sistemas de innovación, proponemos dividir el sistema de innovación en cuatro subsistemas:

- las empresas con sus relaciones interempresariales 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.

El uso de indicadores compuestos o sintéticos como una aproximación holística para medir los sistemas de innovación

Como se ha argumentado previamente, 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, se debe resumir la información contenida en las variables originales, creando un número menor de variables sintéticas de carácter abstracto —a las que denominaremos “factores”—, aunque 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. Por otro lado, los indicadores compuestos resuelven problemas econométricos (como, entre otros, la multicolinealidad y la falta de grados de libertad en los modelos de regresión) o metodológicos (suavizan la existencia de *outliers* o errores en las estadísticas). En esta subsección se debate primero las razones conceptuales, seguidas por las ventajas metodológicas y, al final, se discuten las fortalezas y debilidades de las variables compuestas.

La teoría evolucionista de la economía de la innovación subraya la heterogeneidad del comportamiento innovador como una actividad multidimensional que está

afectada directamente por su entorno económico y social, donde participa un gran número de agentes, instituciones y factores que interactúan en un marco sistémico basado en un gran número de relaciones interdependientes, que a su vez dificultan el establecimiento de relaciones causales unidireccionales. Además, un gran número de aspectos del entorno —no directamente relacionados con la I+D— tiene un impacto directo sobre las actividades innovadoras. Todo ello implica que un sistema de innovación es un concepto abstracto difícil de medir de forma directa a base de variables individuales. De hecho, muchos indicadores individuales reflejan conceptos muy parecidos y pueden ser sustituidos entre ellos. 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 de que su uso por separado marginalizaría la simultaneidad o el carácter holístico del comportamiento innovador. O, como indican Makkonen y Van Der Have (2013, p. 251): “[...] 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, resuelve 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 solo 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 en cuanto a su elaboración o confección (Buesa *et al.*, 2006; Hollenstein, 1996; Makkonen y Van Der Have, 2013; OECD, 2008). El primer problema sería la forma

de ponderación o el peso de cada variable individual en el proceso de su agregación al indicador compuesto (Cerulli y Filipetty, 2012; Grupp y Schubert, 2010). 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 tal vez 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 Van Der 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 las innovaciones de proceso. También la industrialización, comercialización, transmisión y difusión forman 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). Además, existe una amplia heterogeneidad en la *expertise* y competencias de distintos agentes, regiones y sistemas de innovación. Por tanto, sería un error utilizar indicadores (individuales) para medir la capacidad tecnológica de solo 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 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, p. 68):

Resumiendo un debate que viene prolongándose desde hace décadas, la medición de la ciencia y la tecnolo-

gía 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.

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. Este sistema desagrega el territorio de la Unión Europea (UE) en 98 regiones NUTS 1, 273 regiones del nivel NUTS 2 y 1.324 regiones NUTS 3. Con base en este sistema y siguiendo a Baumert (2006), los criterios utilizados para la selección de las unidades geográficas son:

- 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 solo 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.

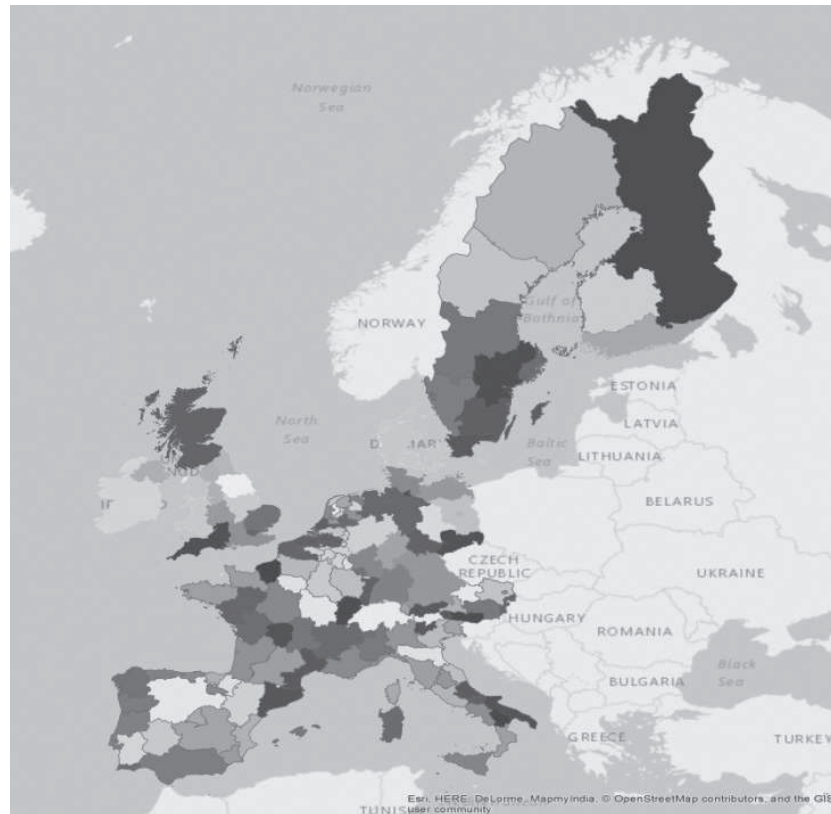
Dada la falta de mucha información relevante se ha decidido sacar a Grecia y a los países del este europeo del estudio, por tanto hablamos de la UE14³. Las 132 unidades geográficas finalmente elegidas se presentan en la figura 1.

Metodología, base de datos y variables utilizadas

En esta subsección 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 subsección 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*.

3 Alemania, Austria, Bélgica, Dinamarca, España, Finlandia, Francia, Irlanda, Italia, Luxemburgo, Países Bajos, Portugal, Suecia y Reino Unido.

Figura 1. Sistemas regionales de innovación en Europa



Fuente: elaboración propia.

Variables del “input” de los procesos de innovación

Medición del esfuerzo de los sistemas de innovación

El insumo de mayor incidencia de acuerdo con 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 bruto⁴. Cabe mencionar aquí, que en gran parte de la literatura acerca de las funciones de producción de conocimiento se ha distinguido entre el uso de los gastos en I+D y el *stock* de I+D, o *stock* de capital tecnológico. Desde un punto de vista teórico, esta última variable es preferible, ya que abarca el *stock* de conocimiento disponible en una economía. En este estudio se incorpora esta variable *stock*, previo cálculo con base en la metodología de Soete y Patel (1985). 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 indicó en la sección teórica, 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.

⁴ Todas las variables monetarias han sido deflactadas y expresadas en euros del 2010.

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 estas. Respecto al tamaño de las regiones en Europa, nos encontramos con regiones de gran extensión pero muy poco pobladas (como Åland en Suecia) frente a las 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 como los salarios, la formación bruta de capital fijo, los empleados y el valor añadido bruto.

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, como se indicó en la sección teórica (entorno regional). 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 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, 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 total.

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 e ingenieros —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. Como afirma la OECD en el *Manual de Frascati*, el personal en I+D no resulta suficiente para medir el rendimiento tecnológico de una región, pues solo representa una parte del *input* humano de un sistema de innovación. El personal científico y técnico contribuye igualmente al avance tecnológico con 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).
- Recursos Humanos en C&T - Ocupación (miles de personas).
- Recursos Humanos en C&T - Core (miles de personas).
- Número de alumnos de tercer ciclo, es decir, de aquellos que están llevando a cabo estudios de posgrado. 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 consideradas 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). 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). Además de lo anterior, para un análisis que comienza en el 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).

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 mediante 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.

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.

Resultados del análisis factorial: las variables sintéticas o abstractas que caracterizan los sistemas regionales de innovación

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 test 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, Anderson, Tatham y Black, 2004):

- 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 comunales, que son las variables encargadas de medir la variabilidad de cada uno de los indicadores reales utilizados que se conservan 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 la tabla 1. La solución adoptada incluye siete factores cuya denominación y participación en la varianza explicada por el modelo se han representado en la figura 2. 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 90,77% de la varianza total de la muestra.
- Todas las comunales son superiores al 85%, excepto cinco.

Las tablas 1 y 2 muestran los resultados 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 comunales (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). En segundo lugar —como veremos a continuación— 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 el modelo resulta sumamente robusto, además de mantener un elevado porcentaje de la varianza original, como se puede apreciar en la tabla 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. El segundo apartado nos muestra el porcen-

taje de la varianza explicado por cada uno de los siete factores extraídos con acuerdo con las especificaciones anteriores, así como el porcentaje acumulado, antes de la rotación. Como podemos apreciar, con siete factores el modelo mantiene el 90,77% de la varianza; es decir, que al pasar de 35 variables a siete factores, solo 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. No obstante, 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, hemos 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.

Como podemos apreciar en la tabla 2, la adscripción de cada variable a un único factor resulta ahora inequívoca, pudiendo distinguirse siete factores claramente interpretables, y que equivalen al entorno regional, las instituciones de enseñanza superior (universidad) —que reflejan la generación específica de conocimiento científico—, el *output* tecnológico a través de las patentes y patentes de alta tecnología, a las empresas (innovadoras) —que recogen la actividad específica de creación de conocimiento tecnológico—, la administración pública, el grado de sofisticación de la demanda (en sentido tecnológico), y el *output* económico mediante el crecimiento económico regional. Los resultados obtenidos con el análisis factorial coinciden, por tanto, básicamente con los determinantes apuntados por la teoría. Nuestro modelo queda, por consiguiente, 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

Tabla 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						

Método de extracción: análisis de componentes principales.

Tabla 2. Matriz de componentes rotados
Matriz de componentes rotados^a

	Componentes						
	1	2	3	4	5	6	7
Remuneraciones de asalariados (millones € 2010)	,979						
Valor agregado bruto (millones € 2010)	,977						
PIB (millones € 2010)	,977						
Número de personas empleadas (miles)	,974						
Recursos Humanos en I+D - empleados (miles de personas)	,969						
Recursos Humanos en I+D – núcleo (miles de personas)	,963						
Población promedio anual (miles)	,962						
Recursos Humanos en I+D - educación (miles de personas)	,951						
Formación bruta de capital fijo (millones € 2010)	,945						
Personal total en I+D (Nº)	,900						
Gasto total en I+D (millones € 2010)	,860						
Personal en I+D empresas Full Time Equivalent (FTE) (Nº)	,849						
Gasto total empresas en I+D (millones € 2010)	,816						
Personal en I+D universidades Head Count (HC) (Nº) % empleo total		,911					
Personal en I+D universidades FTE (Nº) % empleo total		,896					
Gasto universidades en I+D (% PIB)		,853					
Estudiantes tercer nivel regional (% población)		,840					
Stock capital tecnológico universidades per cápita (€ 2010)		,812					
Patentes alta tecnología per cápita (millón de personas)			,922				
Patentes alta tecnología per cápita (millón de activos)			,918				
Patentes per cápita (millón de personas)			,773				
Patentes per cápita (millón de activos)			,756				
Gasto empresas en I+D (% PIB)				,761			
Personal en I+D empresas FTE (Nº) % empleo total				,754			
Personal en I+D empresas HC (Nº) % empleo total				,745			

(Continúa)

Tabla 2. Matriz de componentes rotados *(continuación)*
Matriz de componentes rotados^a

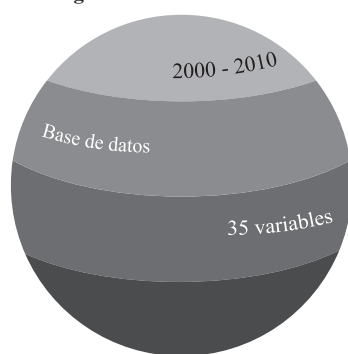
	Componentes						
	1	2	3	4	5	6	7
Empleo regional en manufacturas de alta y media tecnología (% empleo)				,674			
Stock capital tecnológico empresas per cápita (€ 2010)				,672			
Personal en I+D AAPP FTE (Nº) % empleo total					,948		
Personal en I+D AAPP HC (Nº) % empleo total					,927		
Gasto AAPP en I+D (% PIB)					,916		
Stock capital tecnológico AAPP per cápita (€ 2010)					,905		
PIB por trabajador (€ 2010)						,907	
PIB per cápita (€ 2010)						,889	
Tasa de crecimiento PIB per cápita							,988
Tasa de crecimiento PIB							,984

Método de extracción: Análisis de componentes principales.

Método de rotación: Normalización Varimax con Kaiser.

^a La rotación ha convergido en 6 iteraciones.

Figura 2. El modelo factorial



- 1.- Entorno regional (33,88%)
- 2.- Universidades (12,20%)
- 3.- Output tecnológico (11,08%)
- 4.- Empresas innovadoras (10,96%)
- 5.- Administraciones públicas (10,70%)
- 6.- Sofisticación de la demanda (6,34%)
- 7.- Output económico (5,62%)

Fuente: elaboración propia (en paréntesis, el porcentaje de la varianza inicial contenida en el factor).

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.

Índice de capacidad tecnológica y tipología de los SRI en Europa

El índice de capacidad tecnológica busca reflejar cuantitativamente los elementos que pueden mejorar los resultados de los sistemas de innovación, además de permitir realizar comparaciones entre ellos. A partir de la base de datos regionales inicial y del análisis factorial realizado, se logró identificar los factores que componen los sistemas regionales de innovación. Luego se procedió a calcular las ponderaciones (o peso) de los factores y variables, estandarizarlos o normalizarlos para, finalmente, calcular el índice, manteniendo la objetividad estadística. Así, además, se recoge el aporte real de cada variable al sistema de innovación. Se considera que cada factor es un subíndice dentro del índice global y su peso en este es determinado por la variabilidad total explicada por el factor respecto a la varianza total explicada por el modelo factorial. Esto implica que las variables —y, por ende, los factores con mayor variabilidad— tendrán una mayor influencia o peso que aquellas variables que reflejen una

distribución más homogénea entre las regiones (véase Buesa *et al.*, 2006).

Respecto a las variables, su peso dentro de cada factor o subíndice es calculado a partir de la matriz de coeficientes para calcular las puntuaciones factoriales en los componentes. Considerando que cada variable es asignada a solo un factor sobre la base de su grado de correlación con este, el peso relativo es calculado como porcentaje desde las correlaciones entre el factor y cada variable, y la correlación del factor con todas las variables. De acuerdo con los resultados, el factor con mayor ponderación corresponde al “entorno regional” con 37,3% (con 13 variables), seguido por el factor “universidades” con 13,4% (con 5 variables). Más atrás se ubican los factores “productos tecnológicos”, “empresas innovadoras” y “Administraciones Públicas (AAPP)” con 12,2% (4 variables), 12% (5 variables) y 11,7% (5 variables), respectivamente, para quedar los factores “sofisticación de la demanda” y “crecimiento económico regional” (*output* económico) en 7,02% y 6,18%, respectivamente (con 2 variables cada uno).

Finalmente, en la última etapa de construcción del índice, las variables son estandarizadas en orden a oscilar dentro de márgenes establecidos y hacerlas comparables. Para esto se tomaron los valores máximos y mínimos de cada año obteniendo valores estandarizados entre 0 y 100

para la variable correspondiente a la región y al año⁵. La suma de las variables estandarizadas por las ponderaciones dentro de cada factor y multiplicado por cien da los valores para cada factor o subíndice, y estos multiplicados por sus respectivos pesos resultan en los valores del índice para cada región en cada año. A partir del promedio del índice para el período de años de la muestra, se generó el análisis clúster que determina la formación de tres clústeres compuestos por el mismo número de regiones. Para realizar lo anterior, se ha empleado el método

$$5 \quad x_{r,j}^* = \frac{x_{r,j} - x_j^{\min}}{x_j^{\max} - x_j^{\min}} * 100$$

Dónde: $x_{r,j}^*$: valor estandarizado de la región r , en el año j .

$x_{r,j}$: valor observado de la región r , en el año j .

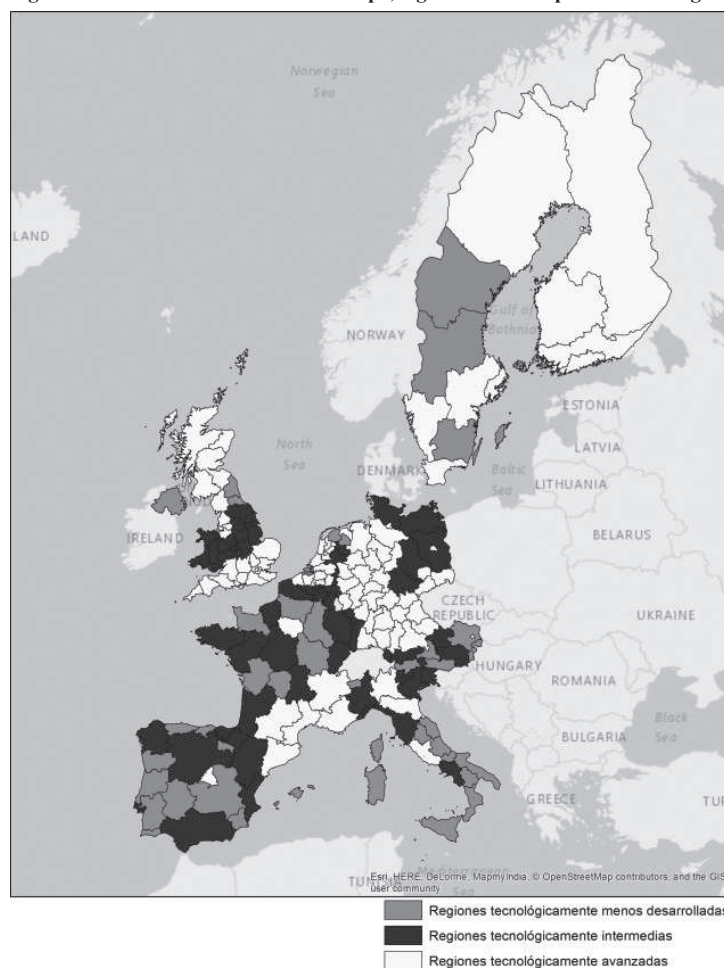
x_j^{\max} : valor máximo observado en el año j .

x_j^{\min} : valor mínimo observado en el año j .

jerárquico aglomerativo “vecino más lejano”, el cual se basa en unir objetos en función de la llamada distancia máxima. La distancia máxima entre individuos de cada clúster se define como “la esfera más reducida (diámetro mínimo) que puede incluir todos los objetos en ambos clústeres”. De esta forma, dos objetos separados por la distancia máxima más corta se colocan en el primer clúster. Luego, se busca nuevamente la distancia máxima más pequeña, uniéndose al clúster existente un nuevo caso, o bien formándose un nuevo clúster. Este procedimiento se repite hasta alcanzar un único conglomerado final.

El resultado deja tres clústeres. El clúster 1 incluye 43 regiones principalmente del sur de Europa con bajos PIB per cápita. El clúster 2 incorpora 43 regiones de desarrollo económico medio y el clúster 3 agrupa a 43 regiones

Figura 3. Sistemas de innovación en Europa, según índice de capacidad tecnológica



Fuente: elaboración propia.

ricas del oeste y centro de Europa y que poseen además sistemas de innovación bien desarrollados⁶.

Se realizó el test de Kruskal-Wallis para corroborar la conformación de los tres clústeres, y se rechazó la hipótesis nula de igualdad de medias entre los tres grupos al 99% de confianza.

Finalmente, con base en las puntuaciones factoriales obtenidas y los clústeres identificados, es posible hacer una breve tipología de los SRI. En la figura 4 se observan las puntuaciones promedio para los clústeres configurados. En esta se aprecia que las regiones más atrasadas son muy fuertes en el factor “crecimiento económico regional” y muy débiles en los factores “entorno”, “universidades” y “empresas”; las regiones intermedias son fuertes en los factores “empresas” y “crecimiento económico regional”, siendo débiles en el factor “sofisticación de la demanda”; mientras que las regiones avanzadas son fuertes en los factores “entorno”, “universidades” y “AAPP”, y débiles en “crecimiento económico regional”. Lo anterior da cuenta de la alta heterogeneidad de los SRI en Europa y la relevancia de aplicar políticas a la medida de cada una de ellas.

6 Se han sacado del análisis las regiones NUTS 1: Irlanda, Luxemburgo y Dinamarca para homogeneizar en NUTS 2 la tipología de los sistemas regionales de innovación en Europa.

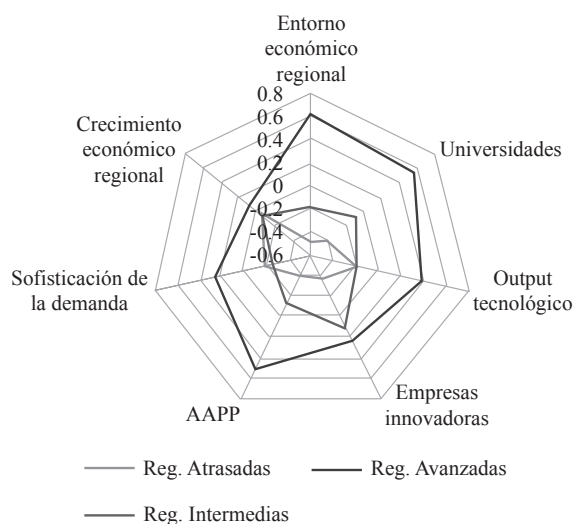
Estimación de una función de producción de conocimiento

Tal y como hemos expuesto en la Introducción, finalmente nos proponemos identificar los determinantes de la innovación regional en Europa y su grado de incidencia sobre el resultado tecnológico de las mismas. Recurrimos para ello a una modelización basada en la función de generación de conocimientos, desarrollada inicialmente por Griliches en 1979, y que desde entonces ha sido aplicada con éxito a toda una serie de trabajos empíricos, tanto en el plano nacional como regional, siguiendo la metodología propuesta por Baumert (2006).

Una revisión de estos estudios, muestra mejores ajustes al trabajar con unidades de análisis agregadas (naciones, regiones o sectores), que al emplear microdatos empresariales (Audretsch, 1998). No obstante, como afirma Fritsch (2002):

[...] la función de generación de conocimientos es una herramienta potente para comparar la calidad de los sistemas regionales de la innovación. [...] Sin embargo, los procesos de innovación son relativamente complejos y no pueden ser medidos de forma comprensiva a través de un único indicador. Es por ello por lo que no se debería recurrir a una única variable, sino a un conjunto de medidas a la hora de comparar la actividad innovadora entre regiones.

Figura 4. Tipología de los sistemas regionales de innovación en Europa



Fuente: elaboración propia.

Por tanto, en el presente estudio trabajaremos con los datos agregados —regionales— y con el número amplio de variables explicativas que conformaron nuestro modelo factorial.

Formulación del modelo

La mayoría de los trabajos empíricos que estudian los determinantes de la innovación tienen su cimiento teórico en la “clásica” función de generación de conocimientos ideada por Griliches (1979), en la que el flujo de nuevas ideas depende del esfuerzo innovador llevado a cabo por una región (u otra unidad geográfica de análisis), tal que:

$$K = f(R) \quad (1)$$

Donde:

K = nuevos conocimientos económicamente valorables.

R = recursos destinados a la investigación.

Este planteamiento concuerda básicamente con el modelo de crecimiento tecnológico endógeno formulado por Romer (1990), en el que la generación de nuevos conocimientos no solo depende del esfuerzo en I+D desempeñado, sino también del stock de conocimientos acumulados, es decir, de la disponibilidad de resultados de anteriores investigaciones. De ello se deriva una función de tipo:

$$A_t = \delta H_{A,t}^{\lambda} A_t^{\varphi} \quad (2)$$

Donde:

A_t = flujo de nuevos conocimientos.

$H_{A,t}^{\lambda}$ = personal dedicado a la I+D.

A_t^{φ} = stock de conocimientos acumulados (a largo plazo λ y $\varphi = 0$).

Un desarrollo más complejo de esta función se encuentra en el modelo ideado por Porter y Stern (1999). De acuerdo con estos autores, la generación de conocimientos resulta de combinar el esfuerzo innovador llevado a cabo con los stocks de conocimientos acumulados nacionales e internacionales, tal que:

$$A_{j,t} = \delta H_{A,t}^{\gamma} A_j^{\varphi} A_{-j}^{\psi} \quad (3)$$

Donde:

$A_{j,t}$ = producción de ideas (patentes).

$H_{A,t}^{\gamma}$ = esfuerzo innovador (recursos humanos dedicados a la innovación).

A_j^{φ} = stock de conocimientos propios del país j .

A_{-j}^{ψ} = stock de conocimientos descubiertos en países distintos al país j y aún no difundidos en este.

Generalizando las funciones (2) y (3) podemos concluir, de acuerdo con el trabajo de Griliches (1990), que el flujo de nuevos conocimientos depende, por una parte, del esfuerzo innovador desempeñado y, por otra, de un conjunto de características propias de la región, que se englobarían en un vector Z_r , de manera que:

$$K_r = f(R_r, Z_r) \quad (4)$$

Donde Z_r puede ser sustituido directamente por una combinación lineal de los indicadores regionales oportunos.

Ahora bien, la pregunta acerca de cuáles son, en concreto, estos indicadores, no ha recibido una respuesta unánime por los diferentes autores que se han enfrentado a ella. Al contrario, las distintas “escuelas” y enfoques teóricos han enfatizado aspectos diversos —hasta el punto de poder ver en ello parte de su “hecho diferencial”—, lo que no obsta para que las distintas aportaciones resulten complementarias entre sí (Grupp, 1997).

En este sentido, el estudio de Furman, Porter y Stern (2002) acerca de los determinantes de la capacidad innovadora nacional, implica un avance notable, al desarrollar un modelo que aúna las contribuciones de distintas escuelas y autores. También en el presente trabajo partimos de un enfoque hasta cierto punto sincrético, tratando de conjugar las distintas perspectivas teóricas para formar el vector Z_r .

Nuestro objetivo consiste, por tanto, en detectar los factores determinantes de la innovación y su grado de incidencia, partiendo —de acuerdo con los postulados teóricos— de la hipótesis, de que todos los elementos del sistema de innovación deberán influir de manera positiva en los resultados del mismo, si bien con intensidades distintas (Baumert, 2006).

Para ello, utilizaremos la metodología de configuración de los SRI expuestos en la primera parte, pero considerando las puntuaciones factoriales obtenidas de un nuevo análisis factorial, esta vez solo incluyendo los elementos de esfuerzo y entorno del sistema, descartando los factores *output*⁷. El *output* tecnológico (patentes totales) será una de nuestras variables dependientes en la regresión de la función de producción de conocimientos. Además, se configuró otra función considerando las

7 La varianza total explicada por este nuevo modelo con cinco factores es del 87,29%, cumpliéndose además todos los requisitos estadísticos y conceptuales.

publicaciones científicas (totales) como variable dependiente, reflejando así el output científico de los SRI.

La función de producción de conocimiento (FPC) con variables regionales combinadas

En esta etapa de nuestro análisis usamos las variables “sintéticas” calculadas previamente para estimar una función de producción de conocimiento por datos de panel. Proponemos un modelo aditivo, siendo común en este tipo de estudios (véanse Acs, Audretsch y Feldman, 1992; Anselin, Varga y Acs, 1997; Feldman, 1994; Furman et al., 2002; Jaffe, 1989), de acuerdo con la siguiente especificación (Baumert, 2006):

$$K_{it} = \beta_0 + \beta_1 RENV_{it} + \beta_2 FIR_{it} + \beta_3 UNI_{it} + \beta_4 ADM_{it} + \beta_5 DEM_{it} + \varepsilon_{it} + \mu_{it} + \nu_t$$

La variable *output* hace referencia al *nuevo conocimiento económicamente valorable* tanto en términos tecnológicos (K_{it} = número de patentes totales) como científicos (K_{it} = número de publicaciones científicas totales), en tanto que las variables explicativas son los cinco factores de esfuerzo y entorno previamente calculados: entorno (RENV), empresas (FIR), universidades (UNI), AAPP (ADM) y sofisticación de la demanda (DEM)⁸.

Algunos puntos deben ser clarificados respecto a esta ecuación. Primero, en la cuarta sección se justificó el uso de las patentes como *output* del sistema. Por otro lado, para subsanar el problema de estudios previos respecto al desconocimiento de que una parte relevante del *output* lo componen las actividades de investigación científica, en este estudio se han incorporado las estadísticas de publicaciones en revistas académicas. La literatura también reconoce ciertos problemas asociados al uso de publicaciones como variable del *output*. Por un lado, está el sesgo del lenguaje, en el sentido de que gran parte de las publicaciones en las revistas científicas más prestigiosas se publican en el idioma inglés, generándose así un sesgo hacia los investigadores cuya lengua nativa sea esta. Otra crítica es que muchas publicaciones son escritas por múltiples autores, muchas veces desde regiones o países diferentes, siendo casi imposible distinguir la contribución individual a la publicación. Sin embargo, ambos problemas pierden fuerza a nivel regional dentro de un país, ya que el sesgo idiomático afecta por igual a todas las regiones, lo mismo que el problema de la coau-

toría⁹. Los datos utilizados para el análisis de la producción científica de las universidades y regiones europeas se han obtenido a partir de una revisión exhaustiva de los artículos de investigación universitarios publicados en revistas indexadas en el *Science Citation Index Expanded* (SCI) en los campos de Ciencias y Tecnología, Ciencias de la Salud e Ingeniería (se han excluido deliberadamente los campos de Ciencias Sociales y de Humanidades para complementar las publicaciones científicas a las patentes y limitar el campo de estudio). Los datos cubren el período 1998-2010¹⁰. El SCI forma parte del *Web of Science* (WoS), que es una base de datos bibliográfica creada por Thomson Reuters. La principal ventaja de WoS es que permite acceder a una lista completa de todos los autores y su institución de afiliación, además de ser considerada por la comunidad científica como la base de datos con la relación de revistas de mayor impacto y calidad en cada campo. Como es lógico, también hay algunas limitaciones para el uso de esta fuente de información; por ejemplo, no incluye todas las revistas y la lista está fuertemente sesgada por las publicaciones en inglés (para un mayor detalle, véanse Bordons, Fernández y Gómez, 2002; Van Raan, 2005; Weingart, 2005).

Otro asunto importante tiene relación con el retardo temporal entre el esfuerzo en I+D y el momento de la solicitud de patentes o la publicación académica. Estudios empíricos parecen demostrar que esta relación es casi contemporánea, al menos por lo que se refiere a las patentes (Hall, Griliches y Hausman, 1986; OECD, 2004; Schmoch, 1999), aunque no cabe decir lo mismo de las publicaciones en revistas científicas. Por otro lado, entre las variables elegidas se encuentra el capital tecnológico, la que de acuerdo con su metodología de cálculo incorpora los gastos en I+D con retardos y las depreciaciones del stock, es decir, implícitamente se usa una estructura de retardos. Finalmente, el análisis factorial suaviza la serie temporal de datos, por lo que las posibles divergencias en los valores de un año a otro de una variable se reducen, haciendo poco relevante la distinción de retardos. De esta manera, los modelos presentados en este estudio solo asumen retardos implícitos entre las variables independientes y los distintos *outputs*¹¹.

8 Los componentes de error global, temporal y específico individual son:

ε_{it} = término de error global.

μ_i = componente de error específico al individuo invariante al tiempo.

ν_t = componente de error específico al tiempo invariante al individuo.

9 Cuando una publicación está asignada a múltiples regiones (los firmantes pertenecen a universidades de diferentes regiones) se ha aplicado el método de conteo completo, siguiendo lo ya realizado por otros autores (Winkler, 2014).

10 Esta base de datos la desarrollaron los profesores M. Acosta y D. Coronado de la Universidad de Cádiz, quienes amablemente la han compartido con nosotros.

11 A modo de comprobación se han calculado modelos con uno y dos retardos, obteniendo resultados muy similares.

Resultados econométricos

Es importante señalar que los resultados aquí presentados tienen como objetivo detectar el peso relativo de los factores determinantes de la innovación y el conocimiento a través de una función de “explicación”, más que obtener una estimación de un *output* futuro, como lo haría una función de “predicción”. Esto tiene importantes consecuencias metodológicas. Además de no hacer necesaria una estructura de retardos o desfases entre insumos y productos, una predicción hace menos aconsejable combinar las técnicas de regresión con otras técnicas estadísticas como el análisis factorial, ya que los coeficientes de regresión (no estandarizados) así obtenidos no harían referencia a la elasticidad de una variable concreta, sino que reflejan la elasticidad de la puntuación factorial, que depende del cambio experimentado por todas las demás variables incluidas en el factor. Además, al trabajar con realidades regionales tan heterogéneas, hace que los errores sean mayores y que la varianza no se distribuya uniformemente a lo largo del plano de regresión, supuestos fundamentales a la hora de estimar regresiones con fines predictivos. Lo anterior no obsta que se hayan optimizado al máximo los modelos aplicando las transformaciones pertinentes (errores robustecidos, test de estacionariedad para panel, test de Hausman, etc.). Sin embargo, por lo señalado anteriormente, las variables dependientes no mantienen una relación lineal con respecto a los factores, lo que se refleja en ajustes subóptimos y en una frecuente ausencia de normalidad en los residuos. Para subsanar estos inconvenientes, realizamos pruebas con distintas transformaciones de las variables dependientes obteniendo los mejores resultados con el tipo raíz cuadrada¹². Esta transformación es la que se aplica a todos los modelos.

Los resultados se presentan en la tabla 3. De acuerdo con los resultados del test de Hausman, se prefiere el modelo de efectos fijos corregido por heterocedasticidad (en negrita)¹³. Los ajustes globales son aceptables con un R^2 de 78,8% en el modelo con patentes y 73,8% en el modelo con publicaciones. En el modelo tecnológico (*output* patentes), todas las variables afectan de manera positiva la producción tecnológica y son significativas estadísticamente a excepción de las universidades. En cambio, en el modelo científico (*output* publicaciones), todos los factores afectan de forma positiva al *output* y todos son significativos estadísticamente (sin incurrir en

colinealidad), lo cual destaca la utilidad de usar factores como regresores. En los dos modelos, el factor “entorno” es el de mayor importancia, resaltando la relevancia que tendría para un SRI alcanzar ciertos umbrales en cuanto a la cantidad de recursos humanos y físicos eventualmente disponibles para ejecutar esfuerzos innovadores. En cuanto a los actores que realizan el esfuerzo, en ambos modelos destacan las empresas, que incluso están sobre las universidades en el modelo científico, lo que dice relación, por un lado, con el tipo de publicaciones elegidas como *output*, a saber, las que pertenecen al campo de las ciencias, la tecnología, la ingeniería y las ciencias de la salud (farmacéuticas), y por otro, al perfil de las regiones líderes como Alemania y Suecia, donde la relación empresa-universidad está muy desarrollada. Para capturar el efecto anterior, se hizo una nueva regresión (publicaciones*), incluyendo una variable que indique el gasto en I+D desarrollado por las universidades y las AAPP financiado por las empresas como porcentaje de su gasto total en I+D; es decir, una variable que dé cuenta de la interacción entre los actores del sistema. Esta variable no fue incluida en el factorial, ya que no hay datos desagregados a nivel regional, pero al incluirla por países se captura la relación empresa-universidad sobre el *output* científico, confirmando la hipótesis anterior¹⁴.

Conclusiones

En este trabajo se presentó 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 unas 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 de forma adecuada aspectos o subsistemas centrales del sistema regional. A partir de esta metodología, se elabora un índice de capacidad tecnológica y se realiza una tipología de los SRI en Europa, para terminar con una estimación de dos funciones de producción de conocimiento: una, tecnológica (*output* patentes) y otra, científica (*output* publicaciones científicas). Ambas estimaciones presentan un buen ajuste global y significancia estadística de las variables independientes para explicar los *outputs* (sin incurrir en colinealidad), lo cual destaca la

12 En concreto, se han realizado pruebas con las siguientes transformaciones: $\ln y - x$; $y - \ln x$; $\ln y - \ln x$; $1/y - x$; $y - 1/x$, así como $y^{1/2} - x$, habiendo obtenido con esta última los mejores resultados.

13 Se aplicó el test de Levin-Lin-Chu para raíces unitarias, confirmando la estacionariedad de los residuos de los modelos finales.

14 Se eliminó el factor AAPP para mejorar el ajuste global del modelo y evitar colinealidad.

Tabla 3. Resultados estimaciones. *Output*: patentes y publicaciones científicas (totales)

	MCO robustos		Panel intergrupos		Panel intragrupos			Panel efectos aleatorios	
FACTORES	Patentes	Publicaciones	Patentes	Publicaciones	Patentes	Publicaciones	Publicaciones*	Patentes	Publicaciones
Entorno regional	10.68 (0.000)	19.35 (0.000)	10.70 (0.000)	19.34 (0.000)	3.78 (0.000)	27.24 (0.000)	26.66 (0.000)	7.48 (0.000)	22.61 (0.000)
Empresas innovadoras	6.15 (0.000)	2.40 (0.000)	6.24 (0.000)	2.26 (0.013)	2.36 (0.000)	8.64 (0.000)	7.75 (0.000)	3.27 (0.000)	6.86 (0.000)
Universidades	0.35 (0.002)	9.21 (0.000)	0.36 (0.365)	9.45 (0.000)	0.35 (0.127)	6.26 (0.000)	5.46 (0.000)	0.06 (0.627)	6.66 (0.000)
AAPP	0.65 (0.000)	1.83 (0.000)	0.62 (0.113)	1.75 (0.054)	1.31 (0.000)	4.91 (0.000)		1.29 (0.000)	4.15 (0.000)
Sofisticación de la demanda	-0.11 (0.439)	-0.58 (0.073)	-0.18 (0.642)	-0.70 (0.443)	1.13 (0.000)	2.79 (0.000)	2.07 (0.008)	0.805 (0.000)	2.51 (0.000)
Constante	15.30 (0.000)	36.18 (0.000)	15.30 (0.000)	36.18 (0.000)	15.30 (0.000)	36.18 (0.000)	34.47 (0.000)	15.30 (0.000)	36.18 (0.000)
Gasto I+D (Univ. + AAPP) / Financ. empresas [%]]							0.280 (0.049)		
Sigma u					9.16	15.31	14.57	4.43	10.18
Sigma i					1.03	2.61	2.71	1.03	2.61
Rho					0.987	0.971	0.966	0.948	0.938
F test	635.83 (0.000)	478.07 (0.000)	203.92 (0.000)	116.81 (0.000)	20.20 (0.000)	55.73 (0.000)	39.31 (0.000)		
Test de Wald								947.68 (0.000)	2293.04 (0.000)
Log-Verosimilitud									
Test de Hausman								459.82 (0.000)	33.67 (0.000)
Chibar2 (01)									
R ²	0.881	0.812							
R ² intragrupos			0.115	0.461	0.192	0.599	0.568	0.175	0.641
R ² intergrupos			0.890	0.822	0.788	0.738	0.738	0.858	0.885
R ² global			0.881	0.812	0.781	0.733	0.733	0.851	0.870
Nº observaciones	1452	1452	1452	1452	1452	1452	1452	1452	1452
Nº grupos			132	132	132	132	132	132	132

En paréntesis los p valores. En itálica los coeficientes no significativos al 99% de confianza.
Fuente: elaboración propia.

utilidad de usar factores como regresores. 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. 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

- Acs, Z., Audretsch, D., & Feldman, M. (1992). Real effect of academic research: Comment. *The American Economic Review*, 82(1), 363-367.
- Aghion, P., Blundell, R., Griffith, R., Howitt, P., & Prantl, S. (2009). The effects of entry on incumbent innovation and productivity. *The Review of Economics and Statistics*, 91(1), 20-32.
- Anselin, L., Varga, A., & Acs, Z. (1997). Local geographic spillovers between university research and high technology innovations. *Journal of Urban Economics*, (42), 422-448.
- Asheim, B., & Gertler, M. (2005). The geography of innovation: Regional innovation systems. En Fagerberg, Mowery, & Nelson (eds.) *The Oxford handbook of innovation*. 1st edition. Nueva York: Oxford University Press.
- Audretsch, D. (1998). Agglomeration and the location of innovative activity. *Oxford Review of Economic Policy*, 14(2), 18-29.
- 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.
- Bordons, M., Fernández, M., & Gómez, I. (2002). Advantages and limitations in the use of impact factor measures for the assessment of research performance. *Scientometrics*, 53(2).
- Buesa, M., Heijs, J., Martínez Pellitero, M., & Baumert, T. (2006). Regional systems of innovation and the knowledge production function: The Spanish case. *Technovation*, 26(4), 436-472.
- Caniëls, M. C. J. (1996). Regional differences in technology: Theory and empirics. MERIT Research Memorandum RM1996-009. <http://merittbbs.unimaas.nl>.
- Cerulli, G., & Filipetty, A. (2012). The complementary nature of technological capabilities: Measurement and robustness issues. *Technological Forecasting & Social Change*, 79, 875-887.
- Edquist, Ch. (ed.). (1997). *Systems of innovation: Technologies, institutions and organizations*. London: Pinter Publishers.
- Edquist, Ch. (2005). Systems of innovation: Perspectives and challenges. En Fagerberg, Mowery, & Nelson (eds.) *The Oxford handbook of innovation*. 1st edition. Nueva York: Oxford University Press.
- Edquist, Ch., & Johnson, B. (1997). Institutions and organisations in systems of innovation. En Edquist (ed.) *Systems of innovation: Technologies, institutions and organizations*. London: Pinter Publishers.
- Etzkowitz, H., & Leydesdorff, L. (2000). The dynamics of innovation: From national systems and "Mode 2" to a Triple Helix of university-industry-government relations. *Research Policy*, 29, 109-123.
- Fagerberg, J., Mowery, D., & Nelson, R. R. (eds.). (2005). *The Oxford handbook of innovation*. 1st edition. Nueva York: Oxford University Press.
- Feldman, M. (1994). *The geography of innovation*. Dordrecht.
- Freeman, C. (1987). *Technology and economic performance: Lessons from Japan*. London: Pinter Publishers.
- Fritsch, M. (2002). Measuring the quality of regional innovation systems: A knowledge production function approach. *International Regional Science Review*, 25(1), 86-101.
- Furman, J. L., Porter, M. E., & Stern, S. (2002). The determinants of national innovative capacity. *Research Policy*, 31, 899-933.
- Griliches, Z. (1979). Issues in assessing the contribution of R&D productivity growth. *Bell Journal of Economics*, 10, 92-116.
- Griliches, Z. (1990). Patent statistics as economic indicators: A survey. *Journal of Economic Literature*, 28, 1661-1707.
- Grupp, H. (1997). *Messung und Erklärung des Technischen Wandels. Grundzüge einer empirischen Innovationsökonomik*. Berlin, Heidelberg et al.
- Grupp, H., & Schubert, T. (2010). Review and new evidence on composite innovation indicators for evaluating national performance. *Research Policy*, 39(1), 67-78.
- Hair, J. F., Anderson, R., Tatham, R., & Black, R. (2004). *Análisis multivariante de datos*. Madrid: Prentice Hall.

- Hall, B., Griliches, Z., & Hausman, J. (1986). Patents and R and D: Is there a lag? *International Economic Review*, 27(2), 265-283.
- Heijs, J. (2001). *Política tecnológica e innovación. Evaluación de la financiación pública de I+D en España*. CES, Madrid.
- 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.
- Hu, M. C., & Mathews, J. A. (2005). National innovative capacity in east Asia. *Research Policy*, 34(9), 1322-1349.
- Jaffe, A. (1989). Real effects of academic research. *American Economic Review*, 79(5), 957-970.
- Kline, S., & Rosenberg, N. (1986). An overview of innovation. En R. Landau, & N. Rosenberg (eds.). *The positive sum strategy: Harnessing technology for economic growth*. 1st edition. Washington, D. C.: National Academy of Sciences, 275-306.
- Krammer, S. M. S. (2009). Drivers of national innovation in transition: Evidences from a panel of Eastern European Countries. *Research Policy*, 38, 845-860.
- Li, X. (2009). China's regional innovation capacity in transition: An empirical approach. *Research Policy*, 38(2), 338-357.
- Lundvall, B. (1992). *National systems of innovation: Towards a theory of innovation and interactive learning*. 1st edition. London: Pinter Publishers.
- Makkonen, T., & Van Der Have, R. (2013). Benchmarking regional innovative performance: Composite measures and direct innovation counts. *Scientometrics*, 94(1), 247-262.
- Malerba, F., & Orsenigo, L. (1995). Schumpeterian patterns of innovation. *Cambridge Journal of Economics*, 19, 47-65.
- Mansfield, E. (1986). Patents and innovation: An empirical study. *Management Science*, 32(2), 173-181.
- Metcalf, J. S. (1995). The economic foundation of technology policy: Equilibrium and evolutionary perspectives. En Stoneman (ed.) *Handbook of Economics of Innovations and Technological Change*. Oxford: Wiley-Blackwell.
- Nelson, R., & Winter, S. G. (1982). *An evolutionary theory of economic change*. Cambridge and London.
- Nelson, R. R. (ed.). (1993). *National innovation systems: A comparative analysis*. Nueva York: Oxford University Press.
- Nelson, R., & Rosenberg, N. (1993). Technical innovation and national systems. En Nelson, R. (ed.) *National innovation systems: A comparative analysis*. Nueva York: Oxford University Press.
- OECD. (2004). *Compendium of patent statistics*. Paris.
- OECD. (2008). *Handbook on constructing composite indicators: Methodology and user guide*.
- Patel, P., & Pavitt, K. (1994). National innovation systems: Why they are important and how they might be measured and compared. *Economics of Innovation and New Technology*, 3(1), 77-95.
- Porter, M. E., & Stern, S. (1999). *Measuring the "ideas" production function: Evidence from the international patent output*. NBER Working Paper 7891.
- Romer, P. (1990). Endogenous technological change. *Journal of Political Economy*, 98, 72-102.
- Scherer, F. M. (1984). *Innovation and growth. Schumpeterian perspectives*. Cambridge (MA): MIT Press.
- Schmoch, U. (1999). Eignen sich Patente als Innovationssindikatoren? En Boch, R. (ed.). *Patentschutz und Innovation in Geschichte und Gegenwart*. Frankfurt am Main et al.
- Schumpeter, J. A. (1935). *The theory of economic development*. English edition. London: Transaction Publishers.
- Soete, L., & Patel, P. (1985). Recherche-développement, importations de technologie et croissance économique. Une tentative de comparaison internationale. *Revue Economique*, 36(5).
- Solow, R. (1956). A contribution to the theory of growth. *Quarterly Journal of Economics*, (70), 65-94.
- Van Raan, A. (2005). Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics*, 62(1), 133-143.
- Vence Deza, X. (1995). *Economía de la innovación y del cambio tecnológico*. Madrid: Siglo Veintiuno de España Editores S.A.
- Weingart, P. (2005). Impact of bibliometrics upon the science system: Inadvertent consequences? *Scientometrics*, 62(1), 117-131.
- Winkler, W. E. (2014). Matching and record linkage. *WIREs Computational Statistics*, 6(5), 313-325.