

UNIVERSIDAD DE CÓRDOBA



**Estudio del uso de los recursos de I+D+i en la
consecución de un Desarrollo Dostenible a través
de una Economía basada en el Conocimiento:
clasificación de países mediante el Análisis
Envolvente de Datos y Técnicas de Clasificación**

TESIS QUE PRESENTA

Mónica de la Paz Marín

PARA OPTAR AL GRADO DE DOCTOR POR LA UNIVERSIDAD DE
CÓRDOBA

CÓRDOBA, MAYO DE 2013

DIRECTORES

Carlos R. García Alonso
Pedro Antonio Gutiérrez Peña

Departamento de Gestión Empresarial y Métodos Cuantitativos

Centro de Ciencias Económicas y Empresariales (ETEA)

TITULO: *Estudio del uso de los recursos I+D+i en la consecución de un desarrollo sostenible a través de una economía basada en el conocimiento; clasificación de países mediante el Análisis Envolvente de Datos y Técnicas de Clasificación*

AUTOR: *MÓNICA DE LA PAZ MARÍN*

© Edita: Servicio de Publicaciones de la Universidad de Córdoba.
Campus de Rabanales
Ctra. Nacional IV, Km. 396 A
14071 Córdoba

www.uco.es/publicaciones
publicaciones@uco.es

La tesis titulada “Estudio del uso de los recursos de I+D+i en la consecución de un desarrollo sostenible a través de una economía basada en el conocimiento: clasificación de países mediante el análisis envolvente de datos y técnicas de clasificación”, que presenta D^a Mónica de la Paz Marín para optar al grado de Doctor, ha sido realizada dentro del programa de Doctorado “Ciencias Jurídicas y Sociales” del Departamento de Gestión Empresarial y Métodos Cuantitativos de la Universidad de Córdoba (ETEA) bajo la dirección del Doctor D. Carlos R. García Alonso y del Doctor D. Pedro Antonio Gutiérrez Peña.

En Córdoba, mayo de 2013

La Doctoranda:

Fdo: Mónica de la Paz Marín

Esta tesis doctoral ha sido financiada en parte con cargo a los Proyectos **TIN2011-22794** de la Comisión Interministerial del Ministerio de Ciencia y Tecnología (MICYT), al Proyecto de Excelencia **P2011-TIC-7508** de la Junta de Andalucía y con fondos FEDER de la Unión Europea.



TÍTULO DE LA TESIS:

ESTUDIO DEL USO DE LOS RECURSOS DE I+D+i EN LA CONSECUCCIÓN DE UN DESARROLLO SOSTENIBLE A TRAVÉS DE UNA ECONOMÍA BASADA EN EL CONOCIMIENTO: CLASIFICACIÓN DE PAÍSES MEDIANTE EL ANÁLISIS ENVOLVENTE DE DATOS Y TÉCNICAS DE CLASIFICACIÓN

DOCTORANDO/A:

MÓNICA DE LA PAZ MARÍN

INFORME RAZONADO DEL/DE LOS DIRECTOR/ES DE LA TESIS

(se hará mención a la evolución y desarrollo de la tesis, así como a trabajos y publicaciones derivados de la misma).

La Tesis Doctoral titulada:

ESTUDIO DEL USO DE LOS RECURSOS DE I+D+i EN LA CONSECUCCIÓN DE UN DESARROLLO SOSTENIBLE A TRAVÉS DE UNA ECONOMÍA BASADA EN EL CONOCIMIENTO: CLASIFICACIÓN DE PAÍSES MEDIANTE EL ANÁLISIS ENVOLVENTE DE DATOS Y TÉCNICAS DE CLASIFICACIÓN

es la culminación de un proceso intensivo de investigación científica sobre la importancia de la investigación científica, el desarrollo y la innovación en el bienestar de los países. Durante su desarrollo, se ha asistido a relevantes eventos científicos y publicado artículos científicos en revistas indexadas. El proceso ha incluido: i) la revisión sistemática e intensiva de la bibliografía científica relacionada con la materia objeto del estudio, ii) el diseño y desarrollo de un modelo causal (Red Bayesiana), iii) la adaptación de herramientas informáticas para la resolución de los distintos problemas matemático-estadísticos planteados, iv) la validación de las distintas metodologías utilizando tanto el conocimiento experto como información secundaria y, finalmente, v) la elicitación de conocimiento experto a partir del análisis de los resultados obtenidos sobre el que se fundamentaron las conclusiones del trabajo presentado.

La presente Tesis Doctoral es, por tanto, resultado de un esfuerzo conjunto entre la universidad y expertos en diferentes organismos, siendo en nuestra opinión un ejemplo de investigación científica que trasciende el mero conocimiento académico y se proyecta sobre la planificación y gestión de situaciones reales. Teniendo en cuenta esta característica, las principales dificultades que se afrontaron fueron: i) la estructura, diseño y temporalidad de las bases de datos disponibles, ii) la complejidad teórica que existe en este ámbito científico, iii) la inexistencia de bases de conocimiento capaces de interpretar los datos secundarios de acuerdo con el conocimiento experto y, finalmente, iv) la adaptación de metodologías de análisis matemático y estadístico para la correcta interpretación del fenómeno. Todas estas dificultades se han abordado con rigurosidad científica y, como ya se ha mencionado, validado analizando las políticas de I+D+i de los países de la Unión Europea.

Las metodologías utilizadas en este trabajo de investigación pueden ser de especial utilidad en el ámbito del estudio y de la planificación de políticas públicas de I+D+i. Basándose en la evidencia de los resultados obtenidos, los entes decisores tienen a su disposición un conjunto flexible de herramientas para apoyar sus decisiones en este marco de referencia.

En la actualidad, el equipo de investigación del cual forma parte esencial la doctoranda continúa trabajando en estos temas, existiendo la posibilidad cierta de extender la investigación utilizando nuevas metodologías de análisis e incorporando información de otras políticas públicas de I+D+i.

Por todo ello consideramos que la Tesis Doctoral presentada tiene todos los avales necesarios para su presentación y defensa.

Congresos:

Comunicación en XVII SIGEF Congress. (Reus-Cambrils, España). Methods for Decision Making in an uncertain Environment. Julio 2012

Ponencia Congreso Internacional. 9th International Conference on Data Envelopment Analysis: DEA and Performance Management, Thessaloniki (Greece). University of Makedonia. Evaluation of R&D public policy in the European Union: an Expert Knowledge-based and two-stage DEA Approach (De la Paz-Marín, Mónica; García-Alonso, Carlos). Agosto 2011.

Artículos

Fernández-Navarro, F., Campoy-Muñoz, P., De la Paz-Marín, M., Hervás-Martínez, C., Yao, X., 2013. Addressing the EU Sovereign Ratings Using an Ordinal Regression Approach, IEEE Systems, Man, And Cybernetics, PART C: Applications And Reviews, D.O.I.: 10.1109/TSMCC.2013.2247595. Índice de Impacto: 2,009.

De la Paz Marín, M., Campoy Muñoz, P., Hervás Martínez, C., 2012. Non-linear multiclassifier model based on Artificial Intelligence to predict Research and Development Performance in European Countries, Technological Forecasting and Social Change 79, 1731-1745. Índice de Impacto: 1,709.

Capítulos de Libros

De la Paz-Marín, M., Campoy Muñoz, P., Hervás Martínez, C., 2012. Evolutionary Neural Network Classifiers for monitoring Research, Development and Innovation performance in European Union Member States, in: Methods for Decision Making in an uncertain Environment. Proceedings of the XVII SIGEF Congress, 309-333.

Por todo ello, se autoriza la presentación de la tesis doctoral.

Córdoba, 13 de Mayo de 2013

Firma del/de los director/es



Fdo.: CARLOS GARCÍA ALONSO Fdo.: PEDRO A. GUTIÉRREZ PEÑA

Dedicada a mi querido padre.

No te olvido...

AGRADECIMIENTOS

Resulta tan difícil agradecer en pocas palabras a todos y cada uno de los que, de alguna u otra forma, me han apoyado y participado en que este trabajo sea posible, que este apartado de agradecimientos ha requerido un verdadero esfuerzo de síntesis.

A Carlos R. García Alonso, por haberme atendido desde el primer momento, por compartir sus conocimientos de esa forma que sólo él sabe y por su infinita paciencia conmigo y buen humor.

A Pedro Antonio Gutiérrez Peña, por su asombrosa capacidad de trabajo, amabilidad y dedicación.

A los miembros del tribunal de esta tesis y, en especial, a César Hervás Martínez: sencillamente, POR SER COMO ES.

Al equipo de profesores de ETEA: qué alegría volver a encontrarme con ellos y comprobar que siguen siendo tan grandes docentes y motivadores después de “algunos” años.

A mis compañeros de investigación y, especialmente, a Pilar Campoy, por estar siempre ahí, y a los miembros del grupo de investigación AYRNA de la Universidad de Córdoba.

A mis amigos: por sus ánimos, desvelos, preocupaciones e inestimables esfuerzos para que desconectara y cogiera fuerzas renovadas de vez en cuando (son demasiados para no dejarme nadie atrás).

A ti, Fernando, por estar siempre dispuesto en la distancia a lo que me hiciera falta.

A mis compañeros de trabajo en la Universidad, especialmente a Francisco Alcántara, por esas charlas tan enriquecedoras y porque fue él precisamente el que me animó desde un principio a que obtuviese el Grado de Doctor. A mi dedicada Gloria, por ofrecerme desinteresadamente tanta ayuda. Y a mi Director, Carlos Lorenzo, por ser más que un jefe.

Al equipo de Doctorado de la UCO: Pilar, Valle, cuánto me habéis apoyado...

Y, como no, a mi madre y hermanas. A toda mi familia. A ella, agradecimiento, pero también ofrecerles mis más sinceras disculpas por no dedicarles durante todo este tiempo la atención que se merecen. Espero que algún día mi querido esposo Miguel y mis adorados hijos, Miguel Ángel y Álvaro Manuel, sepan entender mi esfuerzo y me perdonen tantas horas robadas de un tiempo que les pertenece.

A todos aquellos a los que no he nombrado, pero no por ello menos importantes (seguro que ellos saben quiénes son)...GRACIAS.

ÍNDICE GENERAL

ACRÓNIMOS Y ABREVIATURAS	13
LISTADO DE TABLAS.....	15
LISTADO DE FIGURAS.....	16
GLOSARIO DE TÉRMINOS.....	17

BLOQUE 0. INTRODUCCIÓN

CAPÍTULO INTRODUCTORIO.....	27
-----------------------------	----

BLOQUE 1. MARCO TEÓRICO Y CONCEPTUAL

CAPÍTULO 1. CAMBIO TECNOLÓGICO, INNOVACIÓN Y CAPITAL HUMANO EN EL PENSAMIENTO ECONÓMICO.....	45
1. EL CAMBIO TECNOLÓGICO Y LA INNOVACIÓN EN LAS TEORÍAS DEL CRECIMIENTO ECONÓMICO	45
1.1. LA TEORÍA CLÁSICA DEL CRECIMIENTO ECONÓMICO.....	46
1.2. NEOCLÁSICOS (TEORÍAS DE CRECIMIENTO EXÓGENO Y ENDÓGENO)	48
1.2.1. Robert M. Solow.....	48
1.2.2. Paul M. Romer.....	51
1.2.3. Robert E. Lucas Jr.....	53
1.2.4. Los poskeynesianos.....	54
1.3. TEORÍA ECONÓMICA EVOLUCIONISTA (NEOSCHUMPETERIANOS).....	55
1.4. TEORÍA CLÁSICA DEL DESARROLLO	56
1.5. GENERAL PURPOSE TECHNOLOGIES.....	57
1.6. LOS FUTURISTAS	58

CAPÍTULO II. HACIA UNA ECONOMÍA BASADA EN EL CONOCIMIENTO Y UN MODELO DE DESARROLLO SOSTENIBLE. EL PAPEL DE LA I+D+I	61
1. LA ECONOMÍA DEL CONOCIMIENTO	61
2. EL DESARROLLO SOSTENIBLE.....	64
3. INTERRELACIÓN ENTRE LA I+D+I, LA ECONOMÍA DEL CONOCIMIENTO Y EL DESARROLLO SOSTENIBLE.	67
CAPÍTULO III: SEGUIMIENTO Y EVALUACIÓN DE POLÍTICAS PÚBLICAS. LOS RECURSOS PÚBLICOS DE LA I+D+I Y LA POLÍTICA CIENTÍFICA Y TECNOLÓGICA EN LA UE.	73
1. LA EVALUACIÓN DE POLÍTICAS PÚBLICAS EN GENERAL. CONCEPTO, FUNCIONES Y RELEVANCIA.....	73
2. CONCEPTOS RELACIONADOS CON LA EVALUACIÓN DE LOS RECURSOS EMPLEADOS ...	76
3. EL ESTUDIO DE LOS RECURSOS DEDICADOS A LA I+D+I.....	79
3.1. USO DE INDICADORES DE I+D+I A NIVEL INTERNACIONAL.....	81
4. LA POLÍTICA DE CIENCIA, TECNOLOGÍA E INNOVACIÓN EN LA UNIÓN EUROPEA.....	83
4.1. ANTECEDENTES	83
4.2. LOS PROGRAMAS MARCO DE INVESTIGACIÓN Y DESARROLLO TECNOLÓGICO.....	85
4.3. ACTIVIDAD RECIENTE DE LA POLÍTICA CIENTÍFICA, TECNOLÓGICA Y DE INNOVACIÓN DE LA UE.....	86
4.3.1. Hitos históricos	86
4.3.2. La Estrategia de Lisboa	87
4.3.3. La Estrategia Europa 2020 para un crecimiento inteligente	89

BLOQUE 2. MODELO CONCEPTUAL CAUSAL (RED BAYESIANA)

CAPÍTULO IV. LOS MODELOS CONCEPTUALES Y LA CAUSALIDAD. LAS REDES BAYESIANAS.	95
1. MODELOS CONCEPTUALES EN GENERAL. LA CAUSALIDAD	95
2. LAS REDES BAYESIANAS.....	98
CAPÍTULO V. DISEÑO DEL MODELO CONCEPTUAL: EL ANÁLISIS COOPERATIVO BASADO EN CONOCIMIENTO EXPERTO (EBCA)	103
CAPÍTULO VI. PRESENTACIÓN DEL MODELO CONCEPTUAL CAUSAL (RED BAYESIANA).....	109
1. BASE TEÓRICA Y CONCEPTUAL	109
2. VARIABLES (NODOS) Y ESTRUCTURA DEL MODELO CONCEPTUAL CAUSAL PROPUESTO	111

BLOQUE 3. ANÁLISIS EMPÍRICO

CAPÍTULO VII. EFICIENCIA DE LAS POLÍTICAS PÚBLICAS DE I+D+I DE LA UE EN EL MARCO DE UN DESARROLLO SOSTENIBLE	129
1. INTRODUCCIÓN.....	129
2. EVALUACIÓN DEL MODELO CONCEPTUAL CAUSAL A TRAVÉS DE UN MÉTODO DE ANÁLISIS CLÁSICO: LA EFICIENCIA DE LAS POLÍTICAS PÚBLICAS DE I+D+I EN LA UE	131
2.1. EL ANÁLISIS ENVOLVENTE DE DATOS (DEA).....	131
2.2. ESTRUCTURA DE LOS ESCENARIOS PARA EL DEA: INPUTS Y OUTPUTS SELECCIONADOS.....	133

2.3. EL ENFOQUE DEA EN DOS PASOS ENLAZADOS (2LS-DEA).....	134
2.4. CONJUNTO DE DATOS.....	135
2.5. LA TRANSFORMACIÓN DE INPUTS Y OUTPUTS: LA BASE DE REGLAS.....	137
3. RESULTADOS DEL ANÁLISIS DE LA EFICIENCIA DE LAS POLÍTICAS PÚBLICAS DE I+D+I.....	139
4. CONCLUSIONES DEL ANÁLISIS DE LA EFICIENCIA TÉCNICA RELATIVA DE LAS POLÍTICAS DE I+D+I EN EL MARCO DE UN DESARROLLO SOSTENIBLE	155

CAPÍTULO VIII. ANÁLISIS DE LOS RESULTADOS DE LA I+D+I EN LA UE...157

1. INTRODUCCIÓN.....	157
2. METODOLOGÍA.....	161
2.1. AGRUPAMIENTO K-MEDIAS Y DESCRIPCIÓN DE LAS CLASES	161
2.2. PROBLEMAS DE CLASIFICACIÓN Y REGRESIÓN.....	163
2.2.1. Redes Neuronales.....	168
2.2.2. Clasificación Multiclase y métricas de evaluación	170
2.2.3. Algoritmos de aprendizaje	174
3. DISEÑO EXPERIMENTAL	179
3.1. DESCRIPCIÓN DE LA BASE DE DATOS	179
3.2. RESULTADOS	180
3.3. EL MEJOR MODELO	182
4. DISCUSIÓN.....	184
5. CONCLUSIONES DEL ANÁLISIS DE LOS RESULTADOS DE I+D+I	186

CAPÍTULO IX. ANÁLISIS DESDE EL PRISMA DE UNA ECONOMÍA BASADA EN EL CONOCIMIENTO

1. INTRODUCCIÓN.....	187
----------------------	-----

2. REVISIÓN DE METODOLOGÍAS	189
3. METODOLOGÍA.....	192
3.1. AGRUPAMIENTO JERÁRQUICO Y DESCRIPCIÓN DE CLASES	192
3.2. CLASIFICADORES NOMINALES Y ORDINALES EVALUADOS.....	195
4. DISEÑO EXPERIMENTAL	200
4.1. DESCRIPCIÓN DE LA BASE DE DATOS	200
4.2. MÉTRICAS DE EVALUACIÓN	205
4.3. RESULTADOS	205
5. DISCUSIÓN Y CONCLUSIONES DEL ANÁLISIS DESDE EL ENFOQUE DE UNA ECONOMÍA BASADA EN EL CONOCIMIENTO	210

BLOQUE 4. DISCUSIÓN, CONCLUSIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN

CAPÍTULO X. DISCUSIÓN, CONCLUSIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN	215
1. DISCUSIÓN	215
2. CONCLUSIONES	219
3. LÍNEAS DE INVESTIGACIÓN FUTURAS	227
REFERENCIAS.....	231
ANEXO I. PUBLICACIONES.....	247
1. TRABAJOS PUBLICADOS	247
2. TRABAJOS ENVIADOS Y EN FASE DE REVISIÓN	299

ACRÓNIMOS Y ABREVIATURAS

2LS-DEA: Two linked-steps Data Envelopment Analysis (Análisis Envolverte de Datos en dos pasos enlazados).

APEs: Algoritmos de Programación Evolutiva.

CMMAD: Comisión Mundial sobre el Medio Ambiente y el Desarrollo de las Naciones Unidas.

CE: Computación Evolutiva.

CCR: Correct Classification Rate (tasa de clasificación correcta).

DS: Desarrollo Sostenible.

EANN: Evolutionary Artificial Neural Network (Red Neuronal Evolutiva) .

EA: Evolutionary Algorithm (Algoritmo Evolutivo) .

EC: Economía del conocimiento.

EEMM: Estado Miembro de la Unión Europea.

EPUNN: Evolutionary Product Unit Neural Networks (Redes Neuronales Evolutivas de Unidad Producto).

ESUNN: Evolutionary Sigmoidal Unit Neural Networks (Redes Neuronales Evolutivas de Unidad Sigmoide).

GAD: Grafo Acíclico Dirigido.

GPTs: General Purpose Technologies (Tecnologías de Utilidad General).

I+D+i: Investigación, Desarrollo e Innovación.

IA: Inteligencia Artificial.

IUS: Innovation Union Scoreboard (Cuadro de Indicadores de Innovación de la Unión).

KAM: Knowledge Assessment Methodology (Metodología de Evaluación del Conocimiento).

KEI: Knowledge Economy Index (Índice de Economía del Conocimiento).

MAE: Mean Absolute Error (Media Absoluta del Error).

MLP: Multilayer Perceptrón (Perceptrón Multicapa).

MSE: Mean Square Error (Error cuadrático medio).

NEM: Nuevos Estados Miembros.

NSF: National Science Foundation (Fundación Nacional de la Ciencia).

OECD: Organisation for Economic Co-operation and Development (Organización para la Cooperación y el Desarrollo Económico).

POM: Proportional Odd Model.

PU: Product Unit (Unidad Producto).

RBFs: Radial Basis Functions (Funciones de Base Radial).

RNAs: Redes Neuronales Artificiales.

SU: Sigmoid Unit (Unidad Sigmoide).

SVMs: Support Vector Machines (Máquinas de Vectores Soporte).

SVOR: Support Vector Ordinal Regression (Regresión Ordinal de Vectores Soporte).

SVORIM: Support Vector Ordinal Regression with Implicit Constraints (Regresión Ordinal de Vectores Soporte con restricciones implícitas de orden).

SVOREX: Support Vector Ordinal Regression with Explicit Constraints (Regresión Ordinal de Vectores Soporte con restricciones explícitas de orden).

TICs: Tecnologías de la Información y de la Comunicación.

U.N.C.T.A.D.: Conferencia de las Naciones Unidas sobre Comercio y Desarrollo.

U.N.D.P.: United Nations Development Programme (Programa de las Naciones Unidas para el Desarrollo).

UE: Unión Europea.

WIPO: World Intellectual Property Organization (Organización Mundial de la Propiedad Intelectual).

LISTADO DE TABLAS

Tabla 0.1. Relación entre objetivos e hipótesis planteados en la investigación	36
Tabla 3.1. Principales dificultades para la evaluación de políticas públicas.....	76
Tabla 3.2. Nociones relacionadas con la eficiencia	77
Tabla 6.1. Variables de la Red Bayesiana (RB) y seleccionas para el análisis 2LS-DEA.....	122
Tabla 6.2. Descripción del resto de variables en la Red Bayesiana (RB) no incluidas en ningún escenario analizado en el 2LS-DEA.....	125
Tabla 7.1. Estadísticos descriptivos de los inputs y outputs empleados.....	136
Tabla 7.2. Base de reglas para la transformación de los inputs/outputs de acuerdo a los valores de las variables padre en la red Bayesiana.....	140
Tabla 7.3. Eficiencia técnica relativa de las políticas públicas de I+D+i de los países de la UE (%) (DMUs) por escenario (100 = eficiente). Periodo 2006-2008.....	150
Tabla 8.1. Descripción de las variables, centroides (media) de los grupos \pm Desv. Típicas.....	164
Tabla 8.2. Matriz de confusión de un clasificador	172
Tabla 8.3. Resultados estadísticos del CCR_G , MS_G , K_G y número de conexiones (#conn.)	182
Tabla 8.4. Media y Desviación Típica (DT) de los métodos estocásticos	182
Tabla 8.5. Expresión de la probabilidad y resultados del mejor modelo ESUNN	183
Tabla 8.6. Influencia de cada variable en la probabilidad de pertenecer a un determinado grupo.....	184
Tabla 9.1. Estadísticos descriptivos de los grupos (media o centroide, desviación estándar y patrón país-año de referencia).....	197
Tabla 9.2. Descripción de las variables empleadas y sus estadísticos descriptivos (cont.).....	203
Tabla 9.3. CCR_G y MAE_G para el conjunto de generalización de los métodos evaluados.....	206
Tabla 9.4. Matriz de confusión para el mejor clasificador ordinal (SVORIM).....	208

LISTADO DE FIGURAS

Figura 1.1. Esquema del trabajo de investigación	41
Figura 2.1. Tres visiones de la economía del conocimiento	64
Figura 2.2. Relación anidada entre la I+D+i, la EC y el DS para la consecución del bienestar	69
Figura 3.1. Relevancia de la evaluación de políticas públicas de I+D+i.....	82
Figura 4.1. Resumen de diez años de la Estrategia de Lisboa	88
Figura 5.1. Metodología propuesta para la mejora del diseño del modelo conceptual.....	105
Figura 6.1. Representación del escenario Tradicional en la red Bayesiana	113
Figura 6.2. Representación del escenario de Innovación en la red Bayesiana.....	114
Figura 6.3. Escenarios de Salud, Medio Ambiente, Sociedad de la Información.....	116
Figura 6.4. Escenario Económico representado en la red Bayesiana	117
Figura 6.5. Variables que influyen en la competitividad de un país.....	119
Figura 6.6: Modelo conceptual causal (red Bayesiana) del crecimiento económico basado en el conocimiento y la innovación	121
Figura 7.1. Esquema del enfoque DEA en dos pasos enlazados (2LS-DEA)	135
Figura 8.1. Modelo de red utilizado en problemas de multclasificación.....	168
Figura 9.1. Dendograma obtenido mediante la técnica de agrupamiento jerárquico.....	194

GLOSARIO DE TÉRMINOS

Casi toda palabra que escribimos o pronunciamos, contiene significados múltiples. Esto conduce a interpretaciones distintas, hecho que es especialmente relevante en el ámbito científico y por lo que se requiere mayor precisión. Así, resulta oportuno definir algunos conceptos relacionados con la Ciencia y la Tecnología, sin perjuicio de que en otras secciones se amplíen o maticen.

Cambio tecnológico: Parayil (1991, p. 294) definió el cambio tecnológico como “un proceso temporal y acumulativo que incrementa la habilidad de la gente o de la sociedad para resolver sus problemas sociales, económicos y cotidianos”. El elemento central de su concepción es el conocimiento y en ello se fundamenta el proceso de cambio tecnológico. Si bien la manifestación del cambio es tangible en forma de artefactos materiales, la parte integral de este proceso es de naturaleza cognitiva. En otro sentido, puede ser caracterizado como el efecto combinado de varias actividades tecnológicamente relacionadas como la invención, la innovación, desarrollo, transferencia y difusión. Puede ser visto como un conjunto de actividades enfocadas a la solución de problemas.

Competitividad: este concepto es definido por el Foro Económico Mundial (World Economic Forum o WEF) como el “conjunto de instituciones, políticas y factores que determinan el nivel de productividad de un país” (WEF, 2010). Las economías más competitivas tienden a producir niveles de ingresos mayores para sus ciudadanos. El nivel de productividad determina la tasa de retorno obtenida por las inversiones en la economía (físicas, humanas y tecnológicas) y estas tasas de retorno son otro de los conductores fundamentales del crecimiento; una economía más competitiva es probable que crezca más rápido en el medio y largo plazo (WEF, 2010).

Desarrollo sostenible: definido por las Naciones Unidas como “el desarrollo que satisface las necesidades de las generaciones actuales sin comprometer la habilidad de las generaciones futuras de satisfacer sus propias necesidades” (Brundtland, 1987). Este concepto se refiere principalmente al aseguramiento del bienestar humano a largo plazo que implica afrontar los retos de los recursos naturales limitados y la pobreza global del planeta. En otras palabras, “...una vida larga y saludable, acceso a la educación, participación en la vida política y social y

trabajos bien pagados le dan la oportunidad a las personas de alcanzar sus objetivos, esperanzas y aspiraciones” (Programa de las Naciones Unidas para el Desarrollo, 2011).

Difusión del progreso tecnológico: es el proceso que se manifiesta en el creciente número de los usuarios efectivos de una innovación, o si se prefiere, en la adopción de ésta por los agentes y a través del conjunto de la economía. De esta definición puede inferirse que el impacto económico que una innovación posee depende fundamentalmente del proceso de difusión. Como es natural, no todas las innovaciones acaban siendo difundidas de una forma exitosa. “La idea de difusión hace referencia a filtración de los conocimientos que los agentes (...) generan, los cuales representan una fuente central de sus ventajas competitivas.” (Erbes et al., 2006, p. 52). Es importante incidir en el hecho de que en algunas ocasiones la difusión se realiza por medio de la imitación.

Economía basada en el conocimiento: su definición más conocida es la realizada por la Organización para la Cooperación y el Desarrollo Económico (OECD, en sus siglas en inglés) como “aquella que se basa directamente en la producción, distribución y uso del conocimiento y la información” (OECD, 1999).

Innovación: Debido a la importancia del concepto de innovación para esta tesis, nos detendremos con mayor detalle en él. Existen muchas formas alternativas de definir la innovación sin variar el sentido sustancial del concepto. Según el Manual de Oslo (OECD, 2005) se puede definir como “la implementación de un producto (bien o servicio) o proceso nuevo o con un alto grado de mejora, o un método de comercialización u organización nuevo aplicado a las prácticas de negocio, al lugar de trabajo o a las relaciones externas”.

Una definición célebre de innovación es la de Jacob Schmookler, quien propone la siguiente: “cuando una empresa produce un bien o servicio o utiliza un método o input que le es novedoso, está realizando cambio técnico. La primera empresa que produce un determinado cambio técnico es un innovador. Su acción es la innovación”.

La conceptualización clásica del concepto de innovación se debe a Schumpeter (1961), entendiéndose que existen cinco tipos de innovación:

- 1) La introducción de un nuevo bien, esto es, uno con el que no se hayan familiarizado los consumidores, o la nueva calidad de un bien; 2) la introducción de un nuevo método de producción, esto es, de uno no probado por la experiencia en la rama de la manufactura de que se trate, que no precisa fundarse en un descubrimiento nuevo desde el punto de vista

científico y que puede consistir simplemente en una nueva forma de manejar comercialmente una mercancía; 3) la apertura de un nuevo mercado, esto es, un mercado en el cual no haya entrado la rama especial de la manufactura del país de que se trate, a pesar de que existiera anteriormente dicho mercado; 4) la conquista de una nueva fuente de aprovisionamiento de materias primas o de bienes semimanufacturados, haya o no existido anteriormente, como en los demás casos; 5) la creación de una nueva gran organización en cualquier industria, como la de una posición de monopolio (por ejemplo, por la formación de un trust) o bien la anulación de una posición de monopolio existente con anterioridad.

Lo más interesante de esta clasificación (aunque hay muchas otras) es que, además de seguir siendo actualmente utilizada por la OECD, deja claro que no todas las innovaciones implican algún tipo de cambio técnico. Como puede apreciarse, las tres últimas carecen de contenido tecnológico: son innovaciones no tecnológicas. En este trabajo, nos ocuparemos tanto de las innovaciones tecnológicas como de las no tecnológicas.

Innovación tecnológica: surge tras la utilización de la tecnología como medio para introducir un cambio en la empresa. Se trata del acto o proceso consistente en acoplar en un país o ámbito espacial preciso una nueva oportunidad tecnológica con una necesidad o, en su caso, con una demanda solvente. El desarrollo de la tecnología permite obtener nuevos descubrimientos, nuevas formas de solucionar los problemas que son inherentes a la actividad productiva. A estas manifestaciones se les ha designado con el concepto de innovación tecnológica.

Su origen puede ser una invención o el producto más inmediato de una transferencia de tecnología. Como se ve, si la invención era la generación de nuevas ideas, la innovación abarcaría el desarrollo de nuevas ideas transformadas en productos que pueden intercambiarse en el mercado o en procesos para generar productos. Como es fácil deducir de aquí, no todas las invenciones conducen necesariamente a innovaciones, si las primeras no encuentran una aplicación comercial. La innovación surge de los procesos anteriormente descritos y de la función empresarial que consistiría en la decisión de lanzar al mercado el invento, la obtención de recursos para ello y la organización del modo de llevarlo a cabo. Por otro lado, Afuah (1999) afirma que la innovación no tecnológica actúa como facilitador de la innovación tecnológica.

Investigación y desarrollo experimental (I+D): según el Manual de Frascati (OECD, 2003; p. 30), la investigación y el desarrollo experimental (I+D) comprenden el trabajo creativo

llevado a cabo de forma sistemática para incrementar el volumen de conocimientos, incluido el conocimiento del hombre, la cultura y la sociedad, y el uso de esos conocimientos para crear nuevas aplicaciones. El término I+D engloba tres actividades: investigación básica, investigación aplicada y desarrollo experimental.

- **Investigación básica:** consiste en “trabajos experimentales o teóricos que se emprenden principalmente para obtener nuevos conocimientos acerca de los fundamentos de los fenómenos y hechos observables, sin pensar en darles ninguna aplicación o utilización determinada”. Es la búsqueda de algún conocimiento original para el avance de la ciencia pura, sin pensar en objetivos comerciales específicos. Se trata de una actividad de investigación no dirigida estrictamente hacia aplicaciones prácticas. La investigación básica intenta generar conocimientos científicos sin necesidad de obtener resultados económicos. Sin embargo, el que estos conceptos o principios científicos todavía no posean una utilidad directa en la producción de bienes y servicios no quiere decir que no puedan llegar a tenerla tras un esfuerzo posterior de estudio y experimentación.
- **Investigación aplicada:** es la que va dirigida a descubrir un nuevo conocimiento científico para objetivos específicamente comerciales. Pretende soluciones prácticas a problemas concretos y determinados y su objetivo es plasmar los avances teóricos obtenidos en la investigación básica en nuevos productos o procesos a través de la investigación científica. El resultado de la investigación aplicada es la invención o invento que es susceptible de registro en el sistema de patentes. Consiste también en “trabajos originales realizados para adquirir nuevos conocimientos; sin embargo, está dirigida fundamentalmente hacia un objetivo práctico específico” (OECD, 2003; p. 30).
- **Desarrollo experimental:** actividad técnica de naturaleza no rutinaria dedicada a trasladar los productos y procesos resultantes de la investigación básica o aplicada a la actividad normal de la empresas. El desarrollo introduce nuevos productos en los mercados para su explotación comercial y nuevos procesos. La Investigación y Desarrollo (I+D, en lo sucesivo) constituye un espectro de actividades relacionadas con la actividad económica que sigue la secuencia descrita. Comenzaría con unas actividades básicas de investigación inciertas y poco concretas y avanza por etapas de más concreción y aplicación comercial, a la vez que la incertidumbre técnica y económica asociada a cada una de las fases va disminuyendo, mientras que el grado de

rentabilidad y apropiabilidad de los resultados observa un comportamiento opuesto. Según el Manual de Frascati, consiste en trabajos sistemáticos que aprovechan los conocimientos existentes obtenidos de la investigación y/o la experiencia práctica, y está dirigido a la producción de nuevos materiales, productos o dispositivos; a la puesta en marcha de nuevos procesos, sistemas y servicios, o a la mejora sustancial de los ya existentes (OECD , 2003; p. 30).

A efectos de encuestas, debe distinguirse la I+D de una amplia gama de actividades afines que tienen una base científica y tecnológica. Esas otras actividades están muy relacionadas con la I+D, pero, en lo posible, no deberían tenerse en cuenta para su medición cuando no están directamente relacionadas con proyectos concretos de I+D. Esas actividades se resumen en cuatro apartados: enseñanza y formación, otras actividades científicas y tecnológicas afines, otras actividades industriales y administración y otras actividades de apoyo (OECD, 2003).

Invencción o invento: es la primera creación de un conocimiento que puede o no proceder de las diversas ciencias. La invención es la generación de nuevas ideas (Pampillón, 1991).

Patente: una patente es un derecho exclusivo concedido a una invención, es decir, un producto o procedimiento que aporta, en general, una nueva manera de hacer algo o una nueva solución técnica a un problema (WIPO, 2012). Según el Manual de Estadísticas de Patentes de la OECD (2009), las patentes son derechos temporales, válidos para un plazo de 20 años máximo después de la fecha de la solicitud, después del cual la invención que protege entra en el dominio público. Las patentes son derechos territoriales que sólo se aplican al país para el que las patentes han sido concedidas. Las patentes se conceden a las invenciones en todos los campos de la tecnología. En general, las leyes de la naturaleza, los fenómenos naturales y las ideas abstractas no son patentables (por supuesto, existe un debate sobre dónde están límites).

Progreso tecnológico: comprende ciertas clases de conocimiento que hacen posible producir: 1) un mayor volumen de producción total o 2) una producción total cualitativamente superior, a partir de una determinada cantidad de recursos (Rosenberg 1982, p. 3). Pampillón entiende que el progreso tecnológico está íntimamente relacionado con la tecnología (Pampillón, 1991, p. 5). Stoneman, sin embargo, enfatiza que el cambio técnico se iguala con el avance (progreso) técnico, pero debería quedar inmediatamente claro que ambos términos no son sinónimos. No todo cambio es un avance, aunque todo avance será un cambio (Stoneman, 1987, p. 8). Una definición más pragmática de ambos conceptos, entiende la

tecnología como “el conjunto social del conocimiento de las disciplinas industriales, siendo el cambio tecnológico, el cambio de dicho conjunto” (Schmookler, 1966, p. 1). Entenderemos por tecnologías los “conjuntos de conocimientos y por técnicas los sistemas físicos que los hacen operacionales” (Vegara, 1989, p. 16).

Sistema Nacional de Innovación: El enfoque de los sistemas nacionales de innovación (NIS en sus siglas en inglés) destaca que los flujos de la tecnología y la información entre la gente, las empresas y las instituciones son clave para el proceso innovador. Las definiciones más comunes de NIS son las siguientes:

“... la red de instituciones en los sectores públicos y privados cuyas actividades e interacciones inician, importan, modifican y difunden nuevas tecnologías” (Freeman, 1987).

“... los elementos y relaciones que interactúan en la producción, difusión y uso del nuevo, y económicamente útil, conocimiento... y se encuentran localizados o enraizados dentro de las fronteras de una nación” (Lundvall, 1992).

“... un conjunto de instituciones cuyas interacciones determinan los resultados de la innovación... de las firmas nacionales” (Nelson, 1993).

Técnica: Según Mokyr (2005, p. 286), “La tecnología es conocimiento (...). La unidad básica de análisis de la tecnología es la técnica. Una técnica es un conjunto de instrucciones sobre cómo producir bienes y servicios (...). Mientras la tecnología a menudo depende de los artefactos, los artefactos no son lo mismo que la técnica y lo que define la técnica es el contenido de las instrucciones”. En el mismo sentido, se refiere Agazzi (1998), quien define la técnica como una exhibición de habilidades prácticas que facilitan la realización de ciertas actividades (Agazzi, 1998, p. 2). Para este autor, la tecnología es fundamentalmente el dominio de la técnica, es el uso del conocimiento científico para manipular las diferentes técnicas.

Tecnología: Según Pampillón (1991, p. 5), puede definirse como “un factor de producción formado por un conjunto de conocimientos que sirven para la fabricación de bienes y la prestación de servicios”. La tecnología se diferencia del conocimiento científico –aunque ambos se encuentren interrelacionados-, pues mientras éste se configura con independencia de sus posibilidades de aplicación productiva, aquella sólo adquiere existencia real en su empleo dentro de las unidades de producción (Molero y Buesa, 1997, p. 135).

Otra definición es la aportada por Ernst Braun (1986): “la creación y utilización de artefactos con fines prácticos. (...) la utilización de herramientas en el sentido más amplio de la palabra”.

El trabajo desarrollado por Molero y Buesa considera que la tecnología puede estar *incorporada* en los equipos de producción y *desincorporada* en la experiencia empresarial, en la organización, en los métodos de trabajo, en el diseño de los productos, en planos, en los documentos de patentes, etc.

Se pueden originar nuevos conocimientos tecnológicos basándose en las rutinas y los conocimientos acumulados después de utilizar la tecnología de forma continuada, ya que genera nuevo conocimiento para mejorarla. El aprendizaje que se obtiene por medio de la práctica se le denomina *learning by doing* (Arrow, 1979). Existen otros tipos de aprendizajes, entre ellos, el *learning by failing* o aprendizaje por error, y es que después de que se ha introducido la innovación en el mercado puede tener resultados adversos, no previstos; sin embargo, de ellos se puede obtener la información necesaria para aprender y mejorar. Otro, es el aprendizaje por el uso o *learning by using*, obtenido después de que la introducción en el mercado ha sido exitosa por medio de la información que proporcionen los consumidores para mejorar los productos.

Transferencia de tecnología: en sentido estricto, supone la transmisión y asimilación de conocimientos a través de las fronteras y elementos materiales que hacen posible la innovación.

BLOQUE 0

INTRODUCCIÓN

He ofendido a Dios y a la humanidad porque mi trabajo no tuvo la calidad que debía haber tenido.

Leonardo Da Vinci

Existe consenso sobre la necesidad de cambiar hacia una economía basada en el conocimiento y la presente crisis presenta la oportunidad para hacerlo. Si creen que el conocimiento es caro, prueben con la ignorancia.

Joan Guinovart

Presidente de la Confederación de Sociedades Científicas de España

ROMÁN.- Inventen, pues, ellos y nosotros nos aprovecharemos de sus invenciones. Pues confío y espero en que estarás convencido, como yo lo estoy, de que la luz eléctrica alumbra aquí tan bien como allí donde se inventó.

SABINO.- Acaso mejor.

Miguel De Unamuno

CAPÍTULO INTRODUCTORIO

La economía global continúa capeando uno de los climas económicos más difíciles que se recuerda en generaciones. Lo que comenzó como una crisis del sistema financiero en Estados Unidos (grave crisis de confianza crediticia e hipotecaria, hipotecas *subprime*, el casi colapso del banco de inversión neoyorkino *Bearn Stearn*, así como un valor del dólar anormalmente bajo e instrumentos financieros como el denominado *Credit Default Swaps*), rápidamente se convirtió en una gran recesión a escala mundial dando lugar a la elevación de los precios de las materias primas, a caídas masivas en el consumo, a la escasez de liquidez en la economía, al fuerte incremento del nivel de desempleo, etc. Todo ello ha repercutido negativamente en el acceso a la financiación de las empresas y de los Gobiernos que no tiene precedentes, conduciendo a crecientes presiones proteccionistas en todo el mundo. Pero, lo que resulta incluso más alarmante, es que esta crisis ha calado hondo en la opinión de la sociedad europea de forma que la incredulidad en la clase política y el sistema capitalista, en general, refuerzan la espiral de desconfianza de los mercados.

Como señala el especulador financiero y filósofo George Soros¹, “El estallido de la crisis económica de 2008 puede fijarse oficialmente en agosto de 2007, cuando los bancos centrales tuvieron que intervenir para proporcionar liquidez al sistema bancario”. Aunque este hecho se señala como inicio de la crisis, se trata más bien de su detonante dado que las razones son más profundas y pueden buscarse en la esencia misma del capitalismo, en la necesidad de lucro inherente a este sistema económico. Según este autor, permitir que el capitalismo opere bajo el libertinaje económico se basa o justifica en el paradigma equivocado de que las fuerzas del mercado tienden a producir equilibrios que a veces se obtienen con desviaciones aleatorias. La libertad económica favorece la innovación y, por tanto, el progreso económico, pero dejar al sistema financiero sin regulación produce efectos destructivos tal y como se ha comprobado: la desregulación financiera y la expansión monetaria excesivas han propiciado una ampliación del crédito, su colocación irresponsable en los mercados y el diseño de productos financieros peligrosos como los *swaps*.

¹ Soros, G., 2008. El nuevo paradigma de los mercados financieros, Ed. Taurus (2008).

BLOQUE 0.- INTRODUCCIÓN

En el informe realizado por la U.N.C.T.A.D (Conferencia de las Naciones Unidas sobre Comercio y Desarrollo), para el caso de los países menos avanzados, se afirma que “la actual crisis financiera se debe a las debilidades del modelo neoliberal que ha inspirado las políticas económicas mundiales durante los tres últimos decenios, debilidades magnificadas por el fracaso de las políticas aplicadas y la laxitud de la legislación de los países avanzados. El costo de las intervenciones y recapitalizaciones de bancos ha alcanzado ya un nivel sin precedentes. Sin embargo, la mayor preocupación actual es el impacto negativo en la economía real y el costo que suponen las pérdidas de producción y empleo. La mayoría de las economías avanzadas está en recesión y los mercados emergentes se han desacelerado” (U.N.C.T.A.D., 2012).

Desde la Unión Europea (UE) y en palabras del ex Comisario Europeo de Ciencia e Investigación Janez Potocnick, nos encontramos con que “la situación económica actual es particularmente difícil. La crisis financiera ha alcanzado proporciones históricas. Queda por ver cómo será de severa la caída y cuánto durará. Pero una cosa está clara: la respuesta política de Europa no sólo debe ser fuerte sino que debe estar coordinada y Europa no debería disminuir su inversión en investigación. Por el contrario, los Estados Miembros deberían esforzarse en mejorar su atractivo como zonas para invertir y preparar sus economías e industrias para aprovechar al máximo la recuperación. Esto hace que las actividades relacionadas con el conocimiento y las reformas estructurales sean aún más importantes que durante las fases de expansión económica”.

En relación a este tipo de inversiones fomentadas por las políticas públicas de la investigación, el desarrollo y la innovación (I+D+i), se ha presenciado en las últimas décadas un fuerte impulso por parte de diversas organizaciones internacionales. En particular, los informes anuales de la OECD (Main Science and Technology Indicators) o del Banco Mundial (Building Knowledge Economies: Advanced Strategies for Development; World Bank Institute Development Studies, 2007) y estrategias como la Agenda de Lisboa o la nueva Estrategia Europa 2020 (COM(2010), EUROPE 2020: A strategy for smart, sustainable and inclusive growth), son un claro ejemplo de este impulso.

Una de las principales consecuencias de una recesión como la actual, es la caída del crecimiento de la economía de los países en términos de porcentaje del Producto Interior Bruto (PIB) o de la renta per capita. En este contexto, las actividades de I+D+i tienen como objetivo principal el apoyo a la generación de conocimiento, que se encuentra en la tecnología

BLOQUE 0.- INTRODUCCIÓN

y en las innovaciones incorporadas al tejido productivo, para la creación de riqueza y, por ende, para la consecución de un incremento de la competitividad y la mejora del bienestar de la sociedad en su conjunto.

La innovación es un proceso que resulta de la interacción de muy diferentes actores y políticas, tratándose así de un concepto complejo, multidimensional, transversal, incierto e imposible de medir de forma directa. Pero, resulta un proceso clave ya que se trata de uno de los principales motores del crecimiento de la productividad y de la renta de los países (Griliches, 1998). El crecimiento económico de un país no depende sólo de inputs tradicionales como la tierra, el trabajo y el capital físico, sino también de la eficiencia con la que es capaz de combinarlos. El premio Nobel Robert Solow ya sentó las bases teóricas de este proceso observando que la acumulación de trabajo y capital sólo explica una parte del crecimiento económico de un país, quedando la parte restante, o el conocido como *residuo de Solow* (Solow, 1956), asociada al progreso tecnológico considerado como un factor exógeno a dicho crecimiento. Posteriormente, los modelos de crecimiento endógeno (Romer 1990, Lucas 1998) se centraron en tratar de comprender los determinantes de este residuo y asentaron la importancia del gasto en I+D para el crecimiento de la productividad de un país.

El importante papel que juegan los recursos dedicados a la I+D+i y el análisis de su uso ha llevado a una revisión de los mecanismos de evaluación empleados para superar la tradicional rendición de cuentas o *accountability*, de forma que se conviertan en instrumentos de apoyo a la toma de decisiones y a las estrategias a nivel nacional: se ha pasado de primar aspectos técnicos, científicos y financieros a abordar los impactos socioeconómicos e indirectos derivados de su aplicación. En estos momentos, aparece como especialmente relevante la rendición de cuentas de la gestión de estos recursos y las partidas dedicadas a esta finalidad en los presupuestos, tanto de la UE como de sus Estados Miembros (EEMM), no son una excepción.

En definitiva, la principal razón que puede esgrimirse a la hora de destacar la importancia de la evaluación de los impactos de la política de I+D+i es su contribución a los objetivos públicos generales, tales como el crecimiento económico, la salud o la educación (Boulanger and Bréchet, T., 2005). Debe justificarse por qué invertir en I+D+i y no en otras políticas alternativas. Adicionalmente, estos procesos no dejan de ir acompañados del carácter incierto de los mencionados impactos al tratarse de un proceso de tipo “caja negra” (en particular, si

BLOQUE 0.- INTRODUCCIÓN

consideramos la inseparable relación entre los procesos de cambio social, la evolución de la tecnología y la adopción de innovaciones).

La globalización de la actividad económica, junto con los problemas de sostenibilidad del Estado del Bienestar en la UE, dificultan el logro de estas metas generales ya que los presupuestos públicos están condicionados, por un lado, por el control y el recorte del gasto público, y por otro, tienen que hacer frente a necesidades de gasto derivadas del desempleo y del envejecimiento de la población. Resulta difícil de esta forma satisfacer la demanda de partidas presupuestarias destinadas a fomentar la I+D+i.

Otra razón para realizar un estudio de la eficiencia de las políticas de I+D+i es que la idea de que la rendición de cuentas de la Administración pública debe ser transparente ha emergido fuertemente en la sociedad. Esta evaluación permite crear una sociedad mejor informada a través de sistemas mejorados de rendición de las cuentas públicas.

La ausencia de estudios relativos a la eficiencia del uso de los recursos de I+D+i a nivel macroeconómico se debe, en gran parte, a la citada dificultad de medir sus impactos socioeconómicos. Este hecho es fruto, a su vez, de que el cambio tecnológico alcanzado a través de las innovaciones es el resultado de un proceso dinámico que implica relaciones tanto a corto como a largo plazo (Hausman, 2003) y de que dichas innovaciones pueden proceder directamente de los recursos empleados en la I+D+i (innovaciones tecnológicas) o no (innovaciones no tecnológicas). El proceso dinámico de la innovación no sólo se refiere a la creación de nuevos productos y al aumento del número de patentes, sino también al establecimiento de un ambiente institucional adecuado a la innovación, a sus políticas reguladoras y a la capacitación de los agentes económicos implicados.

Lo expuesto debe entenderse en el marco de la mencionada recesión global, de una caída del crecimiento demográfico de la sociedad europea y de una política comunitaria que plantea como uno de sus objetivos prioritarios reconducir el modelo económico de crecimiento hacia una “economía basada en el conocimiento” o “economía del conocimiento” (EC) y un modelo más sostenible (DS). Sin embargo, el balance de la llamada Estrategia de Lisboa, que perseguía estas metas, refleja que los países de la UE se han encontrado lejos de alcanzarlos. De hecho, la Agenda de la Estrategia de Lisboa estableció varios objetivos relacionados con la innovación (además del 3% para la ratio de gasto en I+D sobre el PIB) como que el 66% del gasto en I+D debía ser financiado por el sector privado y el 7% del mismo debería dedicarse a

BLOQUE 0.- INTRODUCCIÓN

las tecnologías de la información y las comunicaciones (TICs)². Estos objetivos, al día de hoy, no han sido alcanzados por la mayoría de los países miembros de la UE.

Es claro que los responsables políticos se enfrentan con nuevos desafíos en la gestión de la economía de sus países. En España, por ejemplo, a pesar de la recomendación de invertir en conocimiento para salir de la crisis económica, se escucharon las duras críticas a la decisión del Gobierno de recortar³ la inversión pública en I+D+i para el año 2013 a niveles inferiores a los recursos dedicados en los presupuestos del año 2005. Aunque, puede resultar lógico que la opinión pública entienda como necesarios los recortes en este tipo de programas frente a los recortes en otros programas sociales.

La evaluación de la eficiencia de las políticas públicas sigue siendo fuente de gran debate tanto por los diferentes instrumentos e indicadores existentes para realizarla como por los diferentes intentos de proporcionar esquemas de clasificación de los mismos (Vedung, 1998). A pesar de la relevancia del conocimiento, la innovación y el cambio tecnológico, no se encuentran estudios concretos sobre cómo diseñar políticas de innovación (Faberberg y Verspagen, 2009) debido, en gran parte, a las dificultades que comparte con cualquier otro tipo de política pública y a las peculiaridades de la inversión de recursos en I+D+i.

La revisión de la literatura especializada sobre la evaluación de los resultados alcanzados en las políticas públicas de I+D+i revela que la mayor parte de los estudios en este ámbito utilizan indicadores de evaluación limitados, ya que únicamente ofrecen información cuantitativa sobre la cantidad de recursos humanos y financieros y sobre el nivel de desarrollo alcanzado en el ámbito macroeconómico como consecuencia de las medidas públicas implementadas. En consecuencia, dichos indicadores no se pueden catalogar como indicadores de eficiencia propiamente dichos.

La evaluación de políticas de I+D+i implica que, para que los recursos públicos se traduzcan en incrementos del bienestar de la sociedad, existe la obligación de gastar con eficiencia. Es decir, no se trata sólo de cuantificar el esfuerzo inversor en I+D+i, que en definitiva es un dato, sino que es importante analizar el grado de eficiencia con el que se emplean dichos

² Para la economía española, estos objetivos se fijaron en el 2% del PIB para el gasto total en I+D y el 55% para el porcentaje que financia el sector privado, pues el punto de partida era más bajo. Desde entonces, la posición relativa de la economía española ha mejorado sólo ligeramente.

³ La necesidad de mantener las vías de financiación de la investigación está recibiendo el apoyo de instituciones académicas y de investigación a través de una iniciativa europea respaldada por más de 40 premios Nobel:<http://www.no-cuts-on-research.eu/index.php?file=insert.php>.

BLOQUE 0.- INTRODUCCIÓN

recursos. Además, la capacidad innovadora de un sistema (nacional o regional) no depende sólo de su esfuerzo cuantitativo en I+D+i y de su infraestructura en tecnología, sino también de la interacción entre los distintos agentes del sistema –enfoque holístico- como son las empresas (sector privado) y las Administraciones Públicas (sector público).

De lo anterior se desprende que el primer objetivo general de la presente investigación es la evaluación de la eficiencia de las políticas públicas de I+D+i en la UE. El análisis se realizará de forma global evaluando la eficiencia de las políticas de I+D+i de los países de la UE en varios escenarios socioeconómicos, de forma que el foco de atención se centra en la consecución de un desarrollo sostenible. El método empleado para la evaluación de la eficiencia de estas políticas será el Análisis Envolvente de Datos (o DEA, en sus siglas en inglés), con el que se pretende conocer cuál es la posición relativa que ocupa cada una de las políticas de la UE en función de su eficiencia en cada uno de los varios escenarios planteados. En concreto, se ofrece una aplicación novedosa del DEA en dos pasos encadenados (two linked-steps DEA ó 2LS-DEA).

Ante la ausencia de modelos a nivel macro que apoyen la evaluación de los recursos públicos de I+D+i (objetivo anterior), el segundo objetivo general será construir un modelo conceptual causal que refleje de forma global las variables y relaciones implicadas en el crecimiento económico de los países de la UE (o países del entorno de similares características socioeconómicas) desde la perspectiva de las teorías de crecimiento endógeno, así como de forma holística reflejar los diferentes escenarios de sostenibilidad que influyen en el bienestar de una sociedad y que son comunes a una economía basada en el conocimiento. De esta forma, la evaluación de las políticas públicas de I+D+i (primer objetivo) se realizará en base al modelo causal diseñado (segundo objetivo), necesitando abordar en primer lugar éste último para apoyar la consecución del primero.

Dado que los modelos conceptuales son de gran utilidad en el diseño de políticas públicas pero poco utilizados en el área de las políticas de apoyo y fomento de la I+D+i, este modelo pretende ser una herramienta de utilidad para la toma de decisiones por parte de los responsables políticos por la metodología empleada para su diseño, por el gráfico obtenido (una red Bayesiana) y, conceptualmente, por englobar y enlazar tres facetas de una misma realidad. Las facetas íntimamente interrelacionadas entre sí serían las actividades y recursos de I+D+i, la economía del conocimiento (EC) y el desarrollo sostenible (DS), con aspectos o áreas compartidas y aspectos propios diferenciadores de cada una de ellas. En última instancia,

BLOQUE 0.- INTRODUCCIÓN

se trata de distintos puntos de vista que van del más reducido (I+D+i) hasta el más amplio (DS) de una misma realidad poliédrica que es el crecimiento económico como medio para mejorar el bienestar de una sociedad.

Ambos objetivos serán llevados a cabo como parte de la metodología denominada Expert-based Cooperative Analysis (EbCA) propuesta por Gilbert et al. (2010). Esta metodología consiste en incorporar de forma iterativa el conocimiento experto explícito y el conocimiento implícito que surge del proceso de análisis de los resultados del método concreto (DEA, en nuestro caso) de forma que dicho método incorpora el conocimiento experto (DEA modificado).

Fruto de esas iteraciones, el modelo conceptual inicial diseñado también se ve modificado hasta obtener un modelo mejorado. El modelo conceptual final así obtenido recogerá el conocimiento experto (explícito e implícito) sobre las principales asunciones teóricas subyacentes, las variables más empleadas por la literatura especializada y las relaciones de causalidad entre ellas. Siguiendo la Teoría de la Causalidad de Pearl (2009), este modelo cumple la condición necesaria y suficiente para ser considerado una red Bayesiana.

En la literatura existente, no se ha encontrado un marco metodológico de análisis causal que haya abordado el estudio de la eficiencia técnica relativa de este tipo de políticas a nivel europeo y de los factores que contribuyen a alcanzarla. Se espera así proporcionar un soporte teórico sólido para el diseño de los casos de estudio propuestos, sustentando de esta forma todo el bloque experimental.

Tras analizar la eficiencia de estas políticas desde el enfoque global de un DS, el tercer objetivo general de esta investigación consiste en, empleando metodologías alternativas a las existentes en la literatura académica y a las propuestas por organismos oficiales de reconocido prestigio internacional, analizar el uso (en su concepción más amplia) de los recursos dedicados a I+D+i. Se analizarán desde las dos facetas ya comentadas que restan (la I+D+i y la EC), evaluando o monitorizando los resultados de los países seleccionados a través de la clasificación de éstos en grupos homogéneos, según los patrones de comportamiento que se detecten, empleando para ello algoritmos de agrupamiento. Posteriormente, se utilizarán técnicas de clasificación nominal y ordinal pertenecientes al campo de la Inteligencia Artificial y del Aprendizaje Automático para la construcción de modelos capaces de predecir la clasificación de los países en cada una de las clases o grupos obtenidos previamente mediante las técnicas de agrupamiento.

BLOQUE 0.- INTRODUCCIÓN

Como ya se ha adelantado, el modelo diseñado se usará como base en el bloque experimental para la selección de los indicadores de I+D+i más adecuados en cada caso –recursos financieros y humanos, publicaciones científicas, patentes, innovaciones no tecnológicas, etc. Junto al empleo de estos recursos, se seleccionarán variables comúnmente utilizadas para la caracterización de un país según su estado de transición hacia una economía basada en el conocimiento (EC) y/o un desarrollo sostenible (DS). De esta forma, el análisis de los tres conceptos tan íntimamente interrelacionados no sólo permitirá el análisis del empleo, eficiencia, ejecución, resultados o monitorización de los recursos de I+D+i por parte de los países de la UE y del entorno socioeconómico más cercano, sino también su clasificación en función de la consecución de una EC y de un DS en la que las citadas actividades y recursos de I+D+i son una pieza fundamental.

A partir de la amplia revisión bibliográfica realizada, ha sido posible establecer múltiples hipótesis de partida acerca de los objetivos generales planteados. Estas hipótesis se exponen a continuación y se resumen y se relacionan en la tabla 0.1:

A) Variables que influyen en la eficiencia de las políticas públicas de I+D+i en el marco de un desarrollo sostenible.

Hipótesis 1 (H1): Es posible identificar las variables más relevantes y sus relaciones de causalidad para evaluar la eficiencia del uso de los recursos de I+D+i, en base a la literatura existente y la opinión de los expertos permitiendo, así mismo, la construcción de un modelo conceptual causal de red Bayesiana. Este modelo puede recoger las ideas de tres conceptos muy relacionados: los propios recursos de I+D+i, la EC y el DS.

Hipótesis 2 (H2): Se pueden detectar las variables más relevantes en función de la metodología empleada, lo que permitirá conocer las que ejercen una mayor influencia en la mejora de los resultados de eficiencia y de la clasificación de los países según las tres dimensiones consideradas. Estos aspectos podrían ser de gran utilidad para los decisores políticos y los grupos de interés, en general.

B) Interrelación entre los conceptos de I+D+i, economía basada en el conocimiento y desarrollo sostenible.

Hipótesis 3 (H3): Este problema puede ser abordado desde distintas perspectivas, partiendo de una misma realidad poliédrica como es el bienestar último de una sociedad (en este caso, europea) y dada la interrelación entre los tres conceptos.

C) *Relación entre la eficiencia en el escenario Tradicional (recursos y resultados de la actividad científica) y la eficiencia que presentan el resto de escenarios socioeconómicos analizados.*

Hipótesis 4 (H4): El fin último de este tipo de políticas no es la producción científica bibliográfica ni las patentes *per se*, sino el alcanzar un mayor crecimiento económico que, dentro de un modelo de desarrollo sostenible, conduzca al bienestar de la sociedad en cuestión. Por tanto, si un país presenta una política de I+D+i eficiente, también debería serlo en el resto de escenarios socioeconómicos si se quiere que la dedicación de los recursos a este tipo de actividades cumpla su fin último.

D) *Relación entre el valor del gasto en I+D+i y la proporción de recursos humanos dedicados a estas actividades y la eficiencia alcanzada en el escenario Tradicional y en el escenario de Educación.*

Hipótesis 5 (H5): los países que invierten mayores recursos públicos en I+D+i, son los que cuentan con políticas más eficientes en el escenario Tradicional.

Hipótesis 6 (H6): Los países con sistemas educativos conocidos por su gran calidad (p.e., Finlandia), son los más eficientes en el resto de escenarios, especialmente en el Tradicional y en el de Innovación.

E) *Relación entre la eficiencia de las políticas de I+D+i de los países y su localización geográfica.*

Hipótesis 7 (H7): Podría pensarse que los países del Norte de Europa (los países nórdicos), que presentan una mayor tradición de inversión en recursos de I+D+i, que cuentan con un clima especial (menores horas de luz del día, bajas temperaturas, etc.) y una cultura que favorece el estudio y la innovación, serán más eficientes en los distintos escenarios que el resto de países mediterráneos y del Este de Europa.

F) *Países más eficientes.*

Hipótesis 8 (H8): Los países que muestran mejores resultados en relación al uso de los recursos de I+D+i, en general y en el marco de un DS, así como en su transición hacia un modelo de EC, pueden ser los mismos en los diversos casos de estudio propuestos dada la interrelación entre los conceptos. Si esta hipótesis se cumple, estos países podrían utilizarse como referencia a la hora de detectar mejores prácticas entre países y servir de referencia, con la lógica adaptación a las peculiaridades de cada nación (*benchmarking*).

BLOQUE 0.- INTRODUCCIÓN

G) *Presencia de características comunes en los países analizados.*

Hipótesis 9 (H9): Se pueden detectar patrones homogéneos de comportamiento de los países seleccionados a la hora de tratar de clasificarlos en grupos o clases según su ejecución en I+D+i, la EC y el DS y mediante las técnicas de agrupamiento adecuadas. Una vez establecidas las clases, se pueden construir modelos que pueden predecir dicha clasificación con un alto grado de precisión.

H) *Concepto de evaluación de la eficiencia empleado.*

Hipótesis 10 (H10): El concepto de evaluación de la eficiencia puede ser empleado, para el propósito de evaluar la eficiencia de las políticas públicas de I+D+i, desde su sentido más amplio, que va desde el método más específico de evaluación de la eficiencia técnica relativa (DEA) hasta la clasificación de las políticas/países para la monitorización, seguimiento y evaluación de los mismos.

I) *Tipo de problema de clasificación: nominal u ordinal. Técnicas de clasificación de los campos de la Inteligencia Artificial y del Aprendizaje Automático.*

Hipótesis 11 (H11): Con los datos disponibles, es posible abordar el estudio de la clasificación de los países mediante técnicas de clasificación nominales y ordinales de los campos de estudio mencionados. La naturaleza de los problemas prácticos planteados no está determinada *a priori* por lo que se estudiará la naturaleza nominal u ordinal de las clases obtenidas en cada caso, lo que puede ser de gran utilidad para la monitorización y seguimiento por los principales grupos de interés.

Tabla 0.1 Relación entre objetivos e hipótesis planteados en la investigación

OBJETIVOS PLANTEADOS	HIPÓTESIS RELACIONADAS
OBJETIVO 1 (EVALUACIÓN EFICIENCIA RECURSOS PÚBLICOS I+D+I)	H2 , H4, H5, H6, H7 H8, H10
OBJETIVO 2 (MODELO CONCEPTUAL CAUSAL DE APOYO; RED BAYESIANA)	H1, H3
OBJETIVO 3 (TÉCNICAS ALTERNATIVAS DE CLASIFICACIÓN PARA DISTINTAS PERSPECTIVAS)	H2, H8, H9, H10, H11

Respecto a la metodología concreta empleada, la clasificación puede realizarse en cualquier ámbito mediante diversas técnicas. Como se ha ido adelantando, en este trabajo las técnicas

utilizadas fueron, por un lado, la aproximación no paramétrica de programación matemática de análisis de la eficiencia técnica relativa conocida como Análisis Envoltente de Datos (DEA) para la obtención de “rankings” de eficiencia, en el marco de la metodología EbCA (Gibert et al., 2012). Por otro lado, se han empleado técnicas procedentes del campo de la Inteligencia Artificial que permitirán agrupar los países según las características comunes que presenten (algoritmos de agrupamiento) y predecir el grupo o clase al que pertenecen (técnicas de clasificación nominal y ordinal).

En relación al primer método, los métodos envolventes de datos utilizan la programación lineal para estimar de forma relativa –comparando las fronteras de las mejores prácticas– mediante la computación de los valores de los inputs y los outputs (I/O) sin una asunción previa sobre su forma funcional subyacente (modelos no paramétricos).

Los modelos DEA presentan numerosas ventajas en el análisis de la eficiencia técnica relativa (ETR): i) se pueden ejecutar con I/O de diferente naturaleza, ii) no es necesario probar una función que relacione los parámetros y iii) no depende de la escala del análisis. Sin embargo, también tienen importantes deficiencias: i) es necesario conocer los valores de los I/O, ii) el número de unidades a evaluar depende del número disponible de I/O y iii) los I/O son difícilmente interpretables de acuerdo con el conocimiento experto disponible.

En situaciones inciertas, donde los valores de los I/O son variables aleatorias, los modelos DEA tradicionales no son viables. En estos casos existen aproximaciones como Imprecise-DEA, pero que no son de utilidad en problemas completamente estocásticos donde sólo es posible resolver el problema mediante la hibridación de los modelos DEA con un motor de simulación: Monte-Carlo DEA (García-Alonso, 2009).

El planteamiento del presente trabajo de investigación de ofrecer metodologías variadas y adicionales a las empleadas en este campo, contempló la conveniencia de aplicar técnicas de clasificación utilizadas en el campo de la Inteligencia Artificial. En principio, la búsqueda de modelos adecuados para clasificación fue un campo prácticamente exclusivo de la Estadística, pero la necesidad de modelos más avanzados fue reconocida por varios investigadores pioneros (Rosenblatt (1956), Whittaker, (1922)) sugiriendo modelos de mayor complejidad que podrían haber superado a los modelos lineales, pero que, dada la potencia computacional disponible en la época, no pudieron ser implementados. Hoy en día, los recursos computacionales de que disponemos sí que permiten su utilización.

BLOQUE 0.- INTRODUCCIÓN

Al mismo tiempo, cada día es más importante el análisis de los datos y la extracción de conocimiento de los mismos, cobrando protagonismo en un amplio conjunto de áreas (Hastie, 1996). Los conceptos de clasificación y regresión actualmente se engloban en una variada colección de términos como Reconocimiento de Patrones (Pattern Recognition), Análisis Inteligente de Datos (Intelligent Data Analysis), Descubrimiento de Conocimiento (Knowledge Discovery), Aprendizaje Automático (Machine Learning) o Minería de Datos (Data Mining). En este contexto, la clasificación consiste en el establecimiento de una relación (normalmente, utilizando un modelo funcional) entre una variable categórica de interés y un conjunto de variables observadas que se espera que puedan explicarla. Ejemplos de este tipo de metodologías son las redes neuronales artificiales (RNAs) (Bishop, 2006; Haykin, 2008), las Máquinas de Vectores de Soporte (Support Vector Machine, SVM) (Vapnik, 1999) o los árboles de clasificación (Witten, 2005).

Otra idea habitualmente utilizada es la de añadir restricciones adicionales, asociadas al orden, a los clasificadores nominales, lo que se denomina clasificación o regresión ordinal. De entre las diferentes metodologías existentes en clasificación ordinal, los métodos de umbral son los más utilizados en la actualidad. Dentro de estos métodos y basándose en máquinas de soporte vectorial, Shashua y Levin (2003) han propuesto dos métodos: el primero basado en la maximización del margen entre las clases vecinas más próximas y el segundo basado en la maximización de la suma del margen entre todos los pares de clases. Chu y Keerthi (2005) también han propuesto dos métodos de regresión ordinal basados en las ideas de Shashua y Levin. Otra idea alternativa es la reducción del marco de trabajo ordinal a clasificación binaria basada en ejemplos o patrones extendidos con nuevas variables (Lin y Li, 2007). Otro ejemplo perteneciente al ámbito de la Estadística es el modelo Proportional Odds Model (POM) (McCullagh, 1989).

A partir de los objetivos e hipótesis planteados, la presente investigación se estructura de la siguiente forma (ver figura 1.1):

BLOQUE 0. INTRODUCCIÓN. En primer lugar, sirva este Bloque introductorio como presentación de la relevancia del tema de estudio y para la formulación de las hipótesis de partida que se tratarán de contrastar a lo largo de esta investigación, en base a la serie de objetivos propuestos.

BLOQUE 1. MARCO HISTÓRICO Y CONCEPTUAL. La investigación, el desarrollo tecnológico y la innovación son cruciales para el desarrollo y el bienestar de la sociedad, fomentando el cambio tecnológico como principal factor de crecimiento económico. Esta afirmación es comúnmente aceptada en la literatura económica y, por ello, en este bloque se realizará un breve recorrido por las principales teorías del crecimiento económico que consideran como elemento central al cambio tecnológico, al conocimiento o a la innovación. El capítulo I (“Cambio tecnológico, la innovación y el conocimiento en el pensamiento económico”) tiene como objetivo crear el adecuado marco teórico en el que se basará el diseño del modelo conceptual causal. Se trata de una revisión que no pretende conseguir valor añadido en la investigación científica ni ser exhaustiva, sino sacar a la luz algunas de las características y peculiaridades que conforman este tipo de análisis macroeconómico.

En el capítulo II, “Hacia una economía basada en el conocimiento y un modelo de desarrollo sostenible. El papel de la I+D+i”, se comentan brevemente estos conceptos y se justifica la íntima relación entre los mismos, lo que constituye uno de los hilos conductores de la investigación. En el capítulo III, titulado “Seguimiento y evaluación de políticas públicas. Los recursos públicos de la I+D+i y la política científica y tecnológica en la UE”, se destaca la importancia de evaluar políticas públicas en general y, en particular, las que dedican recursos a la I+D+i, objeto de estudio de esta investigación. Dado que se estudiarán principalmente los países de la UE, se presenta una breve panorámica histórica, institucional y normativa de la política científica y tecnológica de la UE, que afecta de forma directa a las respectivas políticas nacionales.

BLOQUE 2. MODELO CONCEPTUAL CAUSAL (RED BAYESIANA). Se trata del bloque donde se presenta el modelo conceptual causal. Tras revisar el concepto de modelo conceptual, la causalidad y las redes Bayesianas, en general (capítulo IV), se presenta la metodología para la elaboración y mejora de su diseño denominada Expert based Cooperative Analysis o EbCA (capítulo V). La estructura del modelo formado por variables (nodos) y relaciones de causalidad (flechas) es explicada con detalle en el capítulo VI. Además, se señalan las principales ideas teóricas que forman parte de la base de conocimiento experto necesaria para el diseño de este modelo.

BLOQUE 3. ANÁLISIS EMPÍRICO. En este bloque experimental se plantean tres análisis empíricos para dar respuesta al objetivo de estudiar la eficiencia del uso de los recursos de I+D+i desde distintas facetas o perspectivas de un mismo fenómeno, interrelacionadas entre

BLOQUE 0.- INTRODUCCIÓN

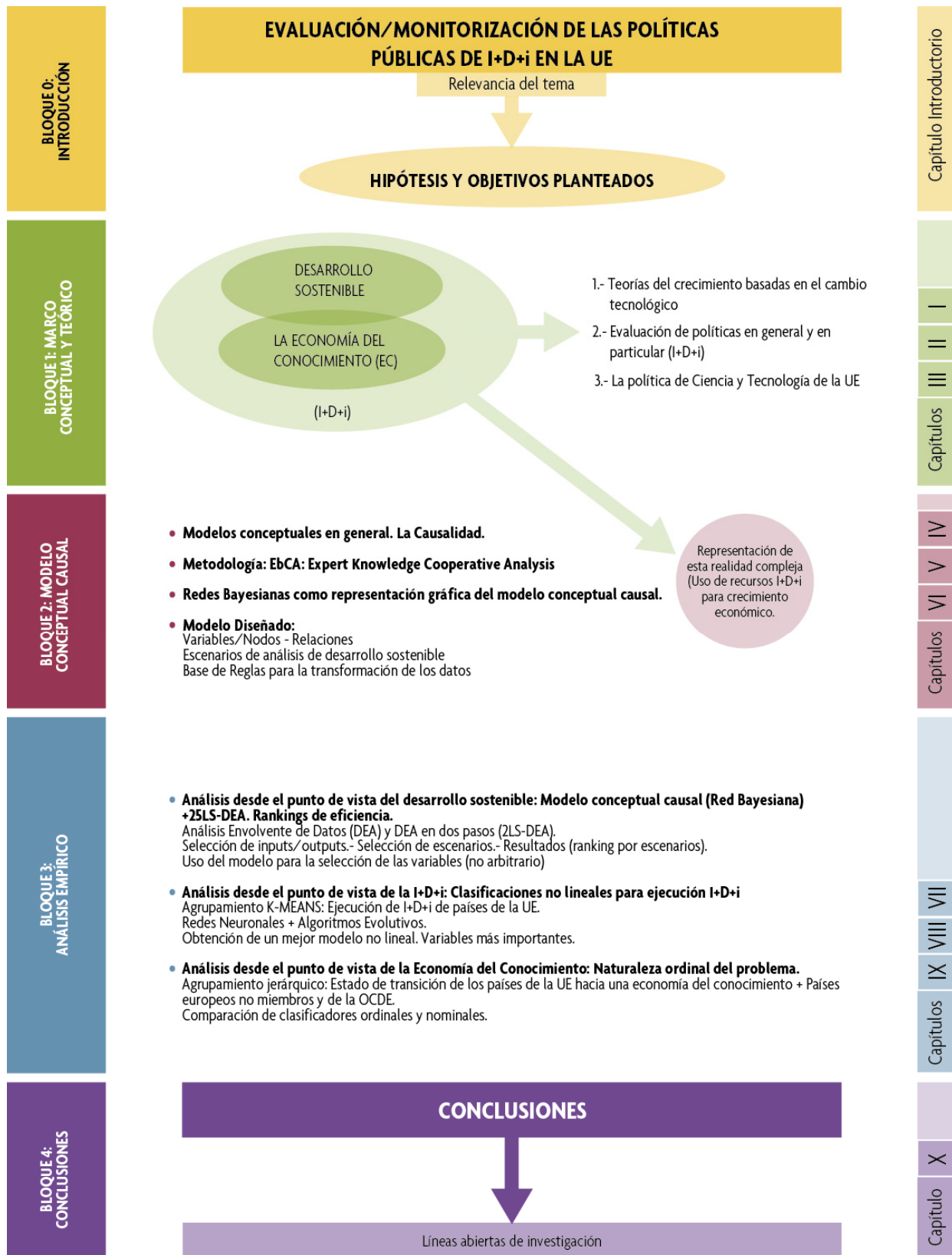
sí, de forma alternativa y complementaria a las metodologías ya existentes. En cada uno de ellos, se seguirá una estructura común: breve presentación de la relevancia del tema, revisión del estado del arte de la metodología relacionada y empleada, así como resultados y conclusiones (en este caso, parciales) de cada uno de los casos planteados.

El primero de los análisis empíricos (capítulo VII), estudiará la eficiencia de las políticas públicas de I+D+i en la Unión Europea desde el prisma de la consecución de un desarrollo sostenible analizando la eficiencia a través del Análisis Envolvente de Datos en dos pasos enlazados (two linked-steps Data Envelopment Analysis o 2LS-DEA). Para ello, se empleó también la metodología EbCA mediante la que se obtiene un método “guiado” gracias a la incorporación del conocimiento experto (explícito e implícito). Además de para el estudio de dicha eficiencia, sirvió al propósito de mejorar el diseño del modelo conceptual causal descrito someramente en los capítulos V y VI del Bloque anterior de forma iterativa. Por lo tanto, el modelo es considerado también un *resultado* del primer análisis, además de emplearse como base conceptual y teórica para los análisis empíricos aquí presentados.

El segundo análisis planteado (capítulo VIII) se centra en los recursos totales dedicados a la I+D+i, relacionándolos con los resultados que se consideran más directamente relacionados con esta clase de recursos. En el tercer caso (capítulo IX), enfocamos el planteamiento desde el punto de vista de lo que caracteriza a una economía del conocimiento por parte de los países europeos y de otros países adicionales del entorno socioeconómico de interés. Respecto a la metodología empleada en el segundo y tercer análisis, se escogieron diversas técnicas pertenecientes a los campos de la Inteligencia Artificial (IA) y del Aprendizaje Automático (tanto para clasificar los países en grupos según el punto de vista empleado de esta compleja realidad como para la obtención de clasificadores/modelos nominales y ordinales que puedan predecir dicha clasificación en el grupo correspondiente, según la naturaleza del problema).

BLOQUE 4. DISCUSIÓN, CONCLUSIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN. En este último bloque, se dará paso a las respuestas que la tesis aporta a las distintas hipótesis y objetivos planteados en la presente introducción, así como las implicaciones y el amplio abanico de líneas futuras de investigación que se abre a raíz de la presente.

Figura 1.1. Esquema del trabajo de investigación



Fuente: Elaboración propia.

BLOQUE 1

MARCO TEÓRICO Y CONCEPTUAL

La productividad no lo es todo, pero a largo plazo lo es casi todo.

La capacidad de un país para mejorar el nivel de vida depende casi totalmente de su habilidad para aumentar el producto por trabajador.

Paul Krugman (1992), *The Age of Diminished Expectations*

"La ciencia de hoy parece estar atrapada en un fuego cruzado entre dos visiones opuestas del mundo. Por una parte, la ciencia es la principal herramienta de la ideología que actualmente dirige la economía mundial denominada sistema de libre mercado, orientada al crecimiento continuo y la búsqueda de riqueza individual. Por otra, la ciencia está llamada, de manera creciente, a producir conocimiento y tecnología que promueva la sustentabilidad ambiental, el desarrollo orientado hacia los pueblos y el manejo de largo plazo de los recursos"

Alberta Kananaskis Village (1998).

The Role of Science and Technology in Society and Governance.

CAPÍTULO I. CAMBIO TECNOLÓGICO, INNOVACIÓN Y CAPITAL HUMANO EN EL PENSAMIENTO ECONÓMICO

1. EL CAMBIO TECNOLÓGICO Y LA INNOVACIÓN EN LAS TEORÍAS DEL CRECIMIENTO ECONÓMICO

En los últimos años, se ha visto acrecentado el interés por las cuestiones relacionadas con los impactos de las actividades relacionadas con la ciencia y la tecnología, así como por las posibles actuaciones del sector público en su promoción y fomento. Ya sea como fuente de crecimiento económico, de competitividad o, más genéricamente, como mecanismo para incrementar el bienestar social, se ha destacado la influencia que los aspectos relacionados con la Investigación, el Desarrollo y la Innovación (I+D+i) poseen sobre las variables económicas. Este fenómeno es propio de los países desarrollados y, sin duda, guarda íntima relación con el auge del fenómeno de la globalización.

Desde sus inicios, se encuentra como reto común en las teorías que tratan de explicar el crecimiento económico el encontrar los factores que expliquen la evidencia empírica de que los países presentan distintas tasas de crecimiento a lo largo del tiempo. Las distintas formas de dar explicación a este fenómeno han dado lugar a otras tantas teorías en la Historia del pensamiento económico.

Así, alrededor del papel de la tecnología y el cambio tecnológico se construyeron cuerpos de teorías del crecimiento económico (los neoclásicos) que consideraban tal cambio tecnológico como una variable exógena en sus funciones de producción (crecimiento económico exógeno) o como una variable endógena (crecimiento económico endógeno), entre otras. Nuevos campos de investigación han ido posteriormente emergiendo en un esfuerzo por descubrir los factores que influyen en el desarrollo económico y por ofrecer consejo a los responsables de la toma de decisiones en el gobierno y la empresa (Fagerberg et al., 2012).

En relación a los aspectos sociales y económicos de la ciencia, tanto la innovación como el conocimiento no son nuevos. Ya en el último cuarto del siglo XVIII, Adam Smith (1776) destacó la importancia de estos fenómenos. Más próximo a nuestros tiempos, durante la

primera mitad del siglo XX, académicos como Schumpeter (1912) aportaron importantes bloques teóricos. Sin embargo, a pesar de estas contribuciones pioneras, una base de conocimiento más sólida sólo comenzó a desarrollarse después de la Segunda Guerra Mundial (en particular, desde los años 60 en adelante). Desde entonces, en Estados Unidos se asumió que la ciencia y la tecnología podían satisfacer las necesidades de la defensa nacional, del crecimiento económico y de la mejora de las condiciones de vida de los ciudadanos (Mukerji, 1989). Los decisores políticos, los científicos y tecnólogos, los agentes económicos y una parte considerable de la opinión pública, tanto norteamericana como de los países de la Europa Occidental, adoptaron la siguiente idea: si la ciencia y la tecnología habían contribuido decisivamente al triunfo de las tropas aliadas, también podrían ser factores determinantes en la competencia económica internacional y en la lucha política entre los modelos capitalista y comunista (Luján y Moreno, 1996).

Tras la importancia que adquirió el cambio tecnológico, se añadieron factores intangibles como el conocimiento, el aprendizaje, la educación, la salud, etc., hasta aparecer un nuevo factor aglutinador de los anteriores, el conocido como “capital humano”. Las investigaciones en las dos últimas décadas estudian los determinantes del crecimiento añadiendo causas aún más particulares como la participación del Estado, la situación política y social, etc.

El objetivo de este capítulo es realizar un breve repaso de las contribuciones teóricas que tratan de explicar la importancia del progreso tecnológico, las innovaciones y el conocimiento en el crecimiento económico como causas fundamentales de éste, prestando mayor atención a los autores o teorías en las que se basó el diseño del modelo conceptual causal.

1.1. LA TEORÍA CLÁSICA DEL CRECIMIENTO ECONÓMICO.

En términos generales, dentro de este ámbito se suelen incluir un amplio conjunto de aportaciones que comprenden desde los escritos de la Escuela Clásica —especialmente A. Smith, T. R. Malthus y D. Ricardo— hasta las aportaciones de J. M. Keynes y la de J. A. Schumpeter.

En concreto, A. Smith (1776) afirma que la riqueza de las naciones depende esencialmente de dos factores: de la distribución que se realice del factor trabajo entre las actividades productivas e improductivas y del grado de eficacia de la actividad productiva (progreso técnico). Estos dos factores se ven influenciados por otros que tienen gran importancia en el

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

proceso: la división del trabajo (especialización), la tendencia al intercambio, el tamaño de los mercados (uso del dinero y el comercio internacional) y, finalmente, la acumulación de capital, que en última instancia se considera el elemento esencial que favorece el crecimiento de un país.

Ahora bien, a pesar de ello, Smith señala que todas las naciones, gracias al desarrollo económico que experimenten, alcanzarán un estado estacionario, ya que las oportunidades de inversión se van agotando y con ello el crecimiento. La llegada a este estado estacionario sólo puede retrasarse con la apertura de nuevos mercados y con la aparición de innovaciones que creen nuevas posibilidades de inversión.

Por su parte, Ricardo (1817) también se referiría al estado estacionario, indicando que se alcanzaría por la existencia de los rendimientos decrecientes y que se podría evitar a través de un aumento del capital y de la implantación del progreso técnico.

Para T. R. Malthus (1820), el crecimiento económico necesita una demanda adicional. Pero, no es suficiente una mayor inversión sino que es conveniente que esa mayor demanda venga acompañada de un incremento de la oferta. De acuerdo con Malthus, la causa de que se llegue al estado estacionario está en la propia dinámica de la población que crea rendimientos decrecientes.

Por lo que se refiere a la importante aportación de Keynes a la teoría del crecimiento económico, son dos las variables a considerar. Por un lado, concede gran importancia al *animal spirit* o espíritu empresarial ya que afecta a las decisiones de inversión y a través de ésta al crecimiento; y por otro lado, al ahorro, ya que va a afectar a la riqueza en función de lo que haga el individuo con él (Keynes, 1936).

Para finalizar con las aportaciones que hemos denominado clásicas, vamos a referirnos a una de las que ha servido de base a este trabajo: la de J. A. Schumpeter (1912). En su modelo, las innovaciones son las que principalmente van a propiciar el crecimiento económico ya que a través de ellas se produce la acumulación. En este proceso, el empresario desempeña un papel esencial al ser quien va a introducir dichas innovaciones en el proceso productivo. En su teoría del crecimiento, la ciencia y la tecnología juegan un papel preponderante.

En concreto, Schumpeter supone que la economía puede encontrarse en dos posibles fases o estados. El primero de ellos es el estado estacionario, donde la economía no crece, y que se caracteriza por presentar un determinado estado tecnológico y por la repetición de los mismos

procesos productivos. La segunda fase, que sería la del crecimiento (“destrucción creativa”), se alcanza gracias a la introducción de ciertos cambios —que denomina innovaciones— en el proceso productivo. Si resultan adecuados, generarán mayores beneficios a la empresa innovadora, lo que animará a las empresas competidoras a introducirlos también en sus procesos productivos con el fin de apropiarse de parte de esos beneficios. Para ello, será necesario aumentar la inversión. El resultado del proceso de incorporación de innovaciones es que durante esta fase la economía experimenta un crecimiento positivo. Cuando todos los agentes han incorporado la innovación en sus respectivos procesos productivos, la inversión se detiene y la economía entra de nuevo en una fase de estado estacionario, al menos hasta que alguien introduzca otra innovación, siendo el denominado empresario innovador el encargado de llevar a cabo esta tarea. Tras dicha actuación, sus rivales tenderán a imitarlo (y así sucesivamente).

1.2. NEOCLÁSICOS (TEORÍAS DE CRECIMIENTO EXÓGENO Y ENDÓGENO)

La teoría moderna del crecimiento económico nace a finales de la década de los 60, con el trabajo seminal de Robert Solow, economista del MIT, premio Nobel de Economía y fundador de lo que hoy se conoce como el «modelo de crecimiento neoclásico». Solow dejó de lado el papel de la tierra como factor de producción dado que con la Revolución Industrial la agricultura había dejado de ser la principal actividad económica de los países más desarrollados. Como veremos a continuación, el desarrollo de la teoría moderna del crecimiento económico se produjo en dos etapas. En un primer momento, surgen trabajos después del estudio de Robert Solow donde se asume que la fuente del crecimiento de las naciones es exógena (es decir, no se explica cuál es el mecanismo interior de la economía que influye en el crecimiento, como parte de las decisiones de las empresas y personas). El segundo momento aparece a partir del estudio de Paul Romer (1990) y se centra en el efecto de la innovación sobre el crecimiento de una economía de una forma endógena.

1.2.1. ROBERT M. SOLOW

El trabajo más importante en la teoría neoclásica moderna es, sin duda, *A Contribution to the Theory of Economic Growth* de Robert Solow (1956). Surgió como una crítica al modelo dominante de Harrod-Domar keynesiano y sigue siendo el punto de partida de la discusión

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

más reciente. Cuestiona “la poderosa conclusión” de que en el largo plazo el sistema económico está, en el mejor de los casos, balanceado sobre el “filo de la navaja” en el crecimiento de equilibrio. Solow desecha las proporciones fijas que consisten en que el producto (Y) es proporcional al capital invertido (K); pero también que la inversión es proporcional al ahorro (rendimientos constantes) y aplica la ley de proporciones variables (rendimientos decrecientes).

El modelo de Harrod (1939) y de Domar (1947), criticado por Solow, relaciona el crecimiento con el ahorro, cuanto más ahorro mayor inversión y aumento de la capacidad productiva, del ingreso y, por tanto, creación de más ahorro. El modelo se basa en esta relación ahorro-inversión-ingreso-productividad del capital, es decir en la acumulación de capital y sin duda los resultados a corto plazo son crecimiento económico, empleo y rendimientos crecientes. Pero, de acuerdo al análisis neoclásico y también clásico -Ricardo, Malthus- no alcanzará una tasa de crecimiento más alto de manera permanente debido a los rendimientos decrecientes. Solow argumenta que la tasa a largo plazo del crecimiento del producto por trabajador (Y/L) depende completamente de la tasa de progreso tecnológico en sentido amplio y, por lo tanto, es independiente de la tasa de ahorro (de la tasa de inversión).

El modelo de Solow acepta todos los supuestos de Harrod-Domar con excepción de las proporciones fijas. Solow explica, en su discurso de aceptación del Premio Nobel de Economía, que él se “enredó con la teoría del crecimiento económico” al considerar que los supuestos de Harrod-Domar eran equivocados y la “receta” del modelo era inverosímil (Solow, 1987).

Para mejorar el modelo, Solow reemplazó las relaciones capital-producto (K/Y) y trabajo-producto (L/Y) constantes por una “más rica y más realista representación de la tecnología” y determinó que la tasa de crecimiento de producto por unidad de trabajo (Y/L) depende completamente de la tasa de progreso tecnológico, en el sentido más amplio.

El segundo gran trabajo de Solow, *Technical Change and the Aggregate Production Function* (1957), también se convirtió en la base de la conformación de la corriente neoclásica moderna. Partiendo de la función de producción Cobb-Douglas⁴, encuentra para Estados Unidos que el crecimiento del producto por hora-hombre (Y/L) se duplicó en cuarenta años. Lo más

⁴El modelo de crecimiento de Solow parte de una función Cobb-Douglas de producción de la forma $Q = A \cdot T^\alpha \cdot K^{1-\alpha}$, donde Q es la producción total en un año, T el factor trabajo, K el factor capital, A el factor de productividad total y α es la elasticidad del capital.

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

sorprendente de los resultados fue que una octava parte del aumento del producto era imputable al incremento del capital (o intensidad del capital) mientras que las siete octavas partes se debían al cambio tecnológico, es decir, el hecho de cuantificar que la inversión en tecnología era más decisiva para el crecimiento que los otros factores; además, la tecnología contrarrestaba los rendimientos decrecientes que llevaban a un estado estacionario.

Posteriormente, Edward Denison (1985) fue el primero en desglosar los determinantes del crecimiento en sus partes constitutivas, en particular el progreso técnico; en su trabajo sobre Estados Unidos (1929-1982) encuentra los siguientes determinantes básicos: el nivel de educación en el trabajo, las calificaciones educativas del trabajador promedio, el capital, la asignación mejorada de los recursos (el movimiento del trabajo de la agricultura de baja productividad a la industria de alta productividad), economías de escala y el crecimiento del conocimiento o progreso tecnológico en su sentido más restrictivo. El 12% de la productividad lo aportó el capital y el 88% el “cambio técnico en sentido amplio”, desglosado en los restantes cinco componentes.

El descubrimiento de que el cambio técnico era el determinante del producto por unidad de fuerza de trabajo (Y/L) fue importante porque nunca antes se había medido el crecimiento a partir de sus componentes. Lo normal consistía en considerar todos los factores productivos como causa del crecimiento. Con el hallazgo, se descartaban relativamente algunos factores y se consideraba al progreso técnico como el determinante para el crecimiento. Aunque, al principio, se llegó al absurdo de minimizar la acumulación de capital (no estando representado por algo físico), reconoce que sin la tasa de inversión el progreso técnico hubiera sido mucho menor o se hubiera reducido hasta anularse y que las innovaciones deben concretarse en nuevas plantas e instalaciones. También afirma que la nueva tecnología puede introducirse en el proceso de producción únicamente a través de inversiones en nuevas obras y equipo industrial (Solow, 1957).

El cambio tecnológico resultó ser un fenómeno “inesperado y raro”, considerándose exógeno o fuera del modelo y del sistema económico, donde el mercado no tiene el control (Ehrlich, 1990). También se le llama “residual” porque no es claramente “observable”, no es “explicado”, comparado con la evidente participación cuantitativa del capital y la mano de obra (se mide la participación de los factores materiales y la diferencia con el 100% es considerado el residuo).

En su tercer ensayo, *Technical Progress, Capital Formation, and Economic Growth*, Solow (1962)

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

afirma que la inversión es una condición necesaria pero no suficiente para el crecimiento e incorpora a su perspectiva la importancia en el proceso del crecimiento económico de actividades como la investigación, la educación y la salud pública. Sin embargo, estima que aún se encuentra lejos de obtener estimaciones cuantitativas de la rentabilidad social de los recursos dedicados a tales actividades (Solow, 1962). Se abrían nuevos campos de investigación para incorporar factores del crecimiento distintos, que se sumarían a su aportación original.

Los estudios de Solow y de sus contemporáneos demostraron de forma cuantitativa la relevancia de la acumulación del capital y del progreso tecnológico en la productividad del trabajo. Pero, sobre todo, supieron ver en el progreso tecnológico, la calificación de la mano de obra y las innovaciones a los principales motores del crecimiento económico. El trabajo de Solow de 1957 y algunos estudios posteriores demostraron que en países capitalistas avanzados el cambio tecnológico contrarresta los rendimientos decrecientes, obteniendo más producción, aún con la misma cantidad de insumos. Factores como la invención, la calificación de la mano de obra y de los procesos productivos, ayudan a contrarrestar la tendencia decreciente de los beneficios.

Es precisamente a Solow a quien se le reconoce la cuantificación del progreso tecnológico (por ello se le conoce como “residuo de Solow”) y en base a sus trabajos ha girado la teoría neoclásica y otras corrientes.

1.2.2. PAUL M. ROMER

En las décadas de los 50 y los 60, el modelo de crecimiento económico básico para el análisis de la realidad de los países industriales era el de Harrod-Domar (Barro et al., 1997). A finales de los 60, se exponía el modelo de Solow, el cual prevaleció en las siguientes dos décadas (Barro et al., 1997). En los noventa, se seguía dicho modelo, a diferencia del de Harrod-Domar y los keynesianos que desaparecieron para dar paso a una nueva interpretación: la teoría endogenista o también conocida como la “nueva teoría del crecimiento” o el “nuevo paradigma” (para una revisión, ver Acemoglu, 2007).

Existe consenso en que esta nueva corriente endogenista la iniciaron Paul M. Romer y Robert E. Lucas Jr. en *The Origins of Endogenous Growth*, donde Romer afirma que esta corriente enfatiza que “el crecimiento económico es un resultado endógeno del sistema económico y no

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

el resultado de fuerzas externas”. Los endogenistas son una nueva generación de neoclásicos que critican importantes aspectos de la teoría neoclásica más antigua, pero a la vez construyen en base a ella y desarrollan ángulos que para muchos son novedosos y punto de partida de innumerables investigaciones prácticas y teóricas, de tal manera que los endogenistas se encuentran hoy en lo más alto de la corriente teórica del crecimiento.

El ensayo seminal de Romer (1986) presenta un modelo de crecimiento a largo plazo, impulsado principalmente por la acumulación de conocimientos. El conocimiento como insumo presenta para este autor un rendimiento creciente y efectos positivos externos (externalidades o *spillovers*) y considera el cambio tecnológico como endógeno. Lo importante de la tesis de Romer es el abandono del supuesto de los rendimientos decrecientes de la teoría de Solow y un regreso a las posiciones clásicas de los rendimientos crecientes (se basa en A. Smith y la creciente especialización y la división del trabajo).

Además, retoma el estudio de Kenneth J. Arrow de 1962 llamado *The Economic Implications of Learning by Doing*. La referencia a Arrow muestra cómo se van enlazando las aportaciones para dar sustento a la nueva corriente y es quizá una de sus principales influencias. Arrow propone un “modelo dinámico de crecimiento” empujado por los rendimientos crecientes del *learning by doing*; según él, los rendimientos crecientes surgen porque se descubre el nuevo conocimiento cuando tiene lugar la inversión y la producción. Arrow reafirma la importancia del cambio tecnológico en el crecimiento económico y su relación con la formación del capital. Además, propone que se agregue a la teoría el hecho, para él obvio, de que el conocimiento es creciente en el tiempo, definiendo al conocimiento como aprendizaje, y éste como producto de la experiencia; también propone la hipótesis de que el cambio tecnológico puede ser atribuido a la experiencia. Reconoce que sería necesario incorporar a su modelo variables adicionales como las instituciones, la educación y la investigación. Éstas variables permiten que el aprendizaje suceda más rápidamente, aumentando la productividad y el crecimiento económico.

Romer critica al modelo de Solow respecto a que la tasa de rendimiento de la inversión y la tasa de crecimiento del producto per cápita son proporcionalmente decrecientes al nivel del capital per cápita. Romer propone un modelo alternativo en el que la productividad del trabajo crece a una tasa creciente en el tiempo sin límites. La tasa de inversión y la tasa de rendimiento del capital puede incrementarse en vez de disminuir con el aumento del stock de capital. Para Romer los resultados dependen principalmente del abandono del supuesto de los

rendimientos decrecientes.

El nuevo modelo propone un modelo de equilibrio con cambio tecnológico endógeno en el cual el crecimiento a largo plazo es impulsado principalmente por la acumulación de conocimientos de “agentes maximizadores de ganancias, progresistas y dinámicos.” Romer considera que, a diferencia del capital físico, el nuevo conocimiento es el resultado de un proceso de investigación y desarrollo (I+D) que exhibe rendimientos crecientes. También supone, que la inversión tiene un “efecto positivo sobre las posibilidades de producción de otras empresas” porque el conocimiento no puede ser absolutamente patentado o mantenido en secreto. Lo más importante para Romer es precisamente, que el conocimiento puede tener un producto marginal creciente, de tal manera que el conocimiento crecerá sin límites y que la economía con más grande stock de capital humano experimentará un crecimiento más rápido (Romer, 1990).

El ensayo de 1986 antes citado y el artículo *Endogenous Technological Change* de 1990 de Romer son considerados los trabajos fundadores del endogenismo. El segundo modelo complementa al primero y, además, añade el capital humano (H) a los insumos del modelo de Solow: capital (K), mano de obra (L) y la tecnología (A).

Su argumento en el ensayo de 1990 se basa en tres premisas: la primera, es que el cambio tecnológico está en la base del crecimiento económico; el cambio tecnológico proporciona el incentivo para una continua acumulación de capital. La segunda premisa es que el cambio tecnológico surge por las acciones deliberadas de las personas que responden a los incentivos del mercado, por consiguiente, “el modelo es de cambio tecnológico endógeno”. La tercera, y más importante premisa, es que las llamadas “instrucciones” son sustancialmente diferentes de otros bienes económicos; una vez que se ha incurrido en el costo de crear una nueva serie de instrucciones, éstas pueden ser usadas una y otra vez sin costes adicionales ya que son costos fijos. Esta es una característica que define a la tecnología.

1.2.3. ROBERT E. LUCAS JR.

El trabajo de Lucas *On the Mechanics of Economic Development* (1988) junto con el de Romer (1986) establecen las bases de la “nueva teoría del crecimiento”. Lucas parte del modelo neoclásico de Solow y Edward Denison realizando adaptaciones para incluir los efectos de la acumulación del capital humano y lo propone como motor de crecimiento alternativo al

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

modelo de Solow. Define como formación de capital humano la escolaridad, el entrenamiento en el trabajo y el *learning by doing*. Plantea que la teoría neoclásica es incapaz de explicar la variación del crecimiento entre los países. Retoma el concepto de capital humano y lo define como el nivel general de destreza de un individuo (Lucas, 1986).

Se observa, a mitad de los años ochenta que existe una complementación de las aportaciones fundamentales: Solow parte del modelo Harrod-Domar, pero lo critica y aporta la influencia decisiva del cambio tecnológico a la teoría del crecimiento a largo plazo. Dicha tesis se mantiene durante varios años, y es retomada por Romer, que la critica y endogeniza, además de agregarle el “conocimiento” y el “capital humano”, al igual que Lucas. El resultado es una teoría endógena donde el crecimiento puede ser impulsado desde el cambio tecnológico, el desarrollo del conocimiento, la educación, el aprendizaje y el capital humano en un mundo competitivo, sin descartar los factores productivos tradicionales aunque se consideran ahora secundarios.

1.2.4. LOS POSKEYNESIANOS

En términos generales, estos autores han perseguido la transmisión y ampliación de las ideas expuestas por Keynes en algunos de sus trabajos. Pero, persiguen una finalidad mucho más amplia que la de centrarse en desarrollar las ideas de Keynes ya que también se basan en las aportaciones de otros autores con ideas más o menos afines a las de aquél. En el ámbito del crecimiento, esta corriente señala que la economía capitalista se caracteriza por ser inestable debido a la divergencia que existe entre los incentivos privados y los sociales, que conducen a fallos en la demanda efectiva. Así pues, el equilibrio es un caso particular de la tendencia general, que es el desequilibrio.

Señalan que la acumulación de capital es importante para la economía, ya que influye en la inversión y, por consiguiente, en el nivel de empleo y aquélla surge cuando los empresarios esperan alcanzar unos mayores beneficios futuros. Ello significa que los *animal spirit* (o expectativas empresariales) son un factor significativo a la hora de potenciar la incorporación del capital, aumentando así el nivel de inversión.

En resumen, los poskeynesianos desarrollan el modelo propuesto por Harrod por otras vías alternativas a las seguidas por Solow e incluyendo algunas hipótesis keynesianas.

1.3. TEORÍA ECONÓMICA EVOLUCIONISTA (NEOSCHUMPETERIANOS)

Los autores Nelson y Winter en su estudio *An Evolutionary Theory of Economic Change* (1982) proponen una nueva teoría para analizar los fenómenos asociados al cambio económico. Consideran como primera premisa que el cambio económico es importante y que “nada es más digno de atención que la comprensión de los cambios acumulados en la organización económica y tecnológica”. La teoría evolucionista (o apreciativa) toma prestada la idea de la selección natural darwinista aplicándola a las empresas y a su habilidad para crecer y sobrevivir en un ambiente de mercado.

Llaman “formal” a la teoría dominante por su alto grado de utilización de modelos matemáticos y los principios básicos que asume, como la maximización de la ganancia, el equilibrio, la competencia, la información y la predicción perfecta, el individuo racional o el pleno empleo. Estos métodos son rechazados por la corriente evolucionista. Los modelos evolucionistas incorporan las innovaciones y el avance tecnológico como los determinantes del crecimiento a largo plazo. En este sentido, la teoría evolucionista retoma las ideas de Schumpeter, definiéndose así mismos como Neoschumpeterianos.

Reivindican a los clásicos y se esfuerzan por demostrar que las nuevas teorías, e incluso la de Solow, no aportan nada a lo que ya antes otros neoclásicos habían formulado. Se basan en estudios empíricos y explicaciones abstractas: “hasta recientemente el modo normal de la teorización económica fue verbal. Smith, David Ricardo, Marshall, Frank Knight, Schumpeter, Keynes, usaron palabras, no usaron matemáticas como su medio teórico dominante” (Nelson y Winter, 1982). La teoría evolucionista pretende proporcionar un análisis que pueda explicar la situación de las empresas y del crecimiento a largo plazo, particularmente, en el manejo de la teoría del avance tecnológico y de las innovaciones, al estilo de Schumpeter.

Se encuentran próximos a la teoría endogenista en cuanto a la importancia del progreso técnico pero consideran que esta teoría sigue considerando el avance técnico y el crecimiento como variables que tienden al equilibrio y con agentes que predicen perfectamente cuando en “la realidad la economía está en continuo desequilibrio y hay incertidumbre”. Los evolucionistas aseveran que los componentes del crecimiento se refuerzan o se acompañan mutuamente, pero el avance técnico es “el principal conductor y catalizador” que induce y apoya nuevas inversiones en tecnologías, en capital físico y humano (Nelson y Winter, 1982).

La corriente evolucionista ha extendido su influencia en base a una floreciente producción

teórica y empírica. Nelson y Winter señalan que los argumentos evolucionistas regresan debido, en parte, a que la teoría neoclásica no es capaz de enfrentar los “desequilibrios dinámicos” (Nelson y Winter, 1996). El pensamiento evolucionista renace en la década de los 80 y 90 con nuevas herramientas analíticas, producto de una amplia investigación interdisciplinaria (Nelson y Winter, 2002).

1.4. TEORÍA CLÁSICA DEL DESARROLLO

El autor Jaime Ros (2001) reivindica la que denomina Teoría Clásica del Desarrollo como más adecuada para el análisis de los países en desarrollo que las dos corrientes principales neoclásicas: la exógena y la endógena. Sugiere combinar el enfoque clásico con los avances en la Teoría Moderna del Crecimiento para responder cuestiones como por qué algunos países son más ricos que otros, por qué algunas economías crecen mucho más rápido que otras, por qué (en definitiva) la enorme mayoría de la población mundial vive en la pobreza si la tecnología moderna ha mostrado su potencial para elevar las condiciones de vida a niveles del primer mundo (Ros, 2001).

El trabajo de Ros tiene su antecedente en las conferencias que impartía Paul Krugman (1992) en el Banco Mundial. Krugman reivindica una corriente de economistas que se dedicó a explicar el retraso de los países pobres y propuso políticas de desarrollo. Como ejemplo, en abril de 1992, ante un auditorio reunido en los salones del Banco Mundial, Krugman declara que la teoría del desarrollo ya no existe, que ha muerto, abandonada en el desván de los trastos viejos, frente a una profesión que se ha volcado al formalismo matemático y al equilibrio general (Krugman, 1992).

El trabajo pionero base de esta corriente es el de Paul N. Rosenstein-Rodan (1943) con respecto a la industrialización de los países de Europa. Krugman y Ros explican que esta corriente pionera se desvaneció por la incapacidad de los clásicos del desarrollo para formular y formalizar adecuadamente sus ideas y por los fracasos prácticos que se mostraron en los países pobres al no superar tal situación.

Ros rescata la competencia imperfecta, el exceso de fuerza de trabajo y los rendimientos crecientes, los cuales considera insuficientemente integrados en la teoría económica. Comparte con la teoría endógena el supuesto de los rendimientos crecientes y las externalidades tecnológicas y, en este sentido, ve la posibilidad de combinar ambas teorías. El modelo de Ros

tiene como motor del crecimiento la acumulación de capital y rechaza la utilización de modelos neoschumpeterianos porque los considera menos relevantes para los países en desarrollo (Ros, 2000).

Ros explica que los países recientemente industrializados, que alcanzaron en los últimos años una convergencia en el ingreso per cápita con los países ricos, son ejemplos que apoyan los modelos de la teoría clásica del desarrollo. El autor, por tanto, propone retomar la teoría clásica, y combinarla con algunos supuestos de otras corrientes para analizar y proponer políticas económicas que ayuden a superar el subdesarrollo sin atender la situación de los países desarrollados.

1.5. GENERAL PURPOSE TECHNOLOGIES

Una de las corrientes más actuales que pretende explicar el crecimiento económico es la conocida como General Purpose Technologies (GPTs) o Tecnologías de Utilidad General. Considera el progreso tecnológico como fuerza motora del crecimiento, en general, y a las innovaciones “drásticas”, en particular, con lo que se aproxima a las ideas clave de los evolucionistas. Esta teoría reconoce que hay antecedentes de economistas e historiadores que proclamaron a la tecnología como una de las principales fuentes del crecimiento económico, por tanto, sus autores se definen como una corriente que reactiva una vieja tradición más que el desarrollo de una nueva. El trabajo que dio origen y nombre a la corriente es el de Bresnahan y Trajtenberg (1992), con desarrollos posteriores como los recopilados por Helpman (1998).

La parte central de las GPTs es que los economistas han prestado poca atención al papel de lo que ellos llaman “innovaciones drásticas.” Éstas son las innovaciones que introducen una discontinuidad en el sentido de que conducen a la sustitución de una vieja tecnología, la cual jugó un papel importante en la industria, por nuevos métodos de producción. Un tipo específico de innovaciones drásticas que califican como *general purpose technologies* serían las innovaciones que tienen potencial para dominar amplios sectores, cambiar drásticamente su modo de obrar, cuyos efectos repercuten a través de la economía entera, afectando a las estructuras sociales, económicas y políticas. Tecnologías como el motor de vapor, la electricidad, el motor de combustión interna, la computación, internet, láseres, cambios organizacionales, son claros ejemplos.

Admiten que hay dos corrientes: una es la evolucionista-apreciativa y la otra es la formal-ortodoxa, con la diferencia de que la primera es capaz de aportar estudios empíricos más valiosos sobre las innovaciones y el progreso tecnológico que la teoría formal, pero al coste de no poder modelarlos matemáticamente.

Por lo tanto, la corriente GPTs se basa en la teoría de los endogenistas y aprovecha los estudios empíricos de evolucionistas como Nelson, Freeman, Soete, Rosenberg y otros para demostrar que el motor del crecimiento se encuentra en el progreso tecnológico; en otras palabras, unen aspectos microeconómicos del proceso de innovación con aspectos macroeconómicos. En esto fundamentan su originalidad y aportación en la comprensión de los determinantes del crecimiento.

1.6. LOS FUTURISTAS

Las teorías expuestas no son las únicas que explican las causas de la riqueza de las naciones. Hay también economistas o sociólogos que no forman parte del círculo de los economistas clásicos que han aportado una visión más amplia del capitalismo. Son autores que han difundido sus ideas por medio de sus libros. Son creadores de términos como la sociedad “poscapitalista”, la “era de la información”, la “era del conocimiento” y “la tercera ola”.

Peter F. Drucker es considerado uno de los pioneros de este tipo de estudios. Drucker (1994) plantea que vivimos una notable transformación que sucede cada cientos de años; según su opinión, se está creando la sociedad postcapitalista desde fines de la Segunda Guerra Mundial. Es una sociedad nueva y distinta donde el recurso económico básico ya no es el capital, ni los recursos naturales, ni el trabajo sino que “es y será el conocimiento”; el valor se crea hoy por la productividad y por la innovación y ambas (son) aplicaciones del conocimiento al trabajo. Drucker ve una sociedad que pudiera ser llamada la “era del conocimiento”. Drucker ubica al conocimiento altamente especializado como el recurso decisivo para la creación de riqueza y para la conformación de la nueva estructura de la sociedad postcapitalista. Así como el capital fue el componente principal en la sociedad industrial capitalista, el conocimiento lo es para la sociedad moderna. Los factores tradicionales “no han desaparecido, pero han pasado a ser secundarios”, y se pueden obtener fácilmente siempre que se tenga conocimiento, por tanto, éste es “el único recurso significativo” (Drucker, 1994).

En su obra, *The Age of Discontinuity* (Drucker, 1969), se plantean ideas como las siguientes:

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

- a) la explosión de la nueva tecnología, cuyo fruto son nuevas industrias importantes,
- b) el cambio de la economía internacional a la economía mundial,
- c) una nueva realidad social y política de instituciones pluralistas y;
- d) el nuevo universo de los conocimientos basado en la educación masiva.

En esta obra ya se habla de la privatización y del “trabajador del conocimiento” y de su impacto en la economía y la sociedad.

Para terminar, Alvin Toffler (1982) es otro conocido autor coincidente con Drucker que fijó una nueva era a partir de los años 50, cuando por primera vez los servicios superaron al resto de los sectores económicos en la producción total de Estados Unidos. Es una nueva civilización que denomina la “tercera ola”, consecuencia de una “primera” que duró diez mil años (revolución agrícola) y de una “segunda ola” (revolución industrial). La tercera ola surge como superación de la etapa industrial y es una nueva sociedad donde “el conocimiento es la clave del crecimiento económico del siglo XXI” (Heidi y Toffler, 1982).

Toffler analiza los cambios de la nueva era a partir del “motor tecnológico” y del conocimiento como su “carburante”. La tercera ola es la sustitución del trabajo físico por aquel basado en el conocimiento; es el surgimiento de una nueva “economía del conocimiento” donde el carácter del trabajo es diferente y, por tanto, se requiere un trabajador distinto, inteligente, más informado, con conocimiento especializado (Toffler, 1980).

Con la presente revisión de las principales teorías relacionadas con el análisis del crecimiento económico, se ha tratado de centrar el marco de referencia teórico fundamental en el que se basará la construcción del modelo conceptual causal (ver capítulo VI, Bloque 2), que a su vez apoyará los análisis empíricos realizados. Las teorías endógenas de Romer, la formación del capital humano de Lucas, los tipos de innovación de Schumpeter, el papel otorgado a la innovación por los evolucionistas como Nelson y Winter y la corriente General Purpose Technologies, principalmente, han sido el punto de partida para la elaboración de dicho modelo y así ha tratado de ser reflejado en la selección de las variables y sus relaciones de causalidad.

CAPÍTULO II: HACIA UNA ECONOMÍA BASADA EN EL CONOCIMIENTO Y UN MODELO DE DESARROLLO SOSTENIBLE. EL PAPEL DE LA I+D+I

1. LA ECONOMÍA DEL CONOCIMIENTO

El concepto de economía del conocimiento o “economía basada en el conocimiento” encuentra sus orígenes a principios de los años 60. En las cuatro décadas siguientes, se ha intentado dar explicación a un fenómeno que se ha convertido en una estrategia de desarrollo en muchos países. Respecto a su definición, para muchos autores sigue siendo una palabra de moda y otros incluso se cuestionan su existencia dado que el conocimiento está presente siempre en cualquier economía, por simple que sea (May, 2002). Pero la diferencia en esta economía es que el grado de incorporación del conocimiento y de la información en la actividad económica es tan elevado que induce a cambios estructurales profundos y cualitativos en la economía, transformando las bases de las ventajas competitivas.

Aunque podría tener sus raíces en el trabajo de Adam Smith, el término fue empleado por primera vez por Machlup (1962) y acuñado por Peter Drucker (1969). Es objeto de especial atención en el informe sobre la economía del conocimiento elaborado por la Organización para la Cooperación y el Desarrollo Económico (OECD), del que se ha seleccionado su definición para el glosario de términos. La definición dada por el Banco Mundial es también muy empleada, estableciendo que “se trata esencialmente de una economía donde el conocimiento es el principal motor del crecimiento” (Chen y Dahlman, 2006, p. 1). En este tipo de economías, por tanto, el énfasis se encuentra más en las capacidades intelectuales que en los factores físicos (Powell y Snellman, 2004). En la realidad, existen muchos ejemplos de la influencia del conocimiento en el actual crecimiento económico: (1) el progreso en las tecnologías de la información y la comunicación (TICs) que permite un acceso rápido y poco costoso al conocimiento y la información; (2) la velocidad creciente en los avances científicos y tecnológicos en todos los campos de conocimiento; (3) el carácter global que ha adquirido la competitividad y (4) las nuevas demandas, gustos y costumbres de los ciudadanos impulsadas por estos avances y cambios tecnológicos, sobre todo en el sector de las comunicaciones. En base a todo ello, el Instituto del Banco Mundial enfatiza el hecho de que la mayoría de los

países que han realizado un rápido progreso han implantado previamente unos programas de cambio a nivel nacional inspirados en la economía del conocimiento (IBRD/World Bank, 2007).

En las últimas dos décadas, la literatura y la investigación surgida alrededor de la EC se ha centrado principalmente en el destacado papel del conocimiento (o capital humano) como fuente del crecimiento económico a largo plazo (Aghion y Howitt, 1997; David y Foray, 1995; Grossman y Helpman, 1991; Leydesdorff, 2006).

Las teorías relacionadas con la economía del conocimiento están íntimamente relacionadas con las teorías vistas en el Capítulo I sobre el crecimiento económico y sus razones y, alguna de ellas, es común a ambos campos al considerar precisamente al conocimiento como un factor para el crecimiento económico. La **nueva teoría del crecimiento** (*New Growth Theory*; Romer, 1986, 1990; Lucas, 1988), por ejemplo, ilustra el esfuerzo de los economistas ortodoxos neoclásicos por incorporar el conocimiento en la función de producción. La **teoría evolucionista del cambio económico** observa la economía como un proceso evolucionario (Nelson y Winter, 1982) y, en este proceso, la innovación ocurre como un proceso impredecible.

Específicamente, teorías relacionadas con la economía del conocimiento son las siguientes:

a) La teoría de los **sistemas nacionales de innovación** (NIS, en sus siglas en inglés) (Freeman, 1987; Lundvall, 1992; Nelson, 1993) establece que los procesos de creación, modificación y difusión de las innovaciones de una economía son los resultados de las actividades y las interacciones de organizaciones diferentes dentro de él, formando un sistema de innovación. En una definición amplia, el sistema de innovación incluye las organizaciones involucradas en la búsqueda y la exploración del conocimiento y todas las partes integrantes de la estructura económica e institucional.

b) La teoría de la **triple hélice** entiende la economía del conocimiento (Etzkowitz and Leydesdorff, 2000; Leydesdorff, 2006) como una triple hélice en dos capas diferentes: capa funcional y capa institucional. Las tres instituciones son la universidad, la industria, y el gobierno. De un lado, las hélices interactúan entre sí, mientras cada una actúa recursivamente; por otro lado, las dos capas de funciones e instituciones interactúan sobre cada una; ambos tipos de interacciones conducen a la dinámica del sistema completo.

c) La teoría de la **brecha del conocimiento** se centra en las diferencias tecnológicas y económicas entre los países. En el marco de esta teoría, existen varios modelos, como el modelo de la brecha tecnológica (Abramovitz, 1986), el modelo de la brecha del conocimiento técnico (World Bank, 1999), y el modelo de la brecha digital (Baskanran and Muchie, 2006). La asunción central es que los países que se encuentran retardados respecto a los líderes es porque los primeros tienen menos conocimientos que los segundos. Los países podrían salir de su posición retardada mediante la explotación del conocimiento y las tecnologías avanzadas desarrolladas por los países líderes. La recuperación de terreno por estos países no es automática, necesita ciertas condiciones denominadas por Abramovitz (1986) “capacidades de absorción” y esfuerzos intensos y generalizados.

En la literatura existen tres puntos de vista basados en las asunciones relacionadas con el conocimiento: conocimiento como un “activo”, conocimiento como una “relación” y conocimiento equivalente a “capacidad” (ver figura 2.1). Entre ellas, el enfoque del conocimiento como una capacidad aparece como el más apropiado en la actualidad.

La nueva teoría del crecimiento y la teoría de la brecha del conocimiento pertenecen al punto de vista del conocimiento como un activo. Los típicos activos de conocimiento son los recursos humanos, prototipos, la tecnología incorporada en las máquinas, equipos, instalaciones y procedimientos tecnológicos de las organizaciones empresariales. En esta vertiente, el conocimiento es un activo que puede difundirse fácilmente de una entidad a otra; cuanto mayor sea el stock de conocimiento de un país, mejor será su posición económica y su ventaja competitiva en comparación con otros.

Una economía puede construir stock de conocimiento a través de sus propias actividades como inversión en investigación y desarrollo (Cowan et al., 2000; Lucas, 1988; Romer, 1990) o puede adquirir conocimiento de otros países (Abramovitz, 1986; Romer, 1994). El tema de la protección de los derechos de propiedad intelectual en estas teorías adquiere una especial relevancia.

Las teorías que entienden el conocimiento como una relación lo consideran como un recurso que se construye y se comparte socialmente. Esto tiene que ver con las conexiones e interacciones sociales, con las redes de la variedad de actores que forman parte de un sistema económico. La teoría del sistema nacional de innovación y la teoría de la triple hélice se relacionan con este punto de vista en diversas formas (Lundvall, 1988; Leydesdorff, 2006).

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

Para el punto de vista del conocimiento como capacidad, las capacidades se refieren a las de las empresas y a las capacidades sociales de una nación. El conocimiento entendido como capacidad ha sido usado para denotar la “habilidad de actuar”.

Figura 2.1. Tres visiones de la economía del conocimiento

Vistas de la Economía del Conocimiento	El Conocimiento como un activo	El Conocimiento como Capacidad	El Conocimiento como Relación
Teorías de la Economía del conocimiento	Nueva teoría del Crecimiento (Romer, 1986; Romer 1990; Lucas, 1998)	Teoría Evolutiva del Cambio Económico (Nelson y Winter, 1982)	Teoría de la Triple Hélice de la Economía del Conocimiento (Etzkowitz y Leydesdorff, 2000; Leydesdorff, 2006)
	Brecha Tecnológica, Teoría de la Brecha del Conocimiento (Abramovitz, 1986; WB, 1999; Baskaran y Muhie, 2006)	Teoría del Sistema de Innovación Nacional (Lundvall, 1992; Nelson, 1993)	

Fuente: Duc Dang y Katsuhiko Umemoto, 2009.

El hecho de que la capacidad de una entidad se atribuya a la misma como un todo, no reducible a ninguna de sus partes, implica que el enfoque de las capacidades puede examinar la EC de forma sistemática. De hecho, muchos estudios sugieren que el concepto de las capacidades puede proveer un enlace entre las tres visiones del conocimiento (i.e. Wright et al., 2001).

Estos puntos de vista, conceptos, ideas y teorías se han tenido en cuenta en nuestro modelo conceptual, tal y como veremos en detalle en el Bloque 2.

2. EL DESARROLLO SOSTENIBLE

El desarrollo sostenible (DS), además de no ser fácil el encontrar una definición consensuada, es un concepto que ha evolucionado a la par que las sociedades que proclaman su intención de optar por este modelo de crecimiento. El debate suscitado en torno al término desarrollo sostenible, originaría la publicación en 1987 del destacado trabajo de la Comisión Mundial de Medio Ambiente de las Naciones Unidas, bajo las órdenes de Gro Harlem Brundtland, entonces presidenta de la Comisión Mundial de Medio Ambiente y Desarrollo (CMMAD),

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

mejor conocido como Informe Brundtland (Brundtland, 1987). Esta obra, marcó el punto de inflexión en el proceso de institucionalización del concepto de desarrollo sostenible que hasta entonces había sido, básicamente, una mera discusión académica y a partir de ese momento será trasladado a la esfera política⁵. Una de las definiciones más conocidas de DS es la de este Informe y se ha recogido en el glosario de términos⁶ (Brundtland, 1987, p. 43).

El mérito de este informe se encuentra en el esfuerzo por hacer operativo el concepto así como por establecer una plataforma para negociaciones internacionales. Con el paso del tiempo, se ha ido ampliando este concepto hasta englobar también los asuntos sociales y económicos. Por otra parte, considera que el desarrollo no es sólo crecimiento, sino que incluye otra serie de factores sociales y ambientales, y exigen que se produzca un cambio en el modelo de crecimiento que repercuta sobre el menor consumo de recursos y energía y sobre una mayor equidad social (CMMAD, 1987; p. 77).

No será hasta el último cuarto del siglo XX cuando el término DS empiece a cobrar notoriedad. A lo largo de la historia, la preocupación por el desarrollo económico ha sido una constante para los economistas y la sociedad en su conjunto pero, por primera vez, los economistas enfocan ese desarrollo desde una perspectiva multidisciplinar que va a englobar aspectos ambientales y sociales.

El objetivo del crecimiento y del progreso ha sido el tema central de los economistas desde que la economía se convirtió en ciencia. Tras la II Guerra Mundial, empezaron a surgir trabajos empíricos que pretendían analizar las diferencias encontradas entre los diversos procesos de crecimiento desarrollados en el mundo. Se pueden encontrar diferentes teorías y distintas formas de clasificar a estos autores, aunque se pueden agrupar en cinco grandes corrientes de pensamiento: Teoría de la Modernización, Teoría Estructuralista, Teoría Neomarxista, Teoría Neoliberal y Teorías Alternativas (Aguado et al., 2008).

El DS ha adquirido una enorme profusión en todos los ámbitos políticos y sociales, llegando a emplear el término de forma abusiva, eludiendo las ideas fundamentales que representaba en su origen (Goldin y Winters, 1995). El Desarrollo Sostenible no es sinónimo de protección ambiental y tampoco hay que confundirlo con crecimiento sostenible, cayendo en el error que

⁵ Sin olvidar la publicación del informe *Los límites del crecimiento* (Meadows et al., 1972) que supuso un aviso sobre las posibles consecuencias no deseadas del crecimiento económico.

⁶ Además de esta definición, existen más de 300 según Pezzey (1992).

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

frecuentemente se da en el ámbito político al hacer uso del término desarrollo sostenible de forma contradictoria y poco rigurosa, e incluso utilizándolos como sinónimos (Allende, 2000).

Aceptando el planteamiento de las líneas anteriores, puede afirmarse que el Desarrollo Sostenible incorporaba de forma implícita una percepción tridimensional en la que se integraban los aspectos económicos, ecológicos y sociales. En este sentido, fue el Banco Mundial quien enfocó el estudio del desarrollo sostenible a través de un marco triangular, englobando aspectos sociales, económicos y ambientales. Así, por lo general, se considera que el desarrollo sostenible está formado por estos tres grandes componentes (Hardi y Zdan, 1997), lo que se conoce como “la triple línea base”. El bienestar en estas tres áreas está entrelazado, y no es independiente. Por ejemplo, una sociedad saludable y próspera depende de un medio ambiente adecuado para que le provea de alimentos y recursos, agua potable, y aire limpio para sus ciudadanos. El paradigma de la sostenibilidad rechaza el argumento de que las pérdidas en los ámbitos ambiental y social son consecuencias inevitables y aceptables del desarrollo económico.

De acuerdo con esta percepción, será el Tratado de la Unión Europea de 1992 el que incluya el desarrollo sostenible entre sus objetivos al asumir el principio de integración ambiental vinculado al DS. En el año 2001, se adoptará la primera estrategia de desarrollo sostenible de la UE. A lo largo de su texto se mantiene que para conseguir una Europa sostenible es necesario tener una visión más amplia a largo plazo, lo cual implica que crecimiento económico, cohesión social y protección ambiental deben estar estrechamente unidos. Años más tarde, el Consejo Europeo de Bruselas (junio de 2006), considera el DS como uno de los objetivos fundamentales de la UE y elabora “una Estrategia Renovada de Desarrollo Sostenible de la Unión Europea ambiciosa y amplia”, lo que ratifica el compromiso de la UE con el desarrollo sostenible. Destacables son también las menciones al desarrollo sostenible en la Carta de Derechos Fundamentales de la UE.

En la actualidad de la UE, los Programas de Medio Ambiente (PMA) vienen orientando su política medioambiental desde principios de la década de los 60. El VI Programa de Medio Ambiente (VI PMA) expiró en Julio de 2012. La Comisión Europea, en respuesta a la demanda de las partes interesadas, entre las que se cuentan el Consejo y el Parlamento Europeo, propuso un Programa destinado a sucederlo (Comisión Europea, 2012).

El nuevo PMA ofrece un marco general para la política medioambiental hasta 2020 y establece objetivos prioritarios (nueve) para la UE y sus Estados Miembros. La propuesta se enmarca en

un cuádruple contexto. En primer lugar, a pesar de los avances, persisten graves problemas medioambientales. En segundo lugar, la UE ha adoptado la Estrategia Europa 2020 para un crecimiento inteligente, sostenible e integrador, que orienta la evolución política hasta el año 2020. En tercer lugar, la necesidad de reformas estructurales de la crisis abre nuevas oportunidades para que la UE avance hacia una economía verde e integradora. Por último, el proceso de «Río+20» ha puesto de relieve la importancia de la dimensión global, ejerciendo una fuerte influencia en la estrategia europea de DS así como en la Estrategia Europa 2020 (lo que supone una muestra más de la interrelación entre los conceptos, como veremos en el apartado siguiente).

3. INTERRELACIÓN ENTRE LA I+D+I, LA ECONOMÍA DEL CONOCIMIENTO Y EL DESARROLLO SOSTENIBLE

Una vez expuestos y contextualizados los conceptos relacionados con el cambio tecnológico y la innovación en general (capítulo I) y de la economía del conocimiento (EC) y el desarrollo sostenible (DS) en los dos apartados anteriores, resulta más fácil captar la íntima conexión entre los tres conceptos y el por qué muchos temas/indicadores relacionados para su medición presentan puntos comunes entre ellos.

Si se consideran los tres conceptos como círculos anidados que se superponen sobre un círculo mayor, que representaría el bienestar humano, quedaría menor superficie libre del solapamiento cuanto mayor sean los círculos de la EC y del DS. Es decir, en la medida en que el medio ambiente, la sociedad y la economía se alinean más, el área de solape aumenta al igual que aumenta el bienestar humano. El concepto más amplio de ellos sería el de desarrollo sostenible, que abarcaría el componente más restringido de economía del conocimiento, siendo ambos impulsados por las actividades de investigación, desarrollo e innovación que conducen al cambio tecnológico. Todo ello influenciado transversalmente por el motor que constituye la educación (ver figura 2.2).

En estos componentes, es clara la influencia que ejercen las actividades de I+D+i para su mejora. La relación de estos conceptos se hace aún más obvia si hablamos de ejemplos concretos de los logros de las actividades de I+D+i: en los innumerables avances y descubrimientos conseguidos en el área de la salud que han incrementado nuestro bienestar y esperanza de vida, en las tecnologías de la información y la comunicación (móviles, internet,

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

con las nuevas formas de relaciones sociales que conllevan) y la influencia de éstas a su vez en las nuevas formas de educación (enseñanza no presencial, tutorías on-line), en el medioambiente (biodiesel, vehículos eléctricos, edificios inteligentes y sostenibles), el papel que ha jugado en todo ello los avances en la nanotecnología (adelantos en la medicina, en el tratamiento del agua y los alimentos), etc.

Para el uso de este conocimiento, las TICs aportan una herramienta de difusión excepcional que se retroalimenta con el propio conocimiento creado y difundido, conocimiento como fuente u origen y como vehículo conductor del mismo a niveles globales nunca antes conocidos.

En la generación de actividades de I+D+i y en la consecución de una economía del conocimiento, otra fuerza motora que impulsa y se despliega entre todas las capas de la sociedad es la educación, sin la que no se puede imaginar la creación del conocimiento necesario empleado en las actividades de I+D+i, ni la consecución de una EC ni, en última instancia, un desarrollo sostenible.

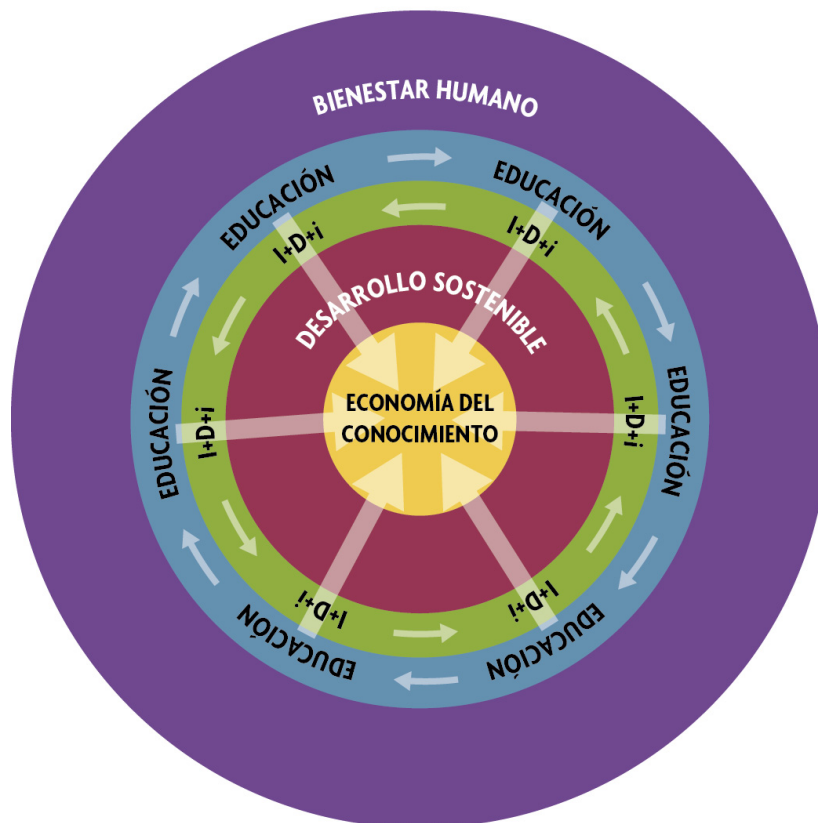
A nivel global, podemos afirmar que para alcanzar el éxito en una economía globalmente competitiva un país debe aprovechar los recursos con los que cuenta y que presentan alguna ventaja comparativa frente a sus competidores, más aún en aquellos países con escasos recursos naturales (Thurow, 1996).

Aunque las actividades de producción siguen siendo un importante componente del crecimiento, la entrada del nuevo milenio ha supuesto que los productos y servicios se centren aún más en actividades basadas en el conocimiento (Abramowitz, 1989). La esencia de una economía basada en el conocimiento, es la capacidad de absorber, procesar y aplicar “conocimiento” y traducirlo en una fuente clave de ventaja competitiva junto a los factores básicos de la producción, que siguen siendo importantes en este tipo de economía.

Como hemos visto de forma concreta en el apartado 1, en una economía de este tipo la productividad y el crecimiento económico derivado están determinados en gran medida por la tasa de progreso técnico, la acumulación de conocimiento y una red o sistemas que pueden de forma eficiente distribuir dicho conocimiento. En términos de trabajo o de empleo, también hay un cambio en los requerimientos. Los trabajadores necesitan ser más creativos e innovadores con el uso del conocimiento adquirido y poder adaptarse a las rápidas tecnologías cambiantes. La ventaja competitiva de una economía dependerá de cuán rápido fabrica bienes

y servicios de la misma calidad con el advenimiento de cambios en la tecnología (OECD, 1996) y, los cambios en la tecnología, vienen auspiciados en la gran mayoría por las actividades de I+D+i, tanto públicas como privadas, así como por el incremento y mejora de la educación a todos los niveles.

Figura 2.2. Relación anidada entre la I+D+i, la EC y el DS para la consecución del bienestar



Fuente: Elaboración propia

De acuerdo a los informes de la OECD, el sistema científico de un país adquiere una importancia cada vez mayor (OECD, 1996; OECD, 1996; Smith et al., 1995). Los laboratorios públicos de investigación y las instituciones de educación superior están en el corazón del sistema público de la ciencia que incluye, de forma más amplia, los ministerios correspondientes y los agentes del conocimiento junto a las empresas y otros organismos privados; todo ello apoyado por la adecuada infraestructura.

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

Desde esta visión global, el sistema científico está relacionado con la EC al contribuir a las siguientes funciones clave:

a) La producción de conocimiento, facilitando nuevo conocimiento a través de la investigación, el desarrollo y la innovación y con la transmisión del conocimiento mediante la educación y el desarrollo de recursos humanos. El sistema científico es un elemento crucial en esta transmisión, en particular la educación y la formación de científicos e ingenieros. En la EC, el aprendizaje se torna extremadamente importante en la determinación del destino de los individuos, empresas y economías nacionales. Las capacidades humanas para adquirir y aprender nuevas competencias y aplicarlas son clave para absorber y emplear nuevas tecnologías. Formados apropiadamente, los investigadores y los técnicos son esenciales para producir y aplicar tanto el conocimiento científico como el tecnológico. Las universidades, de una forma especial, son centros de educación y formación de una fuerza de trabajo especializada en la EC.

b) La transferencia de conocimiento para diseminar éste y facilitar medidas de solución de problemas. El sistema científico contribuye a la transferencia y difusión del conocimiento por toda la economía. Uno de los sellos distintivos de la EC es el reconocimiento de que la difusión del conocimiento es tan importante como su creación, llevando a una atención creciente hacia las *redes de distribución del conocimiento* y a los *sistemas nacionales de innovación* (NIS). Los NIS son los agentes e infraestructuras que apoyan el avance y el uso del conocimiento en la economía y las conexiones entre ellos (Sharif, 2006). Son cruciales para la capacidad de un país para la difusión de las innovaciones y absorber y maximizar la contribución de la tecnología al proceso productivo. En este contexto, el sistema científico desempeña un rol principal en la creación del conocimiento que facilita el proceso técnico y desarrolla una base cultural común para el intercambio de información (David et al., 1995).

En cuanto a su relación con el desarrollo sostenible, la finalidad última de la ciencia económica es entender y promover el desarrollo económico para contribuir al bienestar de la sociedad. Los economistas al centrarse en los mecanismos y los facilitadores del desarrollo pierden a veces el punto de vista de los agentes económicos y su bienestar. Por ejemplo, la innovación es un elemento central del desarrollo económico, como hemos visto en el apartado de la revisión histórica de las teorías del crecimiento, pero no conduce automáticamente a mejorar

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

el bienestar y a distribuir justicia (Hartmann y Pyka, 2013). La comunidad llamada Globelics⁷ ha contribuido a responder algunas cuestiones cruciales sobre la interacción entre aprendizaje, innovación y creación de bienestar prestando atención a la necesidad de una mejor comprensión de la mutua relación entre las capacidades humanas y los sistemas de innovación (Johnson et al., 2003; Cozzen and Kaplinsky, 2009). Con este propósito, los autores sugieren integrar los puntos de vista del enfoque de las capacidades de Amartya Sen (Sen, 1999) y la literatura sobre los sistemas nacionales de innovación (Freeman, 1987, Lundvall, 1992).

A pesar de este cuerpo de doctrina dentro del desarrollo económico y la innovación, una pregunta sigue sin ser contestada y obviada: ¿cuál es el impacto de la innovación sobre el desarrollo humano y el bienestar? El argumento para la ausencia de respuestas puede descansar en el hecho de que existe la asunción implícita de que el progreso, la innovación y la diversificación son positivos *per se*.

La innovación es ciertamente un motor para el desarrollo humano (U.N.D.P., 2001) pero la innovación puede tener también efectos negativos como la destrucción de puestos de trabajo y competencias y la creación de polarización y desigualdades estructurales. Para los responsables políticos, la cuestión esencial debería ser qué tipo de innovación, cambio estructural y diversificación económica debería promoverse, dónde y cuándo.

⁷ The Global Network for Economics of Learning, Innovation, and Competence Building Systems (Globelics) se define a sí misma como una comunidad mundial, abierta y diversa de académicos que trabajan en la innovación y la creación de competencias en el contexto del desarrollo económico. Su objetivo principal es contribuir a la creación de conocimiento académico y el intercambio en el ámbito de la innovación y el desarrollo.

CAPÍTULO III: SEGUIMIENTO Y EVALUACIÓN DE POLÍTICAS PÚBLICAS. LOS RECURSOS PÚBLICOS DE LA I+D+I Y LA POLÍTICA CIENTÍFICA Y TECNOLÓGICA EN LA UE

1. LA EVALUACIÓN DE POLÍTICAS PÚBLICAS EN GENERAL. CONCEPTO, FUNCIONES Y RELEVANCIA

En esta investigación, se abordará la evaluación y monitorización de la eficiencia y de los resultados/ejecución del uso de los recursos públicos, en concreto, de los recursos dedicados a la I+D+i, desde una perspectiva macroeconómica o a nivel de país. Y es que en el ámbito del análisis de políticas públicas, la fase (final) de la evaluación ha sido la menos tratada y estudiada. Sin embargo, la evaluación de programas y políticas públicas es una realidad que ha ido creciendo paulatinamente y que reclama la atención de numerosos estudios.

En general, el análisis de una política pública es una actividad intelectual y práctica llevada a cabo dentro de un sistema político para mejorar la ejecución de esa política con información relevante obtenida en el correspondiente análisis (Degtyarev, 2004). La evaluación de políticas públicas puede definirse como “la producción de información que sirva a los gestores públicos para tener una idea sobre la *calidad* de los servicios que están produciendo, en función de lo que se estime como *valor*” (Xavier Ballat, 1992). Dunn (1994) define la evaluación como un “procedimiento analítico-político usado para producir información sobre el valor o la importancia de líneas de acciones pasadas y/o futuras”. Otra de las definiciones conocidas de evaluación es la adoptada por el *Joint Committee on Standards for Educational Evaluation* que se convierte en referencia en el ámbito de la evaluación. La evaluación queda definida como la investigación sistemática del valor o mérito de un objeto (Joint Committee, 1994).

En este trabajo, la evaluación se entenderá como la producción de información sobre qué resultados de la política contribuyen a alcanzar el bienestar social (Smith and Street, 2005). Cuando contribuyen al logro de esta meta, decimos que la política ha conseguido algún nivel de ejecución significativa (los correspondientes problemas sociales han sido, al menos parcialmente, resueltos) o que ha sido eficaz. También Weiss (1998), que utiliza el término de investigación evaluativa (*evaluation research*), afirma que ésta consiste en la utilización de la

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

metodología propia de las ciencias sociales para hacer que el proceso evaluador sea más exacto y sistemático.

Parece haber un consenso sobre la conveniencia y la necesidad de evaluar la acción de los poderes públicos como una herramienta clave en la gestión pública (Monnier, 1995). Esta percepción de la conveniencia y la necesidad de incorporar la evaluación como una herramienta útil para la mejor gestión de los recursos públicos, hace necesaria una investigación que analice la realidad a nivel macroeconómico.

El análisis de políticas públicas ha ido ampliando su objeto de estudio yendo del simple análisis de la decisión pública y sus implicaciones, al estudio de las complejas acciones que se derivan y sus efectos sobre el planteamiento del problema original (Parsons, 1995).

Decididamente, el sentido último de la evaluación es su utilidad para el proceso de intervención social, siendo capaz de generar información útil y con carácter práctico, es decir, dirigida a la acción (Weiss, 1998). Respecto a las funciones de la evaluación Stufflebeam y Shinkfield (1987) señalan tres funciones según el uso que se quiera hacer de la evaluación: perfeccionamiento o mejora (*improvement*), recapitulación, rendición de cuentas o responsabilidad (*accountability*) y ejemplificación o iluminación para acciones futuras (*enlightenment*).

Desde la perspectiva política, la evaluación es un instrumento clave para que los representantes políticos comprueben que la administración está realmente ejecutando las tareas que se le han asignado. Y es que los instrumentos para la intervención del gobierno deberían estar diseñados para maximizar una función social de bienestar explicada en su programa electoral y sus políticas. Sin embargo, es sorprendente como éstos algunas veces no consideran si esa función institucional, que deriva de las políticas públicas, puede alcanzar efectivamente y eficientemente ese objetivo.

La literatura relacionada con la evaluación habla en general de programas, servicios y políticas públicas no distinguiendo, en la mayoría de los casos, procedimientos distintos para objetos diferentes. Es más, se suele hablar de un objeto genérico de la evaluación, al que suele asignarse el concepto de “programa” pero que pretende designar, en general, políticas, planes, programas, proyectos, servicios, medidas y materiales (Joint Committee, 1994).

La utilidad de la información que arroja una evaluación depende, en gran medida, de que ésta llegue a tiempo para la toma de decisiones sobre las mejoras, las posibles modificaciones y el

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

rendición de cuentas (Chelinsky, 1998). Cuanto más inestable e incierto sea el contexto en el que se encuentra el programa o política a evaluar, más acuciante es la necesidad de que la evaluación cumpla los criterios de temporalidad. Dada la naturaleza específica de la evaluación, su éxito no sólo se debe medir por el grado en que ésta contribuye al aumento o mejora del conocimiento sino, sobre todo, por el grado en el que ha resultado valiosa para cubrir las necesidades específicas de información de un cliente, destinatario o usuario particular.

Por la propia naturaleza de la evaluación, ésta debe contar con un último nivel de análisis que consiste en la traducción de los resultados de la evaluación en recomendaciones. Las recomendaciones son el vínculo formal entre una evaluación y su uso. Además, el ofrecimiento de recomendaciones específicas y orientadas a la acción ha sido considerado desde hace tiempo como algo deseable y relacionado con las buenas evaluaciones (Chelinsky, 1998). Patton reconoce que son, a menudo, la parte más visible de un informe de evaluación (Patton, 1997, p. 324). Evidentemente, las recomendaciones de una evaluación deben estar basadas en las valoraciones realizadas que, a su vez, habrán estado basadas en la recopilación y análisis sistemático de la información.

Uno de los principales problemas se plantea a la hora de evaluar el impacto social de las políticas públicas, en la selección de los posibles indicadores, no siendo aconsejable la medida de la rentabilidad privada sino la medición en términos de rentabilidad social para los individuos, en términos de bienestar social. En este contexto, es difícil encontrar indicadores o variables que resuman el valor social de los servicios públicos. Pedraja et al. (2001) sugieren el uso de indicadores aproximados (*proxy*) de los valores de los outputs intermedios.

Como se puede comprobar, el tema de la evaluación o el seguimiento de las políticas públicas no está exento de dificultades. Pedraja et al. (2001) plantean los principales escollos que resumimos en la tabla 3.1.

Una de las dificultades de la medición de los outputs es que el proceso de producción se ve afectado por variables exógenas, socioculturales, etc. que escapan al control directo del gestor. Otros problemas son la multiproducción de las entidades públicas, la combinación de objetivos de eficiencia con los de equidad, cómo fijar un único criterio de medición de la eficiencia, el estado de la tecnología, etc.

Tabla 3.1. Principales dificultades para la evaluación de políticas públicas

PROBLEMAS	LIMITACIONES DEL ANÁLISIS	RECOMENDACIONES PARA LA MEDICIÓN
<i>Valoración y medición del output</i> (outputs intermedios)	<i>Área mínima de comparación</i> (uso de 1 sólo principio)	Uso de las mejores prácticas
<i>Problemas de control</i> (Monopolios)	<i>Utilizar valoraciones sociales</i> (uso de cantidades)	Uso de variables proxy
<i>Multiplicidad de criterios en la función objetivo</i>	<i>Entorno institucional de control</i> (grupos homogéneos)	<i>Técnicas flexibles</i> (incertidumbre)
<i>Ausencia de mecanismos de control de liquidación automática</i> (quiebra)		

Fuente: Pedraja, F., Salinas, J. y Suárez, J. (2001).

2. CONCEPTOS RELACIONADOS CON LA EVALUACIÓN DE LOS RECURSOS EMPLEADOS

Un aspecto íntimamente relacionado con la evaluación de los recursos es el problema de su asignación en un entorno de escasez, lo que alude directamente al problema del uso eficiente de los mismos. En el análisis de la eficiencia de los recursos públicos, se suelen introducir otro tipo de criterios adicionales a los puramente económicos para definir los indicadores sobre cuya base se evalúa el cumplimiento de los objetivos o se analizan los costes y beneficios asociados a los diferentes programas/políticas. Los modelos de evaluación de la eficiencia del sector público pretenden clasificar u ordenar los resultados de la producción de instituciones públicas más rentables desde el punto de vista social.

En resumen, Dunn (1981) propone seis grandes criterios a seguir a la hora de evaluar, que son:

- **Efectividad:** ¿Se han alcanzado los resultados?
- **Eficiencia:** ¿Cuánto esfuerzo se ha requerido para conseguir dichos resultados?
- **Adecuación:** ¿Hasta qué punto el logro de los resultados ha resuelto el problema?
- **Equidad:** ¿Están distribuidos los costes y los beneficios de forma equitativa entre los diferentes grupos?
- **Capacidad de respuesta:** ¿Satisfacen los resultados de la política las necesidades, preferencias o valores de los grupos implicados?

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

- **Conveniencia o propiedad:** ¿Merecen realmente la pena los resultados (objetivos) deseados?

El concepto de eficiencia no es una noción manejada en exclusiva por el área de la Economía Aplicada. La idea de eficiencia tiene otras acepciones y, en ocasiones, suele utilizarse como sinónimo de conceptos que difieren entre sí. La eficiencia económica, en concreto, introduce la relación entre la producción y los recursos mediante la aplicación de un criterio de comparación entre diferentes comportamientos. En la tabla 3.2, se recogen los conceptos que pueden servir para delimitar la noción de eficiencia.

Tabla 3.2. Nociones relacionadas con la eficiencia

	DENOMINACIÓN	TIPOS
CONCEPTOS BÁSICOS	EFICACIA	
	Productividad	<i>Parcial</i>
		<i>Total</i>
	Eficiencia	<i>Técnica</i>
		<i>Asignativa</i>
<i>Económica</i>		
DERIVADOS DE LA NOCIÓN DE EFICIENCIA	Economías de escala	
	Competitividad	
	Rendimiento	
	Economías de alcance	

Fuente: Elaboración propia

En general, la búsqueda de la eficiencia es la búsqueda de la combinación más apropiada de costes y beneficios. Siguiendo a Farrell (1957), las tres dimensiones de la eficiencia son: eficiencia técnica, eficiencia asignativa y eficiencia económica.

La eficiencia técnica (o física) es el concepto de eficiencia utilizado habitualmente en el ámbito público (Pedraja, Salinas y Suárez, 2001). Ésta intenta analizar los procesos productivos y la organización de tareas centrándose en las cantidades de inputs utilizadas y en las cantidades de outputs producidas. La eficiencia técnica se logra si se alcanza el máximo producto o servicio posible con una combinación concreta de factores de producción (orientación al output); o alternativamente, si se alcanza un nivel de producto o servicio determinado con la mínima cantidad de factores o de recursos combinados en una proporción dada (orientación al input). Por tanto, la eficiencia técnica se basa en la utilización de unidades físicas, lo que implica que queda al margen el análisis del coste o precio de los factores y la valoración de los ingresos obtenidos en la producción.

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

Por otra parte, la eficiencia asignativa (o de precios) (Farrell, 1957) hace referencia a la habilidad para combinar los inputs y los outputs en proporciones óptimas en función de los precios. Esto implica alcanzar el coste mínimo al producir un nivel dado de output cuando se modifican las proporciones de los factores productivos utilizados de acuerdo con sus precios y productividades marginales. Alternativamente, se puede definir como la obtención de una cantidad máxima de producto manteniendo el coste a través del reajuste de los factores de producción, según sus costes de uso.

Por último, Farrell (1957) definió la eficiencia económica (o global), la cual implica que la unidad productiva es técnica y asignativamente eficiente, es decir, la obtención de una determinada producción supone la minimización de la cantidad de factores empleados además de su coste. La eficiencia económica se define como el producto de la eficiencia técnica y la eficiencia asignativa.

Un concepto que es preciso analizar junto al de eficiencia es el de “eficacia” o grado de consecución de los objetivos. La noción de eficacia hace alusión únicamente a la obtención de resultados, sin tener en cuenta los recursos empleados por parte de las unidades productivas. Si se fija un determinado objetivo, la institución que consiga un nivel de realización más próximo al objetivo será la más eficaz.

Para solucionar parcialmente la limitación del concepto de eficacia, se suele acudir al análisis de la productividad parcial de los factores, a través de indicadores que relacionan un output con el recurso principal que ayuda a obtenerlo. En cambio, los indicadores de productividad presentan la limitación de que pueden obtenerse dos indicadores iguales con niveles de input y output totalmente diferentes o clasificar a unidades eficientes en peor posición que otras que no aprovechan todas las posibilidades de producción.

Para terminar este apartado de extensiones de la noción de la eficiencia, cabría citar el término “competitividad”. Todas las nociones hasta ahora manejadas reflejan la dimensión interna de la eficiencia, entendida como la obtenida como consecuencia del mejor aprovechamiento de los recursos en un entorno similar. La competitividad representaría la dimensión externa de la eficiencia. En un entorno institucional que favorezca la competencia, las empresas (naciones) competirán ofreciendo productos de calidad mediante la reducción de sus costes, mediante estrategias de diferenciación de los productos o servicios ofrecidos mediante la calidad, y, por supuesto, mediante la novedad que ofrezcan respecto a los de la competencia (innovaciones).

3. EL ESTUDIO DE LOS RECURSOS DEDICADOS A LA I+D+I

Desde el momento en que la inversión en I+D+i es uno de los elementos cruciales para la promoción del progreso científico y tecnológico, cualquier país que emplee sus recursos ineficientemente podría dar lugar a un progreso mucho menor del esperado. Más allá, si los recursos en I+D+i no son empleados eficientemente, las inversiones adicionales que se realicen pueden servir de poca ayuda en la estimulación de dicho progreso. Sin embargo, los estudios realizados se han centrado más en los esfuerzos en nuevas inversiones de I+D+i y han dedicado menos atención al uso efectivo de estos recursos, particularmente a nivel nacional, y toda una vez que han tenido lugar las inversiones (Wang y Huang, 2007).

¿Cómo contribuye el esfuerzo en I+D al bienestar de la nación? ¿cuánto debe invertir un país? ¿qué proporción debería ser pública y qué proporción privada? John H. Marburger III (editorial de la revista Science, 20 de mayo de 2005) planteó estas cuestiones e hizo un llamamiento a lo que denominó “una ciencia de la política científica” (Leydesdorff y Wagner, 2009).

En los estudios que tratan sobre los resultados de la I+D+i, al igual que en el caso de los estudios de la evaluación de políticas públicas, en general, existen muy diferentes perspectivas y las definiciones de lo que se entiende por “ejecución”, “rendimiento”, “eficiencia”, etc. varían ampliamente. La medida de la ejecución o los resultados de los recursos empleados en I+D+i se refleja en términos tales como *eficacia, eficiencia, productividad, calidad, capacidad, éxito, control, monitorización, seguimiento, evaluación, valoración, evaluación comparativa*, incluso *capacidades tecnológicas*, términos que son empleados a menudo como sinónimos y para los que no existe una nítida línea divisoria (Chiesa et al., 2009).

A lo largo de esta investigación, se utilizará el término de la forma más amplia posible, tratando de hacer referencia a cualquier actuación y método que trate de evaluar si los recursos empleados obtienen los objetivos esperados por la inversión en dichos recursos. Los objetivos pueden variar desde los resultados esperados de la I+D+i, a los objetivos más generales de convertir la economía en una economía basada en el conocimiento, siguiendo el modelo de un desarrollo que se pueda calificar como de sostenible, a través de una serie de temáticas relacionadas o escenarios.

La relevancia de la evaluación de las políticas públicas como es la de Investigación, Desarrollo e Innovación (I+D+i) no es ajena a nivel internacional y es destacada en muchas agendas

políticas e informes oficiales como, por ejemplo, el presupuesto del gobierno estadounidense para el año 2013 (el conocido como el proyecto Apolo) o la Estrategia Europa 2020 de la UE (Comisión Europea, 2010).

Una de las respuestas a la actual crisis es el enfoque neo-conservador de un Estado minimalista, es decir, deberían minimizarse los gastos públicos (Aghion, 2012). A pesar de este enfoque, incluso en tiempos de declive económico es fundamental seguir apoyando la Investigación y el Desarrollo (I+D) así como las actividades de innovación (Filippettia and Archibugi, 2011). Desde un punto de vista empírico, muchos autores también han destacado la relevancia de los retornos del gasto en I+D – directos y socioeconómicos (Archibal and Pereira, 2003; Baumol, 2003; Griliches, 1992; Jones and Williams, 1998).

A la hora de evaluar el uso de estos recursos, debemos destacar que las inversiones en I+D+i presentan algunas diferencias respecto a otro tipo de inversiones. Estas diferencias son (Martín and Scott, 2000): i) indivisibilidad de la inversión, ii) economías de escala, iii) la mayor incertidumbre inherente a la inversión (no es seguro que los esfuerzos en I+D darán lugar a nuevo conocimiento e innovaciones) iv) existencia de información asimétrica, v) fallos de mercado y, finalmente, vi) la limitada apropiabilidad de los resultados por parte de los inversores.

Existen dos principales argumentos económicos para los gobiernos para invertir en I+D+i: i) se trata de uno de los factores más importantes para el crecimiento económico a largo plazo con la creación de nuevo conocimiento que produzca innovaciones comerciales; ii) las diferencias de la inversión en I+D+i citadas en el párrafo anterior (este último argumento explica por qué las políticas públicas promocionan y restauran las inversiones en I+D al nivel deseable socialmente y no cubierto por el sector privado).

Al tratar sobre los impactos sociales y económicos de este tipo de políticas, no puede desconocerse el hecho de que la ciencia es, en sí misma, una actividad social (Woolgar, 1991). Oszlak y O'Donnell (1995) advierten de lo complicado que es el tema de los impactos, cuando se preguntan: “Dado un cambio X en cierta característica Z , ¿qué proporción de ese cambio podemos atribuirle causalmente a políticas estatales, a políticas privadas y a otros factores ajenos a unas y a otras?” Esta pregunta, planteada en relación al impacto de las políticas públicas en general, es obviamente válida para el impacto social de la ciencia y tecnología, si se cuestiona acerca de cuál es la proporción de un cambio social atribuible a los efectos de la investigación y del conocimiento científico y tecnológico, o a otros factores. Este tema se

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

abordará en el Bloque 2, como veremos, a la hora de diseñar el modelo conceptual causal de red Bayesiana que mida la eficiencia del uso de los recursos en diferentes escenarios socioeconómicos incorporando el conocimiento experto.

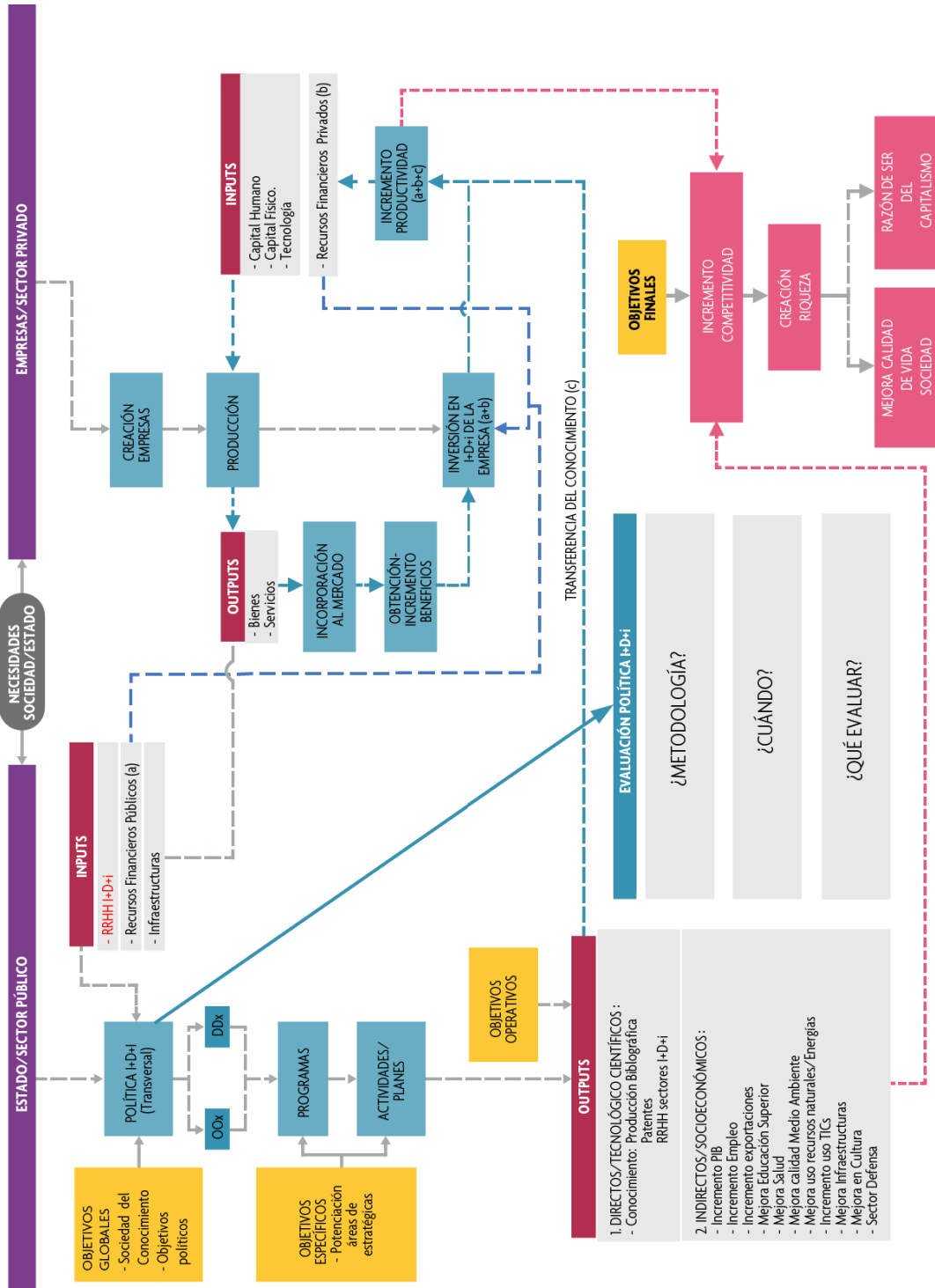
En este ámbito, el conjunto de outputs (impactos) de las políticas públicas de I+D seleccionados durante los últimos veinte años han llevado a una revisión del rol del mecanismo de evaluación. Un rol que ha evolucionado de perseguir objetivos en materia de Defensa nacional a alcanzar metas sociales. Como recapitulación de las ideas expuestas, ver la figura 3.1.

3.1. USO DE INDICADORES DE I+D+I A NIVEL INTERNACIONAL

Cuando los países deciden evaluar de forma cuantitativa su esfuerzo y/o sus resultados en actividades científicas y tecnológicas deben disponer de un conjunto preciso de definiciones y, lo que es más importante, de un conjunto operativo de criterios que permitan diferenciar y clarificar qué constituye (y qué no) cualquiera de los conceptos arriba indicados. De esta forma, precisando y homogeneizando el alcance de conceptos como investigación, desarrollo tecnológico e innovación de tal tipo o de tal otro, etc., se podrá disponer de datos cuantitativos fiables que permitan establecer comparaciones entre países.

El interés, o casi podríamos decir, la preocupación por la medición de la I+D+i no es un tema de actualidad. Desde que la Organización Económica para la Cooperación y el Desarrollo (OECD, 1957) y la National Science Foundation (NSF) americana comenzaron su esfuerzo por crear indicadores estandarizados que permitiesen la comparación internacional de los datos, se ha incrementado el número de este tipo de indicadores así como el número de estudios e informes que tratan de realizar comparaciones entre países. Actualmente, y a nivel europeo, un buen ejemplo es el Innovation Union Scoreboard y el EU Industrial R&D Investment Scoreboard de la Comisión Europea. Otros estudios bien conocidos son el Global Competitiveness Report del Foro Económico Mundial, el World Competitiveness Yearbook del IMD International-Lausana, por sólo nombrar unos pocos. Ese empeño de estandarización ha cristalizado en una serie de documentos de la Organización para la Cooperación y el Desarrollo Económico (OECD), marcando las líneas claves en la medición

Figura 3.1. Relevancia de la evaluación de políticas públicas de I+D+i



de las actividades relacionadas con la ciencia y la tecnología. Estos trabajos constituyen lo que se ha denominado la “familia de los Manuales de Frascati” y son los siguientes:

- a) El Manual de Frascati, que es la guía para la medición y cuantificación de actividades de investigación y desarrollo (I+D) (OECD, 2003).
- b) El Manual de Oslo, que pretende recoger las guías para la compilación e interpretación de los datos de innovación tecnológica. Éste ha sido preparado y publicado en cooperación con la Unión Europea (OECD/Eurostat, 2005).
- c) El Manual de Patentes, que ofrece los criterios para la medición de las actividades científicas y tecnológicas utilizando datos de las oficinas de patentes, ya que las estadísticas de patentes pueden utilizarse como indicadores de ciencia y tecnología (OECD, 2009).
- d) El Manual de Canberra, que se dedica a la medida de los recursos humanos dedicados a las actividades de ciencia y tecnología. Éste, como el de Oslo, ha sido preparado y publicado en cooperación con la Unión Europea (OECD/Eurostat, 1995).

Como señala la propia OECD (OECD/Eurostat, 2005, p. 4), el éxito de estos manuales puede ser atribuido al hecho de que, si bien los datos y los modelos teóricos que contienen no están exentos de limitaciones e imperfecciones, han demostrado ser altamente útiles para los analistas y para los decisores políticos.

De hecho, sin estos manuales resultaría imposible obtener datos estadísticos que pudieran ser comparables a nivel internacional. Y tales datos resultan un prerequisite para el diseño, seguimiento y evaluación de políticas destinadas a promover la innovación tecnológica, particularmente en el ámbito europeo, en el que el proceso de integración y unión económica y monetaria ha generado una necesidad creciente de disponer de datos internacionalmente comparables. Por ello, serán una fuente principal en nuestra investigación a la hora de seleccionar las variables de nuestro modelo conceptual causal y de los casos de estudio.

4. LA POLÍTICA DE CIENCIA, TECNOLOGÍA E INNOVACIÓN EN LA UNIÓN EUROPEA

4.1. ANTECEDENTES

La necesidad de coordinar esfuerzos en investigación de los países miembros estaba ya recogida en los primeros Tratados de las Comunidades Europeas. Aunque, entonces, la

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

investigación tenía un marcado enfoque sectorial, centrada en los campos del acero y del carbón, la energía y la agricultura.

En 1957, los mismos países que formaron la CECA firmaron en Roma el Tratado Constitutivo de la Comunidad Económica Europea (CEE) y el Tratado Constitutivo de la Comunidad Europea de la Energía Atómica (EURATOM). En ambos tratados, los objetivos fueron más económicos y políticos que técnicos y científicos. El tratado de EURATOM hacía referencia al deber de la Comunidad de desarrollar la investigación y asegurar la difusión de los conocimientos técnicos. Por su parte, el Tratado de la CEE sólo hacía referencia explícita a la investigación en agricultura.

Paralelamente, en el año 1954, se creó el primer centro de investigación de ámbito europeo, la Organización Europea para la Investigación Nuclear, cuyos objetivos se dirigían a la realización de investigación básica en el área de la física de partículas. Pero, el principal texto legislativo donde se recoge la política de I+D y se definen sus principales características es el Tratado de Funcionamiento de la Unión Europea (versión consolidada del Tratado de la Comunidad Europea, firmado en Roma el 25 de marzo de 1957 tras la última reforma realizada por el Tratado de Lisboa). En esta versión consolidada de 30 de marzo de 2010, las bases de la política científica y tecnológica se encuentran reguladas en el Título XIX, de los artículos 179 al 190 (antiguos artículos 163 al 173 del TCE).

Se trata de una competencia compartida de la Unión Europea: en los ámbitos de la investigación y el desarrollo tecnológico la Unión dispondrá de competencia para llevar a cabo acciones, en particular destinadas a definir y realizar programas, sin que el ejercicio de esta competencia pueda tener por efecto impedir a los Estados miembros ejercer la suya.

Finalmente, en la Carta de los Derechos Fundamentales de la Unión Europea (artículo 13) se establece que “las artes y la investigación científica son libres”. La Carta de los Derechos Fundamentales de la Unión Europea, que tiene carácter jurídicamente vinculante, confirma los derechos fundamentales garantizados por el Convenio Europeo para la Protección de los Derechos Humanos y de las Libertades Fundamentales, tal y como resultan de las tradiciones constitucionales comunes a los Estados miembros. De ahí la importancia de elevar a derecho fundamental la libertad de la investigación científica.

4.2. LOS PROGRAMAS MARCO DE INVESTIGACIÓN Y DESARROLLO TECNOLÓGICO

Los artículos 163 a 173 del Título XVIII del Tratado Constitutivo de la Comunidad Económica Europea, en la versión del Tratado de Ámsterdam, recogían los fundamentos de la política de investigación y desarrollo de la Comunidad. En ellos se establecía como objetivo fortalecer las bases científicas y tecnológicas de la industria europea y favorecer el desarrollo de su competitividad internacional. Para cumplir este objetivo, la Comunidad realizó un conjunto de acciones recogidas dentro de un Programa Marco plurianual que fijaba los objetivos científicos y tecnológicos a alcanzar.

A lo largo de los siete Programas Marco que han existido y que abarcarán 30 años al finalizar el periodo de vigencia del actual, se han producido considerables cambios en las prioridades de investigación y en los recursos dedicados a cada una de ellas. En un primer momento, el área de Energía concentró gran parte de los recursos como respuesta a la crisis del petróleo de principios de los 70. Aunque el interés en la investigación en energía se ha mantenido a lo largo de los siete programas, otras áreas como las tecnologías de la información y telecomunicaciones y el medio ambiente han visto incrementada tanto su dotación económica como su peso relativo.

Lo más destacable del penúltimo programa, el VI Programa Marco (2002-2006), fue la creación del Espacio Europeo de Investigación con el objetivo de reforzar la relación entre las iniciativas nacionales y las comunitarias, hacer frente al déficit de la balanza comercial tecnológica de Europa frente a sus competidores americanos y asiáticos y al reto de la incorporación a la Comunidad de países con menos capacidad tecnológica.

Respecto al vigente VII Programa Marco (2007-2013), en la historia de la Unión Europea no se han dado unas condiciones más favorables para el desarrollo de la ciencia y la tecnología como las que se han dado bajo su vigencia. Al “objetivo de Lisboa” (Consejo Europeo de Lisboa, 2000) de convertir a la UE en “la economía basada en el conocimiento más competitiva y dinámica del mundo” para el año 2010, y el “Compromiso de Barcelona” de dedicar el 3% del PIB europeo a la Investigación, Desarrollo e Innovación Tecnológica (I+D+I) para ese mismo año, se unieron unas perspectivas económicas favorables a aumentar los fondos europeos dedicados a I+D+i para el periodo de vigencia del VII Programa Marco Europeo (2007-2013).

4.3. ACTIVIDAD RECIENTE DE LA POLÍTICA CIENTÍFICA, TECNOLÓGICA Y DE INNOVACIÓN DE LA UE

4.3.1. Hitos históricos

Hasta hace relativamente poco tiempo, la política de la Unión Europea en el ámbito de la I+D se gestionaba solamente a nivel nacional, lo que suponía pérdida de oportunidades y de eficiencia. Pero, las actividades de investigación y desarrollo tecnológico e innovación resultan cruciales para el desarrollo económico de un país y así es entendido también en la UE, donde sus recursos de investigación en determinadas áreas clave y tecnologías prioritarias deberían incrementar la competitividad de las empresas y generar empleo a nivel comunitario.

En cuanto a los antecedentes históricos de este tipo de política, desde el establecimiento de la Comunidad Económica Europea y de la Comunidad Europea de la Energía Atómica (EURATOM), los principales hitos se pueden resumir cronológicamente de la siguiente forma:

- Programa Esprit (Programa Estratégico Europeo de Investigación sobre Tecnologías de la Información), 1984. Encaminado a fomentar la investigación básica sobre microelectrónica.
- Primer Programa Marco, para el período 1984-1987, al que han ido siguiendo los sucesivos programas hasta el Séptimo Programa Marco actual.
- Acta Única Europea, 1986. Sitúa la ciencia entre las responsabilidades comunitarias.
- Tratado de la Unión Europea (Tratado de Maastricht), 1993. Amplió el papel de la I+D en la UE y subrayó su importancia. Ese mismo año la Comisión publicó su Libro Blanco sobre crecimiento, competitividad y empleo, en el que señalaba la importancia de la investigación, la tecnología y las telecomunicaciones para mantener la competitividad de la industria.
- Libro Verde de la Innovación, 1995. Se publica este trabajo monográfico por parte de la Comisión Europa, que va a marcar la concepción y el desarrollo de la política científica y tecnológica.
- Comunicación “Hacia un Espacio Europeo de la Investigación”, año 2000. Reconocía que la UE debía mejorar sus esfuerzos en el campo de la investigación y la tecnología para seguir siendo competitiva en una economía cada vez más mundializada.

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

- Libro Verde, El Espacio Europeo de Investigación: nuevas perspectivas (Comisión Europea, 2007). Según la introducción de este documento, el Espacio Europeo de Investigación es más que nunca piedra angular de una sociedad europea del conocimiento, es decir, de una sociedad en la que la investigación, la educación, la formación y la innovación actúan con todo su potencial para cumplir las ambiciones económicas, sociales y medioambientales de la UE y las expectativas de sus ciudadanos.
- Tratado de Lisboa, 2009. El Tratado dota a la UE de instituciones modernas y perfecciona sus métodos de trabajo para poder afrontar con eficacia los desafíos del mundo de hoy y para resolver problemas como la globalización, el cambio climático, la evolución demográfica, la seguridad y la energía. El Tratado de Lisboa refuerza la democracia en la UE y mejora su capacidad de defender día a día los intereses de sus ciudadanos.
- La Estrategia Europa 2020. Se trata de la estrategia de crecimiento de la UE para la próxima década, cuya finalidad es doble: superar la crisis y subsanar los defectos de nuestro modelo de crecimiento creando las condiciones para un tipo de crecimiento distinto, más inteligente, sostenible e integrador. Para ello ha fijado cinco objetivos clave que la UE debe alcanzar en los ámbitos de empleo, educación, investigación e innovación, integración social y reducción de la pobreza, y cambio climático y energía, incluyendo siete iniciativas emblemáticas en áreas como la innovación, la economía digital, el empleo, la juventud, la política industrial, la pobreza y el uso eficiente de los recursos.

4.3.2. La Estrategia de Lisboa

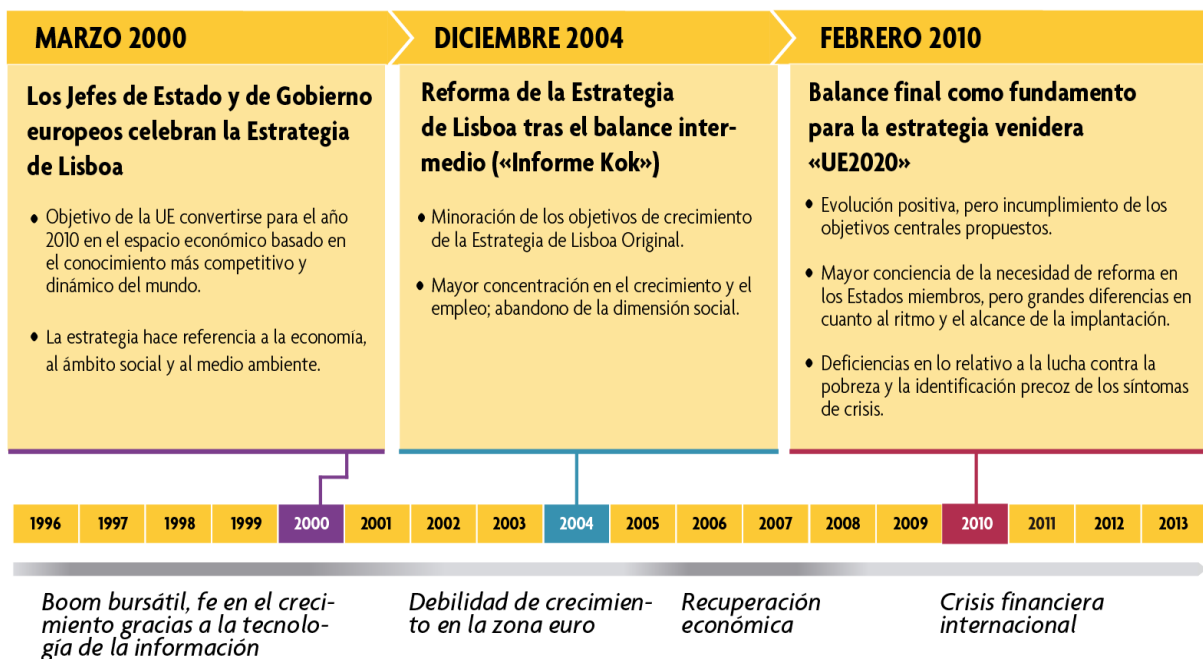
La Estrategia de Lisboa, también conocida como Agenda de Lisboa o Proceso de Lisboa, es un plan de desarrollo de la UE aprobado por el Consejo Europeo celebrado en Lisboa el 23 y el 24 de marzo de 2000 (Comisión Europea, 2000). Los Jefes de Gobierno de la UE se propusieron el siguiente planteamiento: pensaban convertir a la Unión en el «espacio económico basado en el conocimiento más competitivo y dinámico del mundo». Estas palabras surgen cuando la burbuja de las Tecnologías de la Información y la Comunicación (TICs) seguía siendo estable y la gran crisis aún estaba muy lejos. Hasta entrada de los 90, la UE (occidental reducida) había recuperado terreno respecto de los Estados Unidos y el *boom*

sostenido en los Estados Unidos parecía sugerir que el crecimiento que es impulsado por la innovación puede ser ilimitado (Fritz-Vannahme, 2010).

Transcurridos trece años, Europa no ha logrado el esperado objetivo. Antes de la crisis financiera de 2008, la tasa de empleo en la UE se hallaba en el 66% (frente al 70% establecido como objetivo), en Investigación y Desarrollo se invertía el 1,9% del PIB (se pretendía alcanzar el 3%) y el retraso de los niveles de productividad respecto a los Estados Unidos apenas se ha reducido⁸.

El hecho de que el balance final sea negativo se debe a numerosas razones. Entre ellas se encuentran el estallido de la burbuja de Internet, los atentados del 11 de septiembre de 2001 y las ampliaciones de la UE (que han dificultado el proceso ya que sólo había sido concebido para el grupo de los 15 de Europa Occidental). Además, todos los Estados de la Unión se ven inmersos desde 2008 en una crisis económica y financiera mundial.

Figura 4.1. Resumen de diez años de la Estrategia de Lisboa



Fuente: Adaptado de Fritz-Vannahme, J., et al. (2010).

⁸ Datos obtenidos de Eurostat ([www. http://ep.eurostat.ec.europa.eu](http://ep.eurostat.ec.europa.eu)).

Poco después de la aprobación de la Agenda las prioridades políticas empezaron a ser la ampliación pendiente de la Unión y el debate sobre una Constitución común. En la mayoría de los Estados miembros no se percibía una actuación política resuelta a favor de la Estrategia. Además, la Agenda estaba sobrecargada. El informe elaborado por un grupo de alto nivel liderado por Wim Kok (antiguo Primer Ministro neerlandés), así lo evaluaba provisionalmente a finales de 2004⁹. Apenas se hizo uso de las posibles sinergias con programas e instrumentos ya existentes, como el Pacto de Estabilidad y Crecimiento, la Agenda Social o la Estrategia de Desarrollo Sostenible.

A finales de 2010, concluye el proceso iniciado en el Consejo Europeo en Lisboa. En reemplazo de la Estrategia de Lisboa, se dio inicio a una nueva estrategia integrada que estará vigente hasta el año 2020 y en la que la idea fundamental del uso de indicadores de referencia sigue siendo empleada. En la figura 4.1, se muestran los principales hitos desde el año 2000 y la Estrategia de Lisboa.

4.3.3. La Estrategia Europa 2020 para un crecimiento inteligente

Según hemos visto, la Unión Europea necesitaba una nueva estrategia de crecimiento, empleo y proyección social ya que la Agenda de Lisboa no cumplió con las expectativas creadas en torno a la coordinación a nivel europeo. En la comunicación de la Estrategia Europa 2020 de la Comisión Europea, se afirma que la crisis es una llamada de atención, aunque la prioridad a corto plazo sea salir de la crisis. El propósito de Europa 2020 es conseguir un desarrollo sostenible, crear más empleo y lograr mayor bienestar y demostrar que “Europa es capaz de alcanzar un crecimiento inclusivo, sostenible e inteligente, de encontrar el modo de crear nuevos puestos de trabajo y de ofrecer una orientación a nuestras sociedades” (Comisión Europea, 2010).

La Comisión propone para ello cinco objetivos cuantificables para el año 2020: el empleo, la investigación y la innovación, el cambio climático y la energía, la educación y la lucha contra la pobreza. Se trata de objetivos ambiciosos, pero alcanzables. Para alcanzar el éxito, es indispensable que las instituciones y los líderes europeos hagan suyos estos objetivos y

⁹ El informe “Facing the challenge .The Lisbon strategy for growth and employment” (2004), se encuentra disponible en http://europa.eu.int/comm/lisbon_strategy/index_en.html.

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

requiere una respuesta europea coordinada, que incluya a los interlocutores sociales y la sociedad civil.

Europa se enfrenta a una crisis que ha echado por tierra años de progreso económico y social y expuesto las debilidades estructurales de la economía europea. Mientras tanto, los retos a largo plazo (mundialización, presión sobre los recursos, envejecimiento) se intensifican.

Así, la Estrategia Europa 2020 propone tres prioridades que se refuerzan mutuamente:

- Crecimiento inteligente: desarrollo de una economía basada en el conocimiento y la innovación.
- Crecimiento sostenible: promoción de una economía que haga un uso más eficaz de los recursos, que sea más verde y competitiva.
- Crecimiento integrador: fomento de una economía con alto nivel de empleo que tenga cohesión social y territorial.

Con este fin, la Comisión propone los siguientes objetivos principales:

- El 75 % de la población de entre 20 y 64 años debería estar empleada.
- El 3 % del PIB de la UE debería ser invertido en I+D+i.
- Debería alcanzarse el objetivo «20/20/20» en materia de clima y energía (incluido un incremento al 30 % de la reducción de emisiones si se dan las condiciones para ello).
- El porcentaje de abandono escolar debería ser inferior al 10 % y al menos el 40 % de la generación más joven debería tener estudios superiores completos.
- El riesgo de pobreza debería amenazar a 20 millones de personas menos.

Para garantizar que cada Estado miembro adapte la Estrategia Europa 2020 a su situación particular, la Comisión propone que los objetivos de la UE se traduzcan en objetivos y trayectorias nacionales.

Los objetivos son representativos de las tres prioridades de crecimiento inteligente, sostenible e integrador. Además, las siete iniciativas emblemáticas que propone la Comisión Europea para catalizar los avances en cada tema prioritario son (Comisión Europea, 2010):

- Unión por la innovación, con el fin de mejorar las condiciones generales y el acceso a la financiación para investigación e innovación y garantizar que las ideas innovadoras se puedan convertir en productos y servicios que generen crecimiento y empleo.

BLOQUE 1.- MARCO TEÓRICO Y CONCEPTUAL

- Juventud en movimiento, para mejorar los resultados de los sistemas educativos y facilitar la entrada de los jóvenes en el mercado de trabajo.
- Una agenda digital para Europa, con el fin de acelerar el despliegue de internet de alta velocidad y beneficiarse de un mercado único digital para las familias y empresas.
- Una Europa que utilice eficazmente los recursos, para ayudar a desligar crecimiento económico y utilización de recursos, apoyar el cambio hacia una economía con bajas emisiones de carbono, incrementar el uso de fuentes de energía renovables, modernizar nuestro sector del transporte y promover la eficacia energética.
- Una política industrial para la era de la mundialización, para mejorar el entorno empresarial, especialmente para la pequeña y mediana empresa, y apoyar el desarrollo de una base industrial fuerte y sostenible, capaz de competir a nivel mundial.
- Agenda de nuevas cualificaciones y empleos, para modernizar los mercados laborales y potenciar la autonomía de las personas mediante el desarrollo de capacidades a lo largo de su vida con el fin de aumentar la participación laboral y adecuar mejor la oferta y la demanda de trabajos, en particular mediante la movilidad laboral.
- Plataforma europea contra la pobreza, para garantizar la cohesión social y territorial de tal forma que los beneficios del crecimiento y del empleo sean ampliamente compartidos y las personas que sufren de pobreza y exclusión social pueden vivir dignamente y tomar parte activa en la sociedad.

Además, la estrategia Europa 2020 se basa en dos pilares fundamentales: el enfoque temático ya señalado, que combina prioridades y objetivos principales; e informes nacionales, que ayudarán a los Estados miembros a desarrollar sus estrategias para volver a un crecimiento y unas finanzas públicas sostenibles.

Tras este capítulo centrado en el papel de la I+D+i en la UE y su actual Estrategia 2020, se cierra el bloque introductorio que servirá de marco conceptual y teórico del presente trabajo. Se ha pretendido realizar un breve recorrido a través de la naturaleza y problemática del cambio tecnológico y la innovación como factores del crecimiento a lo largo de la Historia del Pensamiento Económico. Todo ello enfatizando aquellas que serán de utilidad (teorías endógenas) a la hora de diseñar el modelo conceptual causal que se presenta en el siguiente Bloque 2.

BLOQUE 2

MODELO

CONCEPTUAL

CAUSAL (RED

BAYESIANA)

All models are wrong, but some are useful.

George E. Box, 1979

*Remember that all models are wrong;
the practical question is how wrong do they
have to be to not be useful.*

Box and Draper, 2000

CAPÍTULO IV. LOS MODELOS CONCEPTUALES Y LA CAUSALIDAD. LAS REDES BAYESIANAS

1. MODELOS CONCEPTUALES EN GENERAL. LA CAUSALIDAD

Uno de los objetivos principales del presente trabajo de investigación es aportar un modelo conceptual causal, representado en forma de red Bayesiana, que llene el vacío existente de un modelo global en la evaluación y monitorización de las políticas públicas de I+D+i. Por ello, resulta conveniente realizar un breve repaso del concepto de modelo conceptual y de su utilidad para el seguimiento de políticas/programas, antes de entrar de lleno en la presentación de la metodología empleada y del modelo obtenido.

En general, la monitorización de políticas públicas se centra en la medición de los recursos en el tiempo, con la intención de ofrecer datos que reflejen cambios en el estado o en las tendencias del sistema o de sus componentes (modelos dinámicos). Considerando que los sistemas de I+D+i son sistemas inherentemente complejos y los gestores se enfrentan a cambios que ocurren en periodos de tiempo cada vez más largos, los modelos conceptuales y declarativos son usados universalmente para sintetizar y comunicar el entendimiento de la dinámica de un sistema, para identificar los componentes clave de dicho sistema y comunicar las interacciones entre los mismos (Gross, 2003).

Muchos de los errores en el diseño de políticas públicas podrían deberse a la falta de modelos conceptuales que identifiquen los elementos clave y señalen las relaciones entre ellos, de forma que se pueda “conocer” lo que sucedería si se influye de una u otra forma en estos elementos, como se destaca, p.e., en el ámbito del diseño de políticas medioambientales (Busch and Trexler 2003). En la fase inicial del seguimiento, unos modelos simples facilitarían un marco para relacionar la información y la literatura experta con un contexto más amplio.

En general, los modelos conceptuales son capaces de (Wu y David, 2002):

- Formalizar el conocimiento de los procesos y de la dinámica de un sistema.
- Identificar las conexiones entre procesos y los límites de diversas disciplinas.
- Identificar los límites y el objeto del sistema de interés y contribuir a su comunicación entre los distintos actores implicados, etc.

Una vez el programa o política está en funcionamiento, la interpretación adecuada de los indicadores de medición seleccionados es posible por las relaciones conocidas entre cada indicador y la función social o el recurso crítico que trata de representar (Kurtz et al. 2001). Estas conexiones deberían ser explícitas en los modelos conceptuales y su correcta articulación sería esencial para justificar los indicadores y su interpretación.

Respecto a la forma que pueden adoptar los modelos conceptuales, puede ser cualquier combinación de texto, tablas, matrices de factores, diagramas de cajas y flechas. Jorgensen (1988) ya hablaba de 10 clases de modelos y evalúa las ventajas y las desventajas de cada uno. Entre ellos, las tablas y las matrices facilitan un medio adecuado para resumir grandes cantidades de información, incluyendo interacciones entre los componentes. La mayoría de los programas de monitorización utilizan diagramas acompañados de narraciones que lo describen, justifican las relaciones entre las variables en la figura y citan las fuentes de información y los datos en los que se basa el modelo.

No hay un único modelo que sea definitivamente el correcto y puede ser incluso de interés explorar las formas alternativas de representar el objeto de estudio. Estas diferentes representaciones pueden ayudar a articular hipótesis para entender cómo opera el sistema. Por otro lado, un modelo universal sería demasiado complejo de entender. Un modelo conceptual útil articulará los procesos y variables más importantes contribuyendo a entender las interacciones y enlaces, facilitará la selección y justificación de las variables para el seguimiento, facilitará la evaluación de los datos, etc.

Un modelo conceptual se presenta como una ayuda a los evaluadores, decisores políticos y resto de grupos de interés. Se trata de una herramienta para visualizar qué fuerzas están operando, qué variables están influyendo y, en particular, las principales relaciones entre ellas. Además, estos modelos son útiles como herramientas de planificación porque ayudan a determinar qué acciones pueden influir más en la variable concreta que se quiera analizar de una forma visual y holística, qué factores se deberían monitorizar a la hora de implantar una política pública que afecta a determinados recursos, etc.

Respecto al tipo de modelo conceptual empleado en esta investigación, damos un paso más al añadir a nuestro modelo las relaciones de causalidad entre las variables seleccionadas. Entre los modelos conceptuales causales nos encontramos los denominados *mapas causales*, también conocidos como *mapas cognitivos*, que aparecen como gráficos directos que representan las relaciones causa-efecto incorporadas en el pensamiento de los expertos. Un mapa cognitivo

puede definirse como un “gráfico directo caracterizado por una estructura jerárquica que la mayoría de las veces toma la forma de un gráfico medios/fines” (Eden et al, 1992).

Los mapas causales expresan el juicio de que ciertos eventos o acciones conducirán a unos resultados particulares. Hay tres componentes principales en un mapa causal:

- El concepto de causalidad
- La conexión causal.
- El valor causal.

Hay que puntualizar que un modelo de red puede ser o un mapa dependiente (D-map) o un mapa de independencias (I-map) (Pearl, 1988). Un D-map garantiza que los vértices que se encuentran conectados son dependientes; sin embargo, puede mostrar un par de conceptos dependientes como un par de vértices separados. Un I-map, por el contrario, garantiza que los vértices o conceptos que se encuentran separados son condicionalmente independientes, dadas otras variables.

Dado que nuestro modelo conceptual pretende recoger las relaciones de causalidad entre las variables, resulta conveniente realizar algunos comentarios sobre la *Teoría de la Causalidad*. El área de la causalidad ha sido muy activa durante los pasados veinte años con numerosas interacciones entre los campos de la filosofía, la estadística y la ciencia de la computación. Esta actividad ha generado una animada controversia en una amplia variedad de temas conceptuales y metodológicos (McKim and Turner, 1997). La causalidad, como un postulado teórico, ha sido objeto de una gran discusión hasta la actual Teoría de la Causalidad de Pearl (Pearl, 2000), que es la que se adopta en este trabajo.

La causalidad probabilística es un paradigma que va desde la determinación absoluta de un efecto dada la ocurrencia de una causa hasta la ocurrencia de una causa que incrementa la probabilidad de que se produzca un determinado efecto. Una asunción subyacente de esta perspectiva es que el conocimiento incompleto de las causas da lugar a relaciones de tipo incierto de causa-efecto (Spirtes et al., 2000) o pseudo-determinismo.

La modelización causal en las Ciencias Sociales resulta de gran interés por su habilidad para combinar información de tipo causa-efecto con datos estadísticos y para proveer de una evaluación cuantitativa de relaciones entre las variables bajo estudio. Aunque no es objeto de esta investigación resolver el modelo causal, sino que se plantea como una de las líneas futuras de trabajo, nos limitamos a comentar en este punto que existen varias aproximaciones para

abordar un modelo causal como son las redes Bayesianas, la metodología de ecuaciones estructurales (SEM), como caso particular de las anteriores, etc.

Las redes Bayesianas modelan el conocimiento incierto para la toma de decisiones, empleadas como un enfoque probabilístico que realizará predicciones de lo que ocurrirá si cambiamos las variables bajo incertidumbre (Newton, 2010). Pueden considerarse un subconjunto de los diagramas de influencia que se centran en el componente de conocimiento incierto (Lazahnno et al., 2007).

Las SEM constituyen un enfoque de modelización causal que combina la información causa-efecto con datos estadísticos para facilitar una evaluación cuantitativa de las relaciones entre las variables estudiadas. Si las relaciones son significativas, la construcción teórica es considerada válida y se puede usar como guía para la aplicación del modelo en la práctica. Aunque esta metodología es adecuada para la validación empírica de relaciones causales teóricas y, en alguna medida, para la predicción, no es adecuada para el diagnóstico de la situación (Gupta and Kim, 2008).

Las RBs, y por extensión las SEM, cubren las aserciones de un modelo causal que puede ser expresado como $M=\{S, \Theta_S\}$, donde S es la estructura de las aserciones causales de un conjunto de variables V representado normalmente por un Gráfico Acíclico y Directo (Directed Acyclic Graph o DAG) y Θ_S es conjunto de parámetros compatible con S (Anderson and Vastag, 2004).

Las redes Bayesianas son especialmente adecuadas para la predicción y el diagnóstico y pueden ser entrenadas sobre la misma estructura con nuevos datos (intervenciones). Para representar nuestro modelo conceptual causal, se ha seleccionado el método de representación de red Bayesiana que describimos con mayor detalle en el siguiente apartado.

2. LAS REDES BAYESIANAS

Las redes bayesianas se desarrollaron a principios de los años 90 basadas en las teorías de los modelos gráficos, particularmente a través de los trabajos de Pearl (1988; 1995, 2000, 2009), quien estableció su posición en la interfaz entre la Estadística, la Inteligencia Artificial y el desarrollo de Sistemas Expertos.

Una red Bayesiana (RB), en general, es un modelo gráfico que incorpora las relaciones causales entre variables estocásticas de interés (nodos), que son típicamente representados en la forma de un diagrama de red (Jensen, 2001). Esta red puede ser usada para la representación del conocimiento experto y los arcos dirigidos (flechas) describen las relaciones.

Así, una RB consta de dos partes: una cualitativa, normalmente un gráfico acíclico directo (DAG) cuyos nodos representan variables aleatorias (continuas en nuestro caso); otra cuantitativa, que es una distribución de probabilidad condicionada para cada nodo o variables X_i , dados sus padres markovianos, $P(x_i|pa(x_i))$. La probabilidad de un nodo que no tuviera padres es su probabilidad *a priori*, $P(x_i)$. El conjunto de padres se representa mediante $pa(x_i)$. Un nodo X_i es un padre markoviano de X_j si y sólo si existe un enlace directo $X_i \rightarrow X_j$. En un DAG no están permitidos ni bucles de retroalimentación ni flechas o arcos con doble punta.

La probabilidad conjunta asociada a una red bayesiana satisface la condición de *d-separación* o independencia condicional (Pearl, 2009). Afirma que un nodo es independiente de los nodos que no son descendientes suyos dados sus padres markovianos. Esta propiedad implica en primer lugar que un enlace $X \rightarrow Y$ en una red Bayesiana representa una dependencia probabilística entre X e Y , lo que significa que la probabilidad de X afecta a la probabilidad de Y .

En segundo lugar, el criterio de *d-separación* implica que, considerando tres variables disjuntas (X, Y, Z) que representan nodos de un DAG G , para comprobar si X es independiente condicionalmente de Y dado Z , en cualquier distribución compatible con G , necesitamos comprobar si los nodos correspondientes a Z bloquean (interrumpen el flujo de información o dependencia) todos los caminos (secuencias de conexiones de cualquier sentido) desde los nodos de X a los de Y .

Un camino p está *d-separado* por un conjunto de nodos Z si y sólo si (Pearl, 2009):

- p contiene una cadena $i \rightarrow m \rightarrow j$ ó $i \leftarrow m \rightarrow j$ tal que m está en Z
- p contiene una cadena $i \rightarrow m \leftarrow j$ tal que m no está en Z y ninguno de sus descendientes tampoco

La ausencia de enlace representa independencia probabilística, es decir, que la probabilidad de X no afecta a la probabilidad de Y , y viceversa.

Como consecuencia de esta propiedad, se tiene que la distribución de probabilidad conjunta de una RB se puede obtener como el producto de las probabilidades de cada nodo

condicionadas en sus padres, lo que se conoce como el “teorema de factorización de la probabilidad” (Pearl, 1988). De esta forma, cualquier distribución que satisfaga:

$$P(x_i | pa(x_i)) = P(x_i | x_1, x_2, \dots, x_{i-1}) \quad (4.1)$$

Permite calcular la función de distribución de probabilidad conjunta P :

$$P(x_1, x_2, \dots, x_{i-1}) = \prod_i P(x_i | pa_i) \quad (4.2)$$

Desde un punto de vista matemático, una RB es un modelo para representar dependencias condicionales e independencias probabilísticas. En este caso, un enlace, considerado por sí mismo, no tiene ningún significado. Sin embargo, cuando se construye una RB como un modelo de un sistema del mundo real, un enlace $X \rightarrow Y$ es causal y, por tanto, X es una causa de Y .

La condición necesaria y suficiente para que una distribución de probabilidad P sea un DAG G (*Markov Relative*) es que, condicionada a sus parentales en G , cada variable debe ser independiente de todos sus predecesores en cierto orden de las variables que concuerda con las flechas en G (Pearl, 2009).

La condición suficiente y necesaria para que una distribución de probabilidad P sea un DAG G (*Markov relative*) es que todas las variables deben ser independientes de todos sus no-descendientes en G , de forma condicionada a sus parentales (Pearl, 2009).

Por todo lo expuesto, podemos afirmar que nuestro modelo conceptual está representado mediante una Red Bayesiana causal.

Las razones para utilizar modelos causales como el propuesto son múltiples:

- Los seres humanos tendemos a interpretar los hechos en términos de relaciones causa-efecto (Pearl, 2009). Por tanto, estos modelos son más fáciles de construir y de modificar (Pennington y Hastie, 1988), así como más fáciles de entender (Pearl, 2009, Pennington y Hastie, 1988).
- Muchos dominios de conocimiento reales se organizan en forma de jerarquías causales (Pearl, 2009).
- La identificación de relaciones causales que no varían permite predecir los efectos tanto de causas espontáneas (que corresponden a variables aleatorias) como de acciones (llamadas intervenciones) (Pearl, 1999).

BLOQUE 2- MODELO CONCEPTUAL CAUSAL (RED BAYESIANA)

- Estos modelos aportan más información que los modelos meramente probabilísticos. Una distribución conjunta aporta información sobre las probabilidades de ciertos eventos y sobre cómo varían después de distintas observaciones; un modelo causal nos indica cómo varían las probabilidades como resultado de intervenciones externas (Pearl, 2009).
- Las propiedades axiomáticas de las redes bayesianas (*d-separación* y propiedad de Markov) corresponden a dependencias e independencias probabilísticas que aparecen en dominios causales (Díez, 1999).
- Las RBs causales proporcionan determinados modelos de razonamiento cualitativo que pueden ser identificados con el objetivo de explicar los resultados de la inferencia (Druzdzel, 1996).
- Aunque hasta hace poco tiempo la causalidad era un concepto vago y abstracto, la teoría de la causalidad tiene una semántica bien definida y una lógica bien fundada (Pearl, 2009).

Las probabilidades iniciales en una red Bayesiana pueden ser mejoradas iterativamente (intervenciones) con nueva información para producir probabilidades posteriores, habilitándola para ser actualizada en respuesta a la información adicional (Charniak, 1991).

Otro aspecto importante de las RBs es el rol que juegan en la formalización de la causalidad y para modelar creencias probabilísticas motivadas por la necesidad de realizar decisiones racionales (Darwiche, 2010). Las RBs son excelentes modelos conceptuales causales para la predicción con las que se podrían usar técnicas de simulación (simulación Montecarlo) y motores de inferencia difusa que evalúen las relaciones de causalidad imprecisas (García-Alonso et al., 2013).

Algunas de las desventajas de las RBs incluyen: la dificultad del conocimiento de las probabilidades para realizar predicciones, que puede ser difícil de manejar debido a las variables aleatorias continuas y, finalmente, que puede no contener las variables apropiadas si el conocimiento explícito previo era erróneo. Además, a medida que los modelos se hacen más complejos, su tratamiento computacional decrece de forma dramática. Actualmente, no existe un modelo Bayesiano que trate con una cantidad de inputs elevada dado que el espacio de búsqueda es simplemente demasiado grande para ser tratable. Las mejoras en hardware y campos como el Aprendizaje Automático han reducido esta limitación pero, por ahora, la

generación de predicciones precisas con ingentes cantidades de datos resulta difícil y costosa (Perfors et al., 2011).

Respecto a su uso, las RBs se emplean en un amplio número de campos incluyendo el desarrollo de Sistemas Expertos, diagnóstico médico, Ciencia de la Computación, Minería de Datos, Aprendizaje Automático, Ciencia Cognitiva y Teoría de la Decisión. Además, son empleadas para construir diagramas de influencia, herramientas de toma de decisiones y clasificadores, siendo muy útiles en el contexto de evaluación del riesgo (Borsuk et al., 2004). Finalmente, recientes revisiones destacan su empleo como herramientas de modelaje medioambiental (Aguilera, 2011).

Dicho modelo nos servirá de base para la realización de los análisis empíricos planteados en el Bloque 3. Y es que, a pesar de los intentos realizados en el campo objeto de estudio, el problema de definir un modelo riguroso para la medición de la ejecución de las políticas públicas de I+D+i no ha sido resuelto aún. Aunque una Red Bayesiana (RB) es la principal aproximación para representar un modelo conceptual causal (Pearl, 2000; Darwiche, 2010), el resultado obtenido no ha sido empleado en la evaluación probabilística de políticas públicas de I+D+i. Sí existen algunos ejemplos de la utilización de RBs simples en la evaluación del impacto de las variables en el proceso de análisis de la colaboración en la innovación (Kim and Park, 2008), sistemas de clasificación como los empleados en la cibermetría (Lamirel et al., 2004) y en la evaluación de la eficiencia técnica relativa de proyectos de I+D (Ruiz et al., 2009).

CAPÍTULO V. DISEÑO DEL MODELO CONCEPTUAL: EL ANÁLISIS COOPERATIVO BASADO EN CONOCIMIENTO EXPERTO (EBCA)

En este capítulo, se detalla la metodología empleada tanto en la construcción y en la mejora del diseño del modelo conceptual causal representado mediante una red Bayesiana como en la incorporación del conocimiento experto a la metodología DEA empleada para el análisis de la eficiencia técnica relativa de las políticas públicas de I+D+i.

La construcción de una red bayesiana implica, principalmente, tres tareas:

- 1.- Identificar las variables.
- 2.- Identificar las relaciones causales entre las variables, lo que completa el gráfico que representa el modelo.
- 3.- Obtener las probabilidades asociadas a cada nodo/variable del grafo y las probabilidades condicionadas de las relaciones causales.

No es objetivo de esta investigación resolver la RB, lo cual sería una tarea muy difícil computacionalmente si consideramos el número de variables implicadas y el que dichas variables son aleatorias y continuas, por lo que no se llevó a cabo la tercera tarea expuesta, quedando como planteamiento de una futura investigación.

En general, las dos primeras tareas pueden realizarse desde dos enfoques principales: el enfoque basado en el análisis de datos y el enfoque basado en el conocimiento experto. Los enfoques basados en los datos usan semánticas de independencia condicional de redes de Bayes para inducir modelos de los datos (Heckerman, 1996). El enfoque basado en el conocimiento usa el conocimiento del dominio de los expertos en su construcción (Laskey y Mahoney, 1997), siendo especialmente útil en situaciones donde el dominio del conocimiento es crucial y la disponibilidad de los datos es escasa. La elicitación del conocimiento cualitativo por parte del ser humano es crítico en la construcción de RBs porque éste encuentra más fácil manejar datos cualitativos que cuantitativos (Nadkarni y Shenoy, 2001).

El segundo enfoque será el empleado para el diseño del modelo conceptual causal en forma de red Bayesiana, es decir, a través de la revisión de la literatura experta y de consultas a expertos.

Una vez diseñado el modelo tras la construcción de una base de conocimiento experto, éste se mejorará y perfeccionará gracias a la aplicación de la metodología que se detalla en los siguientes párrafos.

Obviamente, las actividades de I+D+i pertenecen al ámbito de la actividad humana y se encuentran dominadas por tareas cognitivas y algunos enfoques defienden las ventajas de incorporar conocimiento experto inicial en el propio análisis en sí mismo (Gibert y Sonicki, 1999; Salvador-Carulla et al., 2010). Siguiendo este paradigma, empleamos un nuevo enfoque metodológico denominado Análisis Cooperativo Basado en Conocimiento Experto (Expert knowledge-based Cooperative Analysis o EbCA; Gibert et al., 2010).

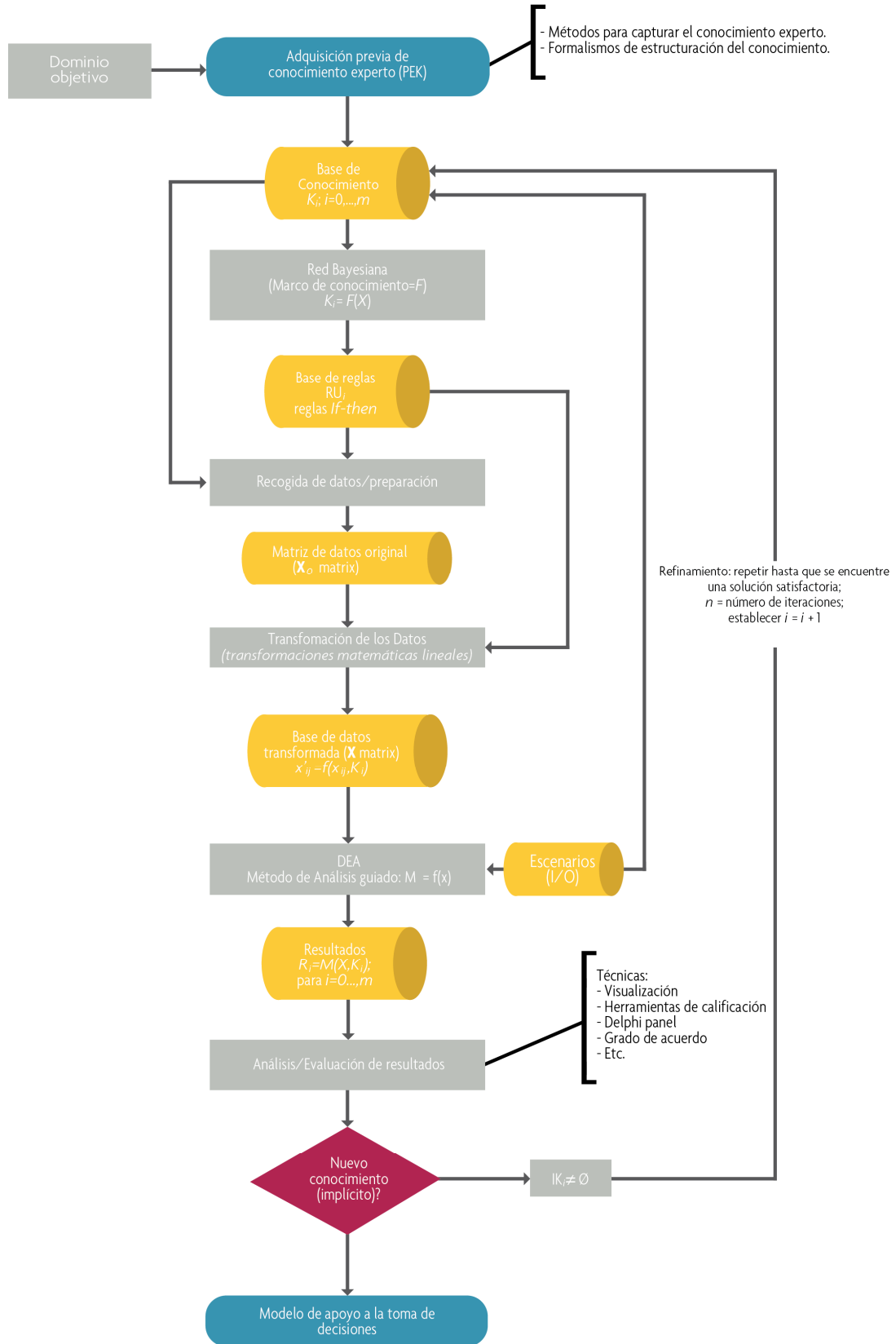
Este enfoque consiste en un marco general adaptable a la investigación de problemas muy complejos donde los enfoques clásicos arrojan pobres e insatisfactorios resultados (Gibert et al., 2010). La metodología EbCA se basa en construir una base de conocimiento experto, representarla, e incluirla formalmente en el método de análisis concreto que se emplee en el estudio de un problema, formando parte esencial del mismo.

De acuerdo con los autores, tener en cuenta el conocimiento experto previo no implica necesariamente una pérdida de rigor científico ya que existen marcos rigurosos de trabajo para tratar con este tipo de conocimiento. Se tiene en cuenta el conocimiento previo de los expertos desde el principio y, siguiendo técnicas estadísticas clásicas, los conduce a elicitar nuevo conocimiento (conocimiento implícito, no incluido en la base de conocimiento experto inicial) de forma que podemos mejorar el modelo conceptual causal propuesto.

En nuestro caso, aplicaremos la metodología EbCA al Análisis Envolvente de Datos o DEA (método de análisis seleccionado en el caso de estudio número 1 del Bloque 3), incorporando el conocimiento experto explícito, que se encuentra recogido de forma estructurada en la red Bayesiana, a dicha metodología. Se considera que completar el método DEA clásico con la metodología EbCA ayuda al análisis de problemas decisionales complejos tanto de evaluación de políticas públicas de I+D+i como del proceso de innovación, tratándolos como un todo y reduciendo de esta forma la incertidumbre asociada a dichos procesos.

Siguiendo esta metodología, se obtiene como producto final del proceso un modelo mejorado para la toma de decisiones y para el análisis de fenómenos complejos y una metodología que incorpora el conocimiento experto.

Figura 5.1. Metodología propuesta para la mejora del diseño del modelo conceptual



Fuente: Elaboración propia a partir de la metodología propuesta por Gibert et al. (2010).

Esto representa una relajación del enfoque de sistemas expertos clásicos basados en el conocimiento, que dependen de construir una correcta y completa descripción del dominio del conocimiento antes del análisis (Gibert et al., 2010).

De acuerdo con Gibert et al. (2010), los pasos seguidos por esta metodología, tanto para guiar el análisis DEA de la eficiencia técnica relativa de las políticas públicas de I+D+i como para mejorar el diseño del modelo conceptual causal (RB) tras analizar los resultados de eficiencia, son los siguientes:

1. Adquisición de conocimiento experto previo (PEK, prior expert knowledge, estableciendo $i=0$ en el diagrama de flujo). Es el nombre dado al proceso de adquirir y representar conocimiento en un dominio especializado de interés (Dialliwal y Benbasat, 1990). Existe un gran número de técnicas y herramientas tanto para la adquisición de conocimiento experto como para su representación (Boose, 1989; Olson and Reuter, 1987). Las técnicas empleadas en este caso son la revisión de la literatura especializada y las entrevistas y consultas a expertos (conocimiento explícito previo). En este paso, la respuesta de los expertos es transcrita para obtener un texto que denominamos “narrativo”. Las narraciones fueron recopiladas y representadas en el modelo conceptual-causal que fue presentado de nuevo a los mismos expertos. El mapa causal obtenido se modifica para conseguir que sea compatible con un DAG.
2. Así, una vez obtenida la base de conocimiento del dominio, el *doctrinae corpus* (K_i) pasa a ser representado bajo el marco general (F), que se trata de un modelo conceptual causal en forma de RB en nuestro caso, definiendo el grafo que representa las variables del problema y las relaciones de dependencia causal entre ellas. Se construye así el modelo (nodos y relaciones) con toda la información del apartado anterior. Estos dos primeros pasos conformarían la construcción manual del modelo inicial de RB.
3. Recopilación de datos de las fuentes de datos oficiales disponibles (base de datos primaria \mathbf{X}_0). En este paso general, se determina la estructura del conjunto de datos, seleccionamos las variables más relevantes y se procede al tratamiento de los datos (lo que incluye su representación para el análisis de outliers, de datos perdidos, etc., escogiendo las técnicas apropiadas para cada caso).
4. De acuerdo al conocimiento K_i representado en el modelo, a nuevas consultas a expertos y la revisión de literatura, se construye una base de reglas de conocimiento

(RU_i). Esta base RU_i puede tener más de un antecedente (variable) pero sólo un consecuente. Formalmente, RU_i se representa de la siguiente forma (iteración $i=0$).

$$K_i : \underset{j=1}{\overset{c}{\text{ALSO}}} [\underset{k=1}{\overset{K}{\text{IF AND}}} (x_{ik} \in X_k \text{ isr } A_{jk}) \underset{k=1}{\overset{K}{\text{THEN AND}}} [x'_{ik} = f(x_{ik})] \quad (5.1)$$

donde c es el número de reglas en K_i , *isr* es el operador relacional que identifica el correspondiente conjunto A_{jk} y f es una transformación matemática (lineal, creciente o decreciente en nuestro caso) sobre cada uno de los datos originales x_{ik} para obtener valores transformados que incorporan el conocimiento experto, x'_{ik} . Las bases de conocimiento representadas en modelos basados en reglas se ajustan bien en la integración e interpretación de diferentes fuentes de conocimiento (Starfield et al., 1989).

5. Aplicación de la base de reglas al conjunto de datos original. El conjunto de reglas RU_i expresadas bajo la sintaxis *If <condition> then <variable transformation>*, como se muestra en la ecuación (5.1), es aplicada a \mathbf{X}_0 obteniéndose la base de datos transformada \mathbf{X} . De esta forma, se lleva a cabo una transferencia de conocimiento experto a la computadora de forma que pueda “comprender” mejor el significado de los datos.
6. Aplicación del método (M) a la base de datos \mathbf{X} transformada. Seguidamente, K_i es empleado para guiar el análisis de la matriz de datos transformados \mathbf{X} dependiendo del método subyacente seleccionado M . Definimos una transformación de M de forma que K_i se tome en consideración y se produzcan nuevos resultados, $R'_i = M(\mathbf{X}, K_i) = R_i | K_i$. Se obtiene M' , un nuevo método que no es más que la generalización de M y que está disponible para futuros usos. En nuestro trabajo, nuestro M sería el método DEA del que los expertos interpretarían los resultados, que ayudará a encontrar incoherencias y mejorará el diseño de nuestro modelo conceptual ya que empleamos la base de datos transformada \mathbf{X} , que incorpora el conocimiento K_i al aplicarse a éste el conjunto de reglas RU_i . Esto significa que $M'(x) = M(f(X))$.
7. Análisis de los resultados (R_i) para la comprobación de sus posibles incoherencias y obtención de nuevo conocimiento no explicitado en la fase inicial.
8. De forma iterativa, a medida que se obtiene nuevo conocimiento implícito, se vuelve al paso uno, modificando el modelo en concordancia con el nuevo conocimiento. La solución se considera satisfactoria si $IK_i = \emptyset$ o $R'_i = R_{i-1}$. Si la solución es satisfactoria, vamos al siguiente paso (código de decisión en la figura 5.1). De otra forma,

BLOQUE 2- MODELO CONCEPTUAL CAUSAL (RED BAYESIANA)

incorporamos IK_i en la base K_i , obteniendo K_{i+1} , formulando esta base de conocimiento nueva en el mismo marco de representación F . Fijamos $i=i+1$ y se repite la secuencia desde el paso 1.

9. Una vez se obtienen resultados adecuados y no se elicitó más conocimiento implícito (es decir, $IK_i = \emptyset$) tras las iteraciones necesarias, se obtiene el modelo definitivo, la última base de reglas aplicada y la última base de datos transformada.
10. Esta base de datos transformada (**X**) será la que permita obtener los resultados definitivos de eficiencia técnica relativa de los 27 EEMM de la UE en el caso de estudio número 1 del Bloque 3 experimental. El modelo causal obtenido servirá de base para la selección de variables del resto de casos de estudio del Bloque experimental (Bloque 3).

Respecto a los dos primeros pasos de la metodología, donde se construye de forma manual el modelo inicial de RB, el principal escollo surgió debido a la escasa literatura sobre el tema, dificultad que se ha tratado de soslayar con el uso de la presente metodología EbCA.

Respecto a la identificación de las variables, un primer modelo validado por los expertos nos sirvió de referencia, y en etapas sucesivas, se fueron eliminando y añadiendo nodos hasta obtener la versión final. Un problema adicional fue decidir el número de variables en la red: demasiado grande dificultan su evaluación y aplicación, aunque el resultado puede ser más preciso al aproximarse a la realidad. Si es demasiado pequeño, puede que no represente adecuadamente la realidad. Nuestro modelo incluye un gran número de variables (52) dado que, como se apuntó, no es objetivo de esta investigación resolver probabilísticamente la red.

Mediante el proceso iterativo propuesto por la metodología EbCA, no sólo se construyó la base inicial de conocimiento experto, sino que a través de refinamientos sucesivos tras analizar los resultados correspondientes al método seleccionado (DEA) para evaluar la eficiencia técnica relativa de las políticas de I+D+i en la UE, se obtuvieron distintas versiones de la red hasta alcanzar la versión definitiva mejorada (ver capítulo 6, Figura 6.6). El resto de pasos de la metodología son empleados en el análisis empírico (capítulo VII, Bloque 3).

CAPÍTULO VI: PRESENTACIÓN DEL MODELO CONCEPTUAL CAUSAL (RED BAYESIANA)

1. BASE TEÓRICA Y CONCEPTUAL

Uno de los aspectos más útiles de los modelos conceptuales a la hora de realizar tareas de evaluación es que pueden ayudar a determinar lo que medir y, lo que es más importante, lo que no medir. Un modelo bien desarrollado en el ámbito de la evaluación de políticas públicas debería representar el conocimiento experto y explícitamente enlazar los objetivos públicos a las principales amenazas y oportunidades que les afectan.

Idealmente, los diseñadores de políticas públicas deberían desarrollar modelos conceptuales para la planificación antes de que se diseñen y se implementen dichas políticas. Estos modelos conceptuales pueden facilitar a los evaluadores las tareas de seguimiento. Sin embargo, las evaluaciones a menudo miden lo que se alcanza frente a lo que los gestores de proyectos esperan lograr (i.e., objetivos y metas de los proyectos) (Margolis et al., 2009).

Las principales prácticas encontradas para evaluar los impactos públicos de las actividades de I+D+i van desde estudios econométricos, modelos generales de equilibrio, que descansan sobre un gran número de asunciones, etc., hasta el uso de indicadores compuestos (Grupp, 2010; Freeman y Soete, 2009), estudios basados en encuestas y estudios de caso¹⁰.

A pesar de la variedad de prácticas citadas, hay pocos estudios que consideren la inversión en I+D+i desde una perspectiva macroeconómica (David and Hall, 2000), cuando puede ser una herramienta útil de apoyo a la toma de decisiones diseñar un modelo conceptual que identifique las variables clave y sus relaciones en los procesos de I+D+i. El principal escollo para el diseño de un modelo conceptual de este tipo es, precisamente, la existencia de múltiples variables e interacciones entre ellas en un medio socioeconómico incierto. Además, las aproximaciones clásicas de modelado presentan muchas dificultades para capturar la estructura completa del dominio bajo estudio (Gibert et al., 2010), particularmente cuando la información es incompleta. Los formalismos algebraicos clásicos no tienen en cuenta el acervo

¹⁰ Para una revisión, ver Fahrenkrog, G.; Polt, W.; Rojo, J.; Tübke, A.; Zinöcker, K. (2002). RTD Evaluation toolbox, assessing the socioeconomic impact of RTD policies. Seville, Joint Research Centre, Institute for Prospective Technological Studies.

de conocimiento necesario para diseñar y desarrollar un modelo conceptual que incluya tanto aspectos cuantitativos como cualitativos del fenómeno causal (Pearl, 2000).

El diseño de un modelo conceptual, como se ha expuesto, es útil para identificar las variables clave y sus relaciones a la hora de evaluar los impactos socioeconómicos de los procesos de I+D+i y de los recursos empleados para ello. Enlazando con el marco teórico recogido en el Bloque 1, la investigación asume varias teorías o concepciones básicas:

La teoría del crecimiento económico endógeno. La RB se basó en las teorías económicas que incorporan el cambio tecnológico como factor endógeno en el proceso de crecimiento y revisadas en el Bloque 1. Estas teorías inicialmente dieron énfasis al papel del conocimiento y al capital humano como un input de la producción. Desde entonces, numerosos estudios empíricos exploraron como y en qué medida la innovación y el cambio tecnológico, considerados como un resultado de las actividades e inversiones en I+D, podía contribuir al crecimiento económico.

El modelo del proceso de innovación. El modelo lineal de innovación (Balconi et al., 2010) asume que un descubrimiento procedente de la investigación es seguido de actividades de desarrollo como un proceso secuencial hasta que la introducción satisfactoria de la innovación en el mercado sucede. Por el contrario, el modelo encadenado de la innovación (Kline and Rosenberg, 1986) y modelos más complejos como el modelo cíclico (Padmore, 1998) establecieron que el elemento clave del proceso de la innovación era que la investigación, el desarrollo y la innovación están unidos de una forma no lineal e interactiva. A pesar de ello, se adoptó la asunción de la linealidad en esta investigación porque los responsables del diseño de políticas públicas siguen entendiendo que las inversiones en I+D generan innovaciones de una forma lineal y, consecuentemente, las bases de datos oficiales son diseñadas siguiendo este modelo (Godin, 2006).

La tradición Schumpeteriana. El llamado paradigma Schumpeteriano (Aghion and Howitt, 1997), que surgió de la teoría moderna de la organización industrial y que colocó a empresas y emprendedores en el corazón del proceso de crecimiento, resulta un marco útil sobre el proceso de crecimiento. El trabajo de Schumpeter (1934) estableció cinco tipos de innovación, recogidos a su vez en el Manual de Oslo (OECD, 2010): organizacional, de producto, de mercado, de proceso y de los inputs. Además, analizó la influencia del tamaño de las empresas y la estructura del mercado sobre la innovación. Esta clasificación se ha tenido en cuenta a la hora de diseñar la red Bayesiana.

Las interacciones entre los agentes (modelo de Quintuple Hélice). En un primer momento, el modelo de Triple Hélice (Etzkowitz and Leydesdorff, 1998) fue considerado como un modelo que engarza sinergias entre los tres actores principales de la I+D+i: las universidades, las industrias y los gobiernos, éstos últimos como responsables del diseño de políticas. El modelo de Cuádruple Hélice añade una cuarta hélice como es “la sociedad pública y civil basada en la cultura y los medios de comunicación”. El modelo de Quintuple hélice añade la perspectiva del “medio ambiente natural de la sociedad” (Carayanis and Campbell, 2010). Este último es el que asumimos en el diseño de la red ya que representa la transición socio-ecológica desde la economía del conocimiento.

2. VARIABLES (NODOS) Y ESTRUCTURA DEL MODELO CONCEPTUAL CAUSAL PROPUESTO

Siguiendo la metodología establecida, se diseñó y desarrolló la RB (figura 6.6) como modelo conceptual de acuerdo al marco teórico del Bloque 1, al estudio de los Manuales elaborados por la OECD, a los informes oficiales de la Comisión Europea (Comisión Europea, 2011), a la literatura científica seleccionada (en especial, ver Godin, 2002 y 2007) y a la opinión del panel de expertos. Así, se identificaron las variables relevantes para los procesos de I+D+i (ver tabla 6.1)¹¹, así como sus relaciones de causalidad, como queda representado gráficamente en el modelo conceptual causal (RB) que se muestra en la figura 6.6. En este punto, resulta fundamental insistir en que el modelo final obtenido puede considerarse también como un resultado del caso de estudio nº 1 (Capítulo VII), al ser mejorado tras el análisis de los resultados de eficiencia técnica relativa por parte de los expertos. Sin embargo, se considera necesario incluirlo en este bloque de forma separada por tratarse la base del análisis experimental.

Para el estudio de la eficiencia técnica relativa de las políticas públicas de I+D+i en los EEMM de la UE, se empleó el método DEA en dos pasos enlazados (2LS-DEA, de las siglas en inglés two linked steps-Data Envelopment Analysis). El análisis de esta eficiencia implicó la selección de los inputs y outputs apropiados (variables en la RB) en cada paso y para cada escenario bajo estudio. Se denomina “escenario” a una selección específica de inputs y outputs que muestra una visión particular de la eficiencia de cada *Decision Making Unit* o DMU (esta

¹¹ Los indicadores seleccionados deben entenderse como representativos de las variables y no como idénticos a los outputs de la I+D+i (*proxy*).

unidad de decisión se corresponde con cada una de las políticas de I+D+i de los países de la UE) según la terminología del método DEA.

De la RB se seleccionarán, por tanto, las variables que se usarán como inputs y outputs de los escenarios socioeconómicos en el análisis 2LS-DEA. Por motivos operativos y descriptivos, se hará referencia a los escenarios simultáneamente a la descripción del contenido del modelo de RB cuando resulte procedente. Todos los inputs y los outputs son variables que se recogen en la RB, pero no todas las variables de la RB son utilizadas como inputs y outputs en los escenarios seleccionados. Una descripción detallada de estos inputs y outputs por escenario puede encontrarse en la tabla 6.1. El resto de variables incluidas en la RB (no seleccionadas) queda recogido en la tabla 6.2.

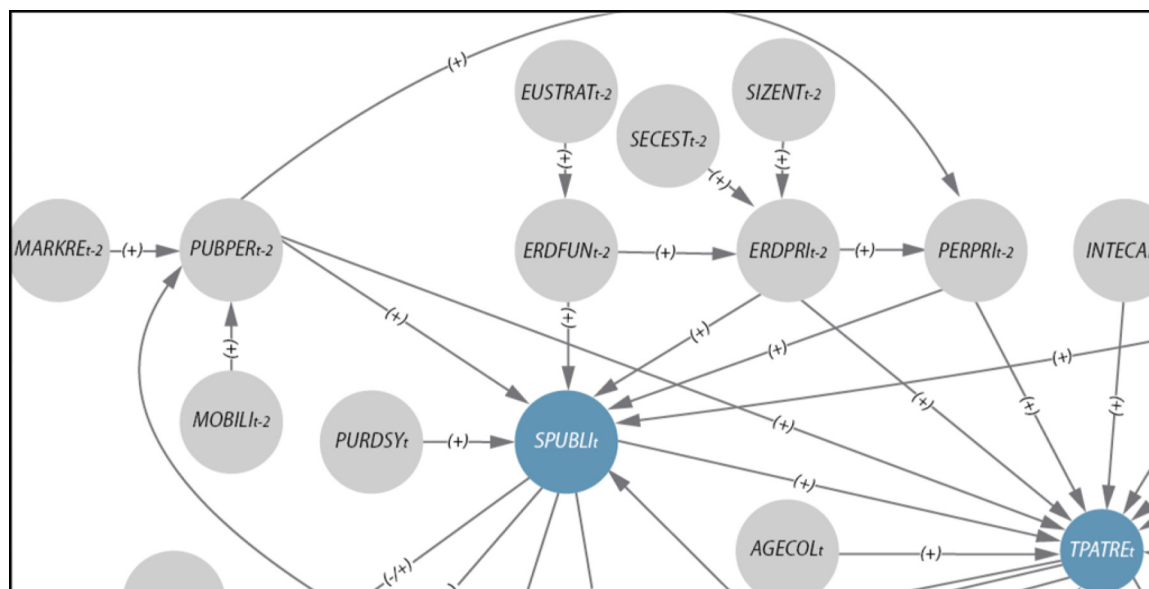
En el modelo se considera el conocido como *time-lag* o retardo (Greiff, 1985), representado por el subíndice temporal $t-2$ en la variable correspondiente. Dicho retardo trata de representar el hecho de que las inversiones en I+D+i reflejan sus resultados dos años de media más tarde a la realización de dichas inversiones.

De esta forma, y pasando a describir el diseño gráfico del modelo conceptual causal, prestamos atención al denominado escenario Tradicional – el escenario analizado en el primer paso del 2LS-DEA. Los principales indicadores *proxy* para medir la innovación (identificados por la RB) son el gasto en I+D+i y el personal en I+D+i. El primero refleja el grado de importancia que los agentes otorgan a las actividades de I+D+i y su capacidad para generar innovaciones. El segundo refleja el número de personas implicadas o dedicadas en la investigación fundamental, la investigación aplicada y el desarrollo experimental (Comisión Europea, 2010; OECD, 2010).

Estos indicadores se diferencian a su vez en sector público ($ERDFUN_{t,2}$ para representar los fondos públicos en I+D+i y $PUBPER_{t,2}$ para representar el personal público dedicado a la I+D+i como porcentaje del empleo total) y en sector privado (la cantidad de gasto del sector empresarial en I+D+i, $ERDPRI_{t,2}$, y el personal dedicado a la I+D+i en este sector, $PERPRI_{t,2}$). Para nuestro propósito, escogeremos sólo las variables de la vertiente pública ($ERDFUN_{t,2}$ y $PUBPER_{t,2}$). La variable $ERDFUN_{t,2}$ se encuentra condicionada por las líneas maestras de la estrategia europea existente ($EUSTRAT_{t,2}$). Por otro lado, $PUBPER_{t,2}$ depende de: i) el nivel de movilidad de los investigadores y de los estudiantes de doctorado ($MOBILI_{t,2}$) y ii) el atractivo del mercado de investigadores, básicamente debido al nivel de los salarios del país de

destino ($MARKRE_{t,2}$). Para facilitar la explicación, se detalla la parte de la red que refleja estas variables en la siguiente figura 6.1:

Figura 6.1. Representación del escenario Tradicional en la red Bayesiana



Fuente: Elaboración propia.

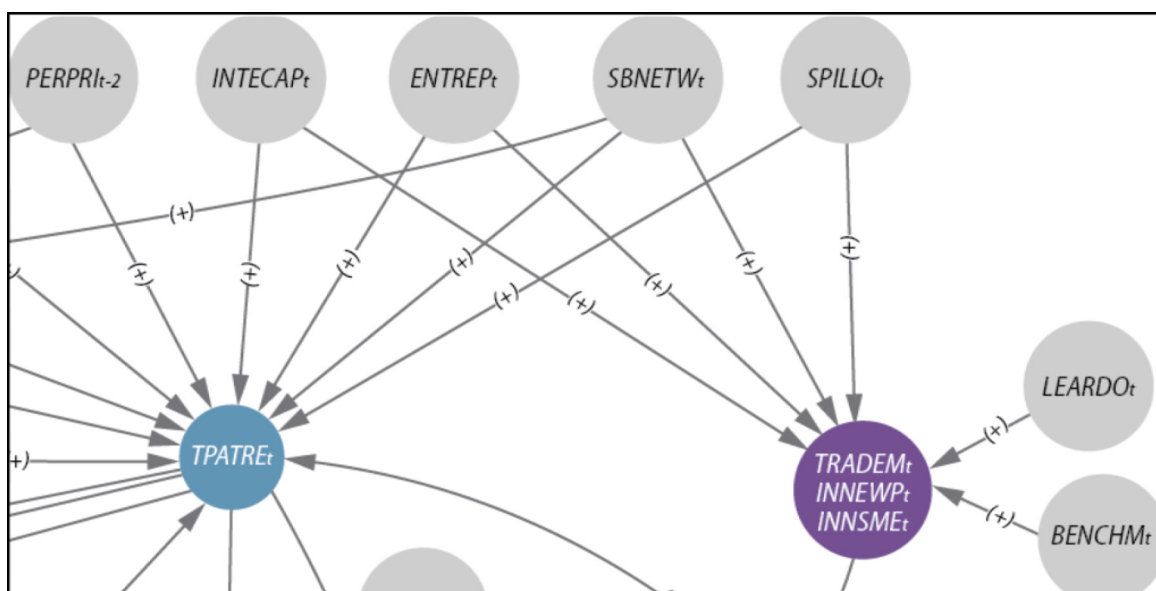
Para el sector privado, el gasto privado en I+D+i ($ERDPRI_{t,2}$) está condicionado por el apoyo público o $ERDFUN_{t,2}$ (en forma de subsidios, incentivos fiscales, etc.), por el tamaño de las empresas ($SIZENT_{t,2}$), y por la estructura del sector privado en el correspondiente país ($SECEST_{t,2}$) (Yu, 2005). El personal implicado en las actividades de I+D+i ($PERPRI_{t,2}$) proviene principalmente del personal recogido en la variable $PUBPER_{t,2}$ y, lógicamente, también depende del gasto privado en I+D, $ERDPRI_{t,2}$.

En el escenario Tradicional, los principales outputs considerados en la literatura son el número de publicaciones científicas ($SPUBLI_t$) y las patentes concedidas ($TPATRE_t$) (OECD, 2010). Ambas variables dependen también, aparte de sus correspondientes inputs, de las infraestructuras generales del país ($INFRAE_t$) que incluyen la maquinaria, el equipamiento, todo tipo de infraestructuras de comunicación, vehículos, utilería y otros edificios (laboratorios y otras infraestructuras científicas). Además, algunos estudios destacan que $TPATRE_t$ puede verse altamente influenciado a su vez por $SPUBLI_t$ (Meyer, 2000).

De acuerdo a la tradición Schumpeteriana, el modelo (ver figura 6.2) incorpora la diferencia entre innovaciones tecnológicas medidas a través de las patentes ($TPATRE_t$) y las

innovaciones no tecnológicas. El avance tecnológico no sería posible sin el desarrollo de nuevas formas de organización que guíen y apoyen la I+D+i y que permitan a las empresas obtener beneficios de esta clase de inversiones (Nelson, 1996). Estas innovaciones no tecnológicas son medidas por indicadores indirectos como las marcas comerciales comunitarias ($TRADEM_t$), número de pequeñas y medianas empresas o PYMEs que introducen innovaciones de marketing u organizacionales ($INNSME_t$) y porcentaje de las ventas de innovaciones nuevas para el mercado y nuevas para la empresa sobre el total de sus ingresos ($INNEW_t$) (IUS, 2010). Las variables $TRADEM_t$, $INNEW_t$ y $INNSME_t$ constituyen, como veremos en el apartado correspondiente, los outputs del denominado escenario de Innovación.

Figura 6.2. Representación del escenario de Innovación en la red Bayesiana



Fuente: Elaboración propia.

Seguindo la figura, también queda recogido el hecho de que $TPATRE_t$ y los indicadores de las innovaciones no tecnológicas ($TRADEM_t$, $INNEW_t$ y $INNSME_t$) dependen de una variedad de variables no medibles fácilmente de forma directa como son $INTECAP_t$ (capital intelectual o la diferencia entre los valores de mercado de los activos tangibles físicos y financieros y el capital humano, estructural y relacional) (Bontis et al., 1999), los $SPILLOV_t$ (spillovers, efecto derrame, o el hecho de que algunos de los beneficios económicos de las actividades de I+D+i llegan también a agentes económicos distintos a los que llevaron a cabo la investigación) (Block, 2012) y las redes sociales y de negocios, $SBNETW_t$, que se refiere a las conexiones

entre individuos y organizaciones (Pavitt, 1998; Parker, 2008). Estas variables tratan de enfatizar el hecho de que la innovación también puede darse sin un proceso sistemático de investigación y desarrollo previo (Hirsch-Kreinsen, 2008), por ejemplo, a través de la adquisición de conocimiento tácito y práctico y a través de la difusión formal e informal del conocimiento entre empresas.

Otras variables en la red que influyen sobre las innovaciones no tecnológicas (*TRADEM_t*, *INNEW_t*, y *INNSME_t*) serían: el aprendizaje llamado *learning-by-doing* (*LEARDO_t*) (Jensen y otros, 2007), el *benchmarking* (*BENCHM_t*), o “el proceso continuo de identificar, comprender y adaptar la práctica y los procesos que conducirán a una mejor ejecución” (Auluck, 2002; p. 111) y aspectos socioculturales, como la existencia o de una cultura empresarial en el país representado mediante la variable *ENTREP_t* (von Hippel, 1988).

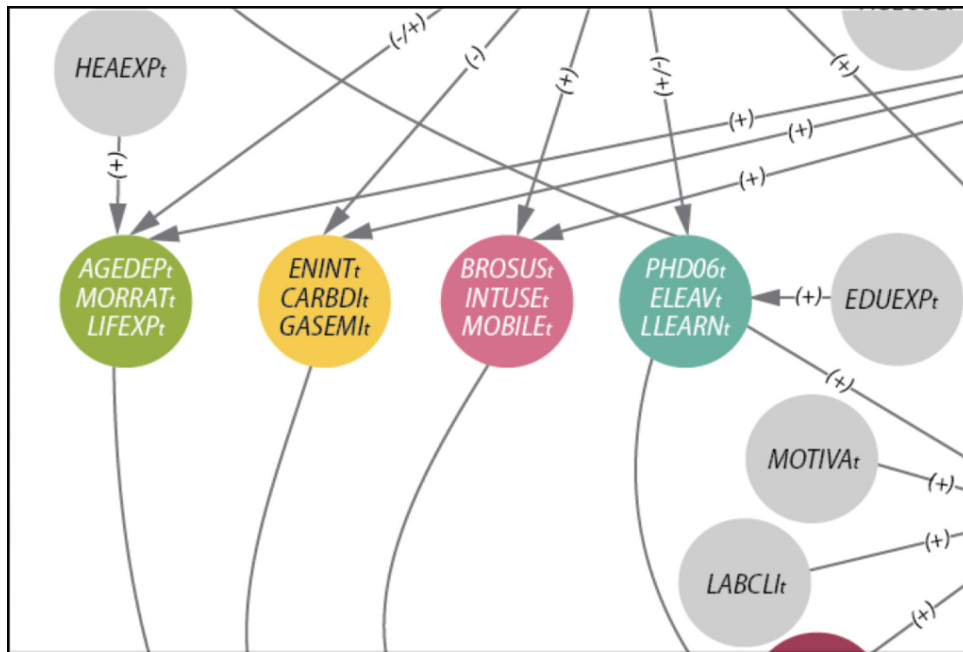
La localización de recursos en I+D+i que desemboca en la obtención de patentes y publicaciones científicas (escenario Tradicional) sería deseable que conllevara efectos positivos sobre el bienestar de la sociedad en general (*WELFAR_t*). Como se comentó al principio de este apartado, estos efectos socioeconómicos positivos tratan de ser reflejados en el modelo conceptual y pueden agruparse en cinco escenarios principales. Estos escenarios socioeconómicos, junto al de innovación, son los empleados a la hora de seleccionar los outputs que evaluarán la eficiencia técnica relativa de las políticas públicas de I+D+i de los 27 países de la UE (como se señaló, los inputs en todos estos escenarios serán los outputs del escenario tradicional, es decir, el número de publicaciones científicas y el número de patentes concedidas). Siguiendo los indicadores del sitio web oficial de Eurostat para la monitorización de los avances sobre desarrollo sostenible, los indicadores seleccionados fueron (figura 6.3):

Salud. Las mejoras en el sector de la salud debidas a las inversiones en I+D+i son ampliamente conocidas y su gasto es usualmente bien recibido por la sociedad. Tres indicadores, de entre los disponibles, han sido seleccionados para el análisis de la eficiencia en este escenario: *AGEDEP_t* (índice o tasa de dependencia de la población), *MORRAT_t* (tasa de mortalidad infantil) y *LIFEXP_t* (esperanza de vida).

Medio Ambiente. Obviamente, para el bienestar de una sociedad resulta fundamental un medio ambiente adecuado. Las variables que serán consideradas outputs a la hora de evaluar la eficiencia de este escenario fueron *ENINT_t* (intensidad energética de la economía de un país), *CARBDI_t* (daños del de dióxido de carbono) y *GASEM_t* (emisiones totales de gases invernadero).

Sociedad de la Información. El sector más relevante en el que los outputs de la I+D+i han tenido una influencia exponencial es el sector de las Tecnologías de la Información y la Comunicación (TICs). Con el incremento del peso relativo del sector servicios inducido por la economía del conocimiento en los países desarrollados, las TICs presentan una importancia incalculable tanto en la forma de vida de la sociedad como en los nuevos métodos de trabajo, conllevando a mejoras en el nivel de productividad por hora trabajada ($RLPGH_t$). Los indicadores seleccionados para este escenario fueron $BROSUS_t$ (suscriptores de banda ancha de internet), $INTUSE_t$ (usuarios de internet) y $MOBILE_t$ (número de suscriptores de teléfono móvil).

Figura 6.3. Escenarios de Salud, Medio Ambiente, Sociedad de la Información y Educación

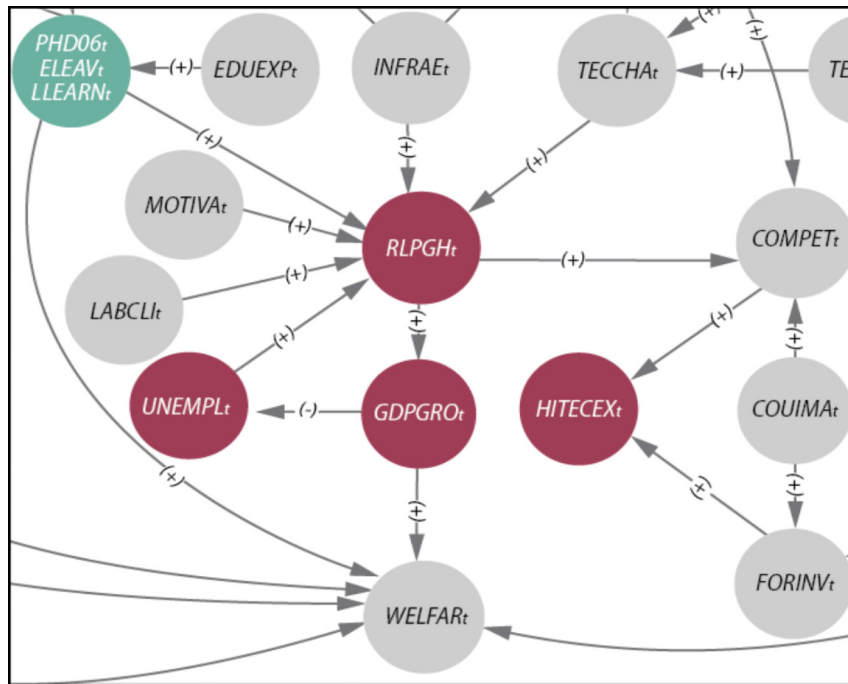


Fuente: Elaboración propia.

Educación. La población europea tiene que adquirir un amplio rango de habilidades y adaptarse a las nuevas situaciones en una sociedad cada vez más basada en el conocimiento y, por ende, en las TICs. La relevancia del conocimiento y la difusión del mismo juegan un papel muy importante para el acceso a un empleo ($UNEMPL_t$). Las variables seleccionadas en esta ocasión fueron $PHD06_t$ (número de doctores), $ELEAV_t$ (estudiantes que abandonan de forma temprana el sistema educativo) y $LLEARN_t$ (el aprendizaje a lo largo de la vida).

Económico. Este escenario centra su atención en la prosperidad económica del país, objetivo clave y último reflejado en las estrategias europeas para la mejora del bienestar de la sociedad. Los indicadores/outputs de este escenario en la red (figura 6.4) fueron el crecimiento del producto interior bruto ($GDPGRO_t$), el crecimiento de la productividad real por hora trabajada ($RLPGH_t$), la tasa de desempleo ($UNEMPL_t$) y las exportaciones de alta tecnología ($HITECX_t$).

Figura 6.4. Escenario Económico representado en la red Bayesiana



Fuente: Elaboración propia.

Siguiendo con la explicación de la red, desde el momento en que el crecimiento en la productividad juega un rol principal en el incremento de los ingresos de los ciudadanos, medido a través del crecimiento del PIB, es importante señalar los factores principales que subyacen en el crecimiento de dicha productividad: las infraestructuras del país ($INFRAE_t$), la motivación de los trabajadores ($MOTIVA_t$), el clima laboral ($LABCLI_t$) y el cambio tecnológico ($TECCHA_t$) muestran un papel muy importante en nuestro modelo de RB para explicar el crecimiento de la productividad real por hora trabajada ($RLPGH_t$). Otras variables que consideramos que influyen en el crecimiento de la productividad son aquellas relacionadas con el escenario de la educación ($PHD06_t$, $ELEAV_t$ y $LLEARN_t$).

Más allá, el nivel del cambio tecnológico ($TECCHA_t$) depende de la innovación tecnológica medida a través del número de patentes ($TPATRE_t$), el gasto total en I+D+i como porcentaje del PIB ($GERDIO_t$) y la balanza de pagos tecnológica ($TECACQ_t$) que incluye las diversas formas de adquirir tecnología como la compra de bienes de capital o servicios tecnológicos, incluyendo la licencia de la propiedad intelectual.

Como establece el Informe Global de Competitividad del Foro Económico Mundial (WEF 2010, p.4), la productividad determina la tasa de retorno de las inversiones (físicas, humanas y tecnológicas) realizadas en una economía. La $RLPGH_t$ influye en la competitividad general de un país y establece el nivel sostenible de prosperidad que puede ser alcanzado por una economía.

Los impactos económicos mostrarán la medida en la que los outputs tradicionales (utilizados como inputs de los escenarios analizados, que son $SPUBLI_t$ y $TPATRE_t$) influirán indirectamente y conseguirán el crecimiento de la productividad real por hora ($RLPGH_t$), el correspondiente crecimiento del PIB ($GDPGRO_t$) así como sus implicaciones en las exportaciones de alta tecnología ($HITECH_t$) y la reducción de la tasa de desempleo ($UNEMPL_t$)

El marco institucional de un país ($INSTFR_t$) también juega un papel relevante en el bienestar de una sociedad. El Foro Económico Mundial la define como “el marco administrativo y legal dentro del cual los individuos, las empresas y los gobiernos interactúan para generar ingresos y bienestar en la economía” (WEF 2010, p. 4). La importancia del marco institucional incluso se ha visto incrementada durante la crisis económica dado el papel directo jugado por los Estados en sus economías.

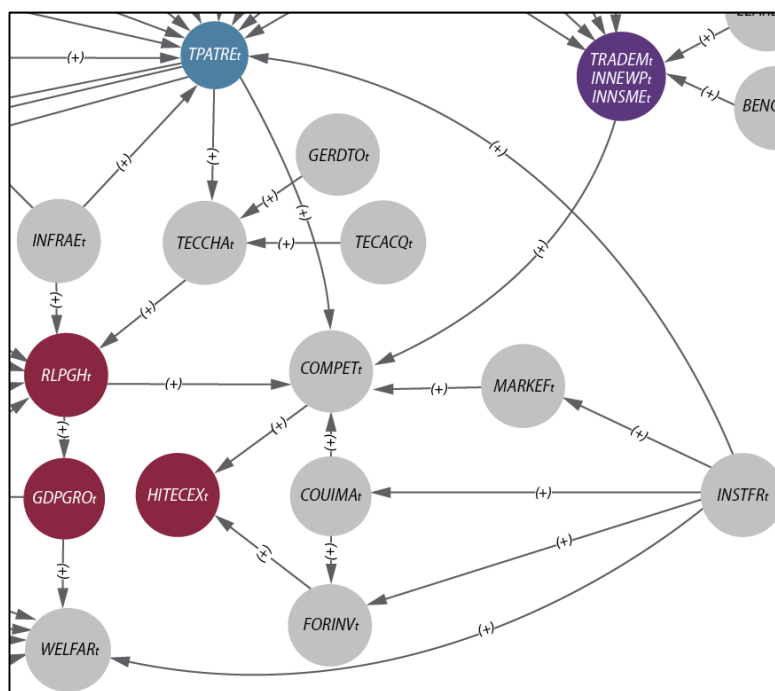
Como ejemplo de esta variable, podemos citar las leyes de patentes (como el Acta Bayh-Dole en Estados Unidos), la introducción de un mercado único para bienes y servicios en Europa, la liberalización del comercio (lo que influye también en la competitividad), la política macroeconómica (lo que afecta a los tipos de interés y el acceso de las empresas al crédito), la política de educación (que influye en el coste de la I+D+i y la formación). Todas estas políticas tienen un efecto potencial sobre las iniciativas de innovación y, por tanto, en el crecimiento económico a la largo plazo.

Este marco legal influye sobre la variable $TPATRE_t$ (ver figura 6.5) dado que los derechos de propiedad intelectual deben ser protegidos por la ley para que sean efectivos. Y esto sugiere un

segundo papel para el Estado como garante: los gobiernos necesitan encontrar el equilibrio adecuado entre preservar los beneficios de la innovación y al mismo tiempo no impedir o disuadir la entrada de futuras innovaciones.

En la RB, $INSTFR_t$ condiciona la imagen en el exterior del país ($COUIM_t$), la inversión directa extranjera ($FORINV_t$) y la eficiencia de los mercados ($MARKEF_t$). $COUIM_t$ puede influir también sobre la última variable. Ejemplo de ello es el efecto causado en los mercados financieros de un país por la calificación de la deuda soberana de éste por parte de las agencias de calificación como Moody's, Standard and Poor's o Fitch.

Figura 6.5. Variables que influyen en la competitividad de un país



Fuente: Elaboración propia.

En la figura 6.5, además, se recogen los efectos sobre la competitividad que pueden ser alcanzados gracias al crecimiento de la productividad ($RLPGH_t$), indicadores de innovaciones tecnológicas ($TPATRE_t$) y no tecnológicas ($TRADEMt$, $INNEWPt$ y $INNSMEt$) y la eficiencia de los mercados ($MARKEF_t$).

Para terminar, todas estas variables y su relaciones de causalidad convergen en la medida más popular, ya citada, para determinar los niveles de bienestar de una sociedad, o lo que es lo

mismo, el crecimiento del PIB (*GDPGRO t*). El crecimiento económico ha sido objeto común de atención por parte de economistas y de decisores políticos (Galindo, 2011). Este crecimiento representa la prosperidad de una sociedad y no resulta extraño que se considere como la medida comúnmente utilizada del bienestar de un país y un objetivo a alcanzar. En ocasiones, se considera una señal de la adecuación de la actividad de un gobierno (un crecimiento económico mayor nos indica una política económica adecuada)¹².

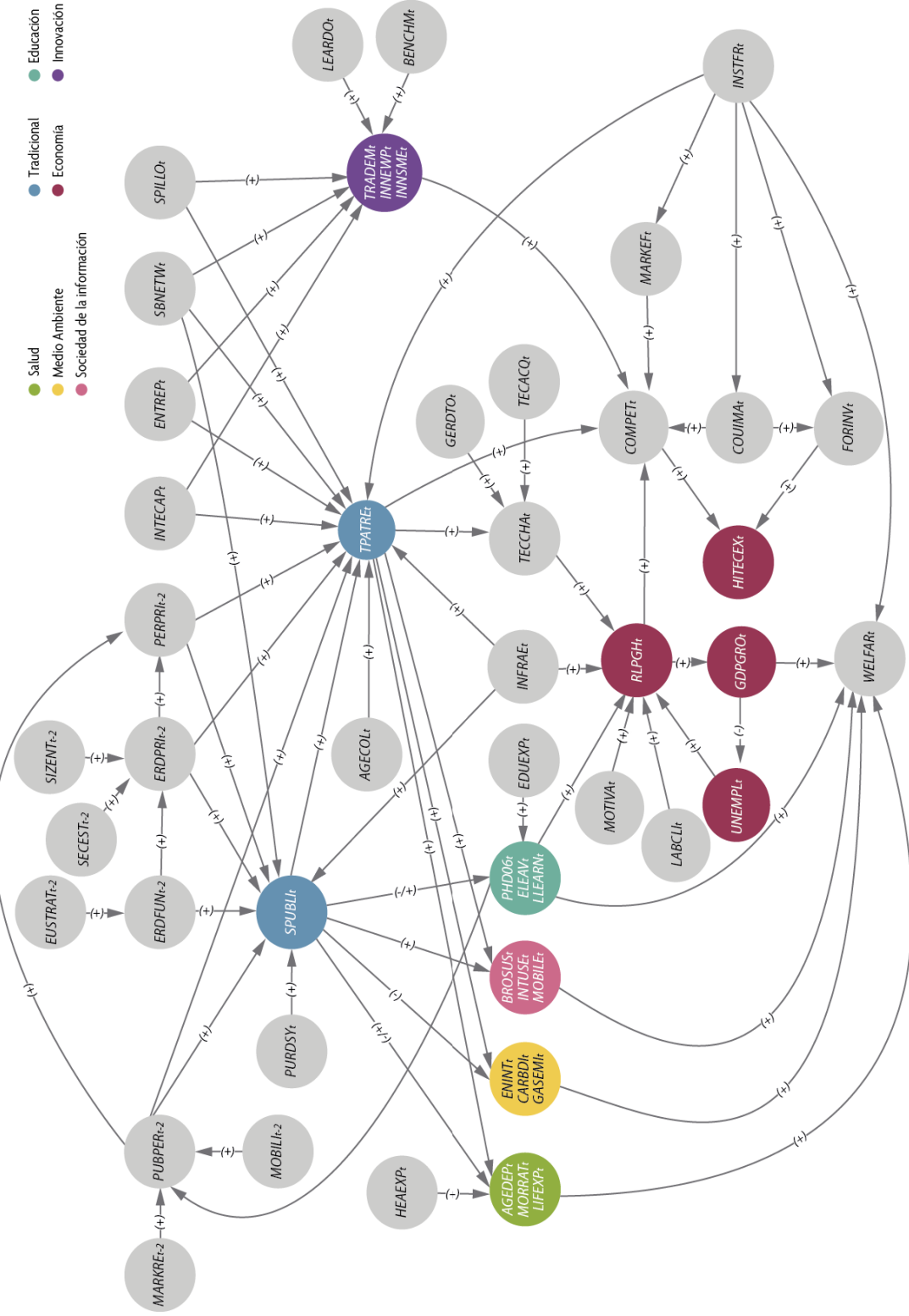
Por lo que se refiere a su medición, por regla general, se suele utilizar la tasa de crecimiento del PIB en términos reales para eliminar los efectos de la inflación. Aunque este indicador no está exento de críticas¹³ por no considerar aspectos como la desigualdad económica, las externalidades, los costes del crecimiento, los cambios en la población o la economía informal, introduciéndose indicadores alternativos destinados a medir el desarrollo humano (Gadrey y Jany-Catrice, 2006), se ha seleccionado por seguir siendo el indicador más empleado y comparable a nivel internacional.

A lo largo de la explicación detallada de la RB, se puede concluir que se trata de una aproximación híbrida para evaluar la eficiencia técnica relativa de las políticas de I+D+i, dado que se emplean variables tanto de a nivel micro como a nivel macroeconómico.

¹² Sin embargo, esta visión ha sido objeto de críticas que se fundamentan en aspectos relacionados con los problemas de distribución de la renta, costes medioambientales, efectos sociales, etc. (McMahon y Squire, 2003).

¹³ Se pueden clasificar en cuatro grupos: los referentes al desarrollo humano y progreso social, como el Índice de Desarrollo Humano (IDH) o el Index of Social Health (ISH); los que formulan un PIB ampliado, como el Measure of Economic Welfare de Nordhaus y Tobin; los que incluyen aspectos medioambientales como el Index of Sustainable Economic Welfare (ISEW), el Genuine Progress Indicator (GPI) o Sustainable National Income (SNI); y, finalmente, índices de bienestar económico, de los que también existen múltiples ejemplos.

Figura 6.6: Modelo conceptual causal (red Bayesiana) del crecimiento económico basado en el conocimiento y la innovación



Fuente: Elaboración propia.

Tabla 6.1. Variables de la Red Bayesiana (RB) y selecciones para el análisis 2LS-DEA

ETIQUETA DEL NODO	VARIABLE (a)	DESCRIPCIÓN	I/O	ESCENARIO	% (b)	INTERVALOS (c)
<i>ERDFUN</i>	(1) Gasto público en I+D+i	Gasto en I+D+i por fuente de fondos: sectores Gobierno y Educación Superior como % del PIB.	I	Tradicional	100	Bajo: [0-0,55]; Alto: (0,55-1)
<i>PUBPER</i>	(2) Personal público dedicado a la I+D+i	Personal dedicado a la I+D+i por sectores de ejecución: Gobierno y Educación Superior considerados como porcentaje del total de empleados (en el denominador, equivalentes a tiempo completo).	I	Tradicional	100	Bajo: [0-0,7]; Alto: (0,7-1,5)
<i>TPATRE</i>	(3) Total de solicitudes de patentes (***)	Solicitudes de patentes mundiales pedidas a través del procedimiento del Tratado de Cooperación o en una oficina de patentes nacional (residentes) por cada 1000 habitantes.	I/O	Tradicional: output Resto escenarios: input	10%	Bajo: [0-1]; Alto: (1-4]
<i>SPUBLJ</i>	(4) Total de publicaciones científicas (c)	Número de documentos científicos del país según el Scientific Journal Ranking (SJR) de SCImago por cada 1000 habitantes.	I/O	Tradicional: output Resto escenarios: input	90%	Bajo: [0-1,5]; Alto: (1,5-3]
<i>GDPGRO</i>	(5) Tasa de crecimiento del PIB per capita	Tasa de crecimiento del PIB per capita: porcentaje de cambio respecto al periodo anterior, el PIB incluye bienes y servicios que tienen mercados (o que podrían tenerlos) y productos que se producen por el gobierno o instituciones sin ánimo de lucro. El PIB real per capita se calcula como el ratio del PIB real entre la media de población de un año concreto.	O	Económico	10%	Bajo: <=1; Medio: (1-3]; Alto:>3
<i>UNEMPL</i>	(6) Tasa de desempleo	Porcentaje de la fuerza de trabajo basado en la definición de la Oficina Internacional del Trabajo (ILO). La fuerza de trabajo es el número total de personas empleadas y desempleadas.	O	Económico	10%	Bajo: <=5; Medio: (5-8]; Alto:>8
<i>HITECEX</i>	(7) Exportaciones de alta tecnología (**)	Porcentaje de todas las exportaciones de productos de Alta Tecnología sobre el total de las exportaciones. El total de exportaciones de la UE, no incluye el comercio intracomunitario.	O	Económico	10%	Bajo: [0-10]; Medio: (10-15]; Alto: (15-100]
<i>RLPGH</i>	(8) Crecimiento de la productividad real del trabajo por hora trabajada	La productividad real del trabajo por hora se calcula como el output real (PIB deflactado medido en volúmenes encadenados, año de referencia 2000) por unidad de input de trabajo (medido por el número total de horas trabajadas). La medida por hora de trabajo ofrece una mejor imagen del desarrollo de la productividad en la economía que la productividad por persona empleada ya que elimina las diferencias entre a tiempo completo/a tiempo parcial de la fuerza de trabajo a través de los países y los años. El % de cambio se calcula sobre el año anterior, índice 2000=100.	O	Económico	10%	Bajo: <=1; Medio: (1-3]; Alto: >3

Tabla 6.1. Variables de la Red Bayesiana (RB) y seleccionas para el análisis 2LS-DEA (cont.)

ETIQUETA DEL NODO	VARIABLE (a)	DESCRIPCIÓN	I/O	ESCENARIO	%(b)	INTERVALOS(c)
<i>ELLEV</i>	(9) Abandonos tempranos de la educación y la formación	El abandono temprano en la educación y la formación se refiere a las personas de edad entre los 18 y los 74 que cumplen las siguientes dos condiciones: primero, el nivel más alto de educación alcanzada es ISCED(9) 0, 1, 2 o 3c, segundo, los encuestados declaran no haber recibido ningún tipo de formación en las cuatro semanas que preceden a la encuesta (numerador). Tanto el numerador como el denominador provienen de la Encuesta Comunitaria de Fuerza de Trabajo.	O	Educación	5%	Bajo: [0-8]; Alto:>8
<i>PHD06</i>	(10) Graduados Doctores	Número de graduados doctores por 1000 habitantes.	O	Educación	90%	Bajo: [0-0,15]; Alto: >0.15
<i>LLFARN</i>	(11) Aprendizaje a lo largo de la vida	Aprendizaje a lo largo de toda la vida (% de personas entre los 18 y los 64 años)	O	Educación	50%	Bajo: [0-15]; Alto: > 15
<i>INNSME</i>	(12) Innovaciones de marketing y organizacionales	PYMES que introducen innovaciones de marketing u organizacionales como porcentaje del total de PYMEs	O	Innovación	40%	Bajo: [0-30]; Medio: (30-50); Alto: >50
<i>INNEIWP</i>	(13) Ventas procedentes de innovaciones	Ventas de innovaciones nuevas para el mercado y nuevas para la empresa como % de los ingresos.	O	Innovación	40%	Bajo: [0-10]; Medio: (10-15); Alto:>15
<i>TRADEM</i>	(14) Marcas comunitarias (**)	Marcas comunitarias por billón de euros del PIB.	O	Innovación	20%	Bajo: [0-5]; Medio: (5-10); Alto:>10
<i>INTUSE</i>	(15) Usuarios de internet (**)	Personas con acceso a internet (por 1000 habitantes)	O	Sociedad de la Información	80%	Bajo: [0-70]; Alto: >70
<i>BROSUS</i>	(16) Suscriptores de internet de banda ancha.	La tasa de penetración de banda ancha describe el número de conexiones de alta velocidad por 1000 habitantes. Este indicador muestra la amplitud de la penetración del acceso a banda ancha en el país a nivel general, sin especificar por grupo de usuarios. Las líneas de banda ancha se definen como aquellas con una capacidad igual o a 144 Kbits/s (cubre tecnologías como el ADSL, cable modem así como otro tipo de líneas de acceso a internet).	O	Sociedad de la Información	80%	Bajo: [0-275]; Alto: >275
<i>MOBILE</i>	(17) Suscriptores de teléfonos móviles.	Número de suscripciones de teléfono móvil por 1000 habitantes.	O	Sociedad de la Información	80%	Bajo: [0-1000]; Alto: >1000
<i>LIFEXP</i>	(18) Esperanza de Vida	La media de años que un recién nacido puede esperar vivir si estuviera sujeto a las actuales condiciones de mortalidad en su vida.	O	Salud	70%	Alto: [60-80]; Muy Alto: >80
<i>AGEDEP</i>	(19) Ratio de dependencia (% de la población con edad de trabajar)	Ratio entre el número total de personas mayores de una edad cuando están generalmente económicamente inactivas (más de 65 años) y el número de personas en edad de trabajar (de los 15 a los 64 años).	O	Salud	70%	Bajo: [0-50]; Alto:>50
<i>MORRAT</i>	(20) Tasa de mortalidad, niños (por cada 1000 nacimientos)	El ratio entre el número de muertes infantiles por debajo de un año de edad durante todo un año en relación con el número de nacimientos en ese año. El valor está expresado por 1000 nacimientos.	O	Salud	70%	Bajo: [0-4]; Alto:>4

Tabla 6.1. Variables de la Red Bayesiana (RB) y seleccionas para el análisis 2LS-DEA (cont.).

ETIQUETA DEL NODO	VARIABLE (a)	DESCRIPCIÓN	I/O	ESCENARIO	% (b)	INTERVALOS (c)
<i>ENINT</i>	(21) Intensidad energética de la economía	Ratio entre el consume interno bruto de energía y el Producto Interior Bruto (PIB) en un año determinado. Mide el consumo de energía de una economía y su eficiencia energética de forma global. El consume bruto interno de energía se calcula como la suma del consume bruto interno de cinco energías: carbón, electricidad, petróleo, gas natural y fuentes de energía renovables. Los datos del PIB se tomaron como volúmenes encadenados con el año 2000 como referencia. Como el consumo bruto interno se mide en kgoe (kilogramos equivalentes de petróleo) y el PIB en 1000 eur, la unidad de media es kgoe/1000 eur.	O	Medio Ambiente	50%	Bajo: [0-150]; Medio: (150-300]; Alto: >300
<i>GAEMI</i>	(22) Emisiones totales de gases de efecto invernadero	Emisión de gases de efecto invernadero por sector. Total de emisiones por 1000 habitantes. La emisión anual de gases invernadero se estiman y reportan bajo el marco de la Convención de las Naciones Unidas sobre cambio climático, el Protocolo de Kyoto y la Decisión 280/2004/EC.	O	Medio Ambiente	50%	Bajo: [0-10]; Medio: (10-15]; Alto: >20
<i>CARBDI</i>	(23) Daños del dióxido de carbono (d)	Ahorro ajustado: daños del dióxido de carbono (% del PIB).	O	Medio Ambiente	50%	Bajo: [0-0,2]; Medio: (0,2-0,4]; Alto: >0,4

(a) Fuente: Eurostat, excepto: (a) Publicaciones científicas (web SCImago Journal & Country Rank), (b) Base de Datos del Banco Mundial, (c) WIPO y (d) el Informe de Global de Competitividad del Foro Económico y Mundial (WEF).

(b) Porcentaje del valor de cada variable en el 2LS-DEA para el correspondiente escenario.

(c) Interpretación del valor de las variables: intervalos de los inputs/outputs.

(d) La clasificación de los niveles de educación se basan en la Clasificación Estándar Internacional de Educación (ISCED, 1997) en todas las variables relacionadas con la Educación a lo largo de este trabajo.

Tabla 6.2. Descripción del resto de variables en la Red Bayesiana (RB) no incluidas en ningún escenario analizado en el 2LS-DEA

ETIQUETA DEL NODO	VARIABLE (6)	DEFINICIÓN	INTERVALOS(6)
<i>AGECOL</i>	(24) Colaboración entre los agentes	Colaboración entre los agentes del sistema de innovación.	Bajo=0;Alto=1
<i>BENCHM</i>	(25) Benchmarking (análisis comparativo)	El proceso continuo de medición de los productos, servicios y prácticas contra los competidores más fuertes o aquellas compañías reconocidas como industrias líderes o la búsqueda de mejores prácticas de la industria que conducen a una mejor ejecución.	Bajo=0;Alto=1
<i>COMPET</i>	(26) Competitividad (****)	El conjunto de instituciones, políticas y factores que determinan el nivel de productividad de un país (ranking del Informe de Competitividad Global del WEF).	Alto=1-25; Bajo=27-133
<i>COUIMA</i>	(27) Imagen exterior del país	Confianza y reputación del país en el exterior.	Bueno=1; Malo=0
<i>EDUEXP</i>	(28) Gasto en educación como % del PIB	Gasto público total en educación como % del PIB, para todos los niveles educativos.	Bajo: [0-4,75]; Alto>=4,75
<i>ENTREP</i>	(29) Cultura emprendedora del país	La cultura se usa para referenciar el conjunto de valores duraderos de un país. El emprendimiento se define como el acto y el proceso por el que una sociedad o los individuos identifican y llevan a cabo oportunidades de negocio para crear riqueza.	Si=1; No=0
<i>ERDPRIV</i>	(30) Gasto privado en I+D+i	Gasto en I+D+i, por fuente de fondos: sector empresas como % del PIB.	Bajo: [0-1]; Alto: [1-3]
<i>EUSTRAT</i>	(31) Estrategia Europea para la política de I+D+i	Principales objetivos de las estrategias de la UE para alcanzar una economía del conocimiento, un crecimiento económico inteligente, más competitiva, etc., como la Estrategia de Lisboa (al menos el total de gasto en I+D tiene que alcanzar el 3%del PIB y el sector privado debe representar al menos el 60% del total del gasto público en I+D+i).	Muy bueno: ERDFUN/GERDT<=0,4 AND GERDT>=3% Muy malo: ERDFUN/GERDT>0,55 AND GERDT<1,5%
<i>FORINV</i>	(32) Inversión extranjera	Presencia de empresas multinacionales en el país	Bajo=0; Alto=1
<i>GERDIO</i>	(33) Gasto total en I+D+i	Gasto en I+D+i (total) como % del PIB	Bajo:[0-1,5];Medio(1,5-3); Alto:>=3
<i>HEAEXP</i>	(34) Gasto en salud	Gasto total en salud como % del PIB	Bajo: [(0-8]; Alto:>8
<i>INFRAE</i>	(35) Infraestructuras	Formación bruta de capital fijo que consiste en adquisiciones de productores residentes, menos las disposiciones, activos fijos tangibles e intangibles. Esto cubre en particular la maquinaria y el equipamiento, vehículos, utilaje y otros edificios, incluyendo laboratorios y otras infraestructuras científicas e inversiones.	Satisfactoria=1 No satisfactoria=0
<i>INSTPR</i>	(36) Marco institucional y político.	El marco legal y administrativo en el que los individuos, las empresas y los gobiernos interactúan para generar ingresos y salud en una economía determina el ambiente institucional.	Estable=1; Inestable=0
<i>INTECAP</i>	(37) Capital intelectual	Diferencia entre el valor de los activos tangibles (físicos y financieros) y valor del mercado (incluyendo capital humano, estructural y relacional).	Bajo=0;Alto=1
<i>LABCLI</i>	(38) Clima laboral	Clima de gestión del trabajo en las organizaciones.	Malo=0;Bueno=1
<i>LEARDO</i>	(39) Aprender haciendo (learning by doing)	Cuanto más temprana es la innovación, antes los empleados de una empresa pueden aprender haciendo, lo que conduce a innovaciones y costes más bajos.	Bajo=0; Alto=1
<i>MARKEF</i>	(40) Eficiencia de los mercados	Ranking del Informe Global de Competitividad.	Bajo=0;Alto=1
<i>MARKRE</i>	(41) Mercado para los investigadores(****)	Atractivo del mercado para investigadores (salarios)	Atractivo=1; No atractivo=0
<i>MOBILI</i>	(42) Movilidad de los investigadores y de los estudiantes	Nivel de movilidad de los investigadores y estudiantes de otros países.	Alto, Bajo
<i>MOTIVA</i>	(43) Motivación	Habilidades, motivación y compromiso como vital para la productividad de las personas.	Motivado=1; No motivado=2

Tabla 6.2. Descripción del resto de variables en la Red Bayesiana (RB) no incluidas en ningún escenario analizado en el 2LS-DEA (cont.).

ETIQUETA DEL NODO	VARIABLE (a)	DEFINICIÓN	INTERVALOS (b)
<i>PERPRI</i>	(44) Personal privado de I+D	Personal de investigación y desarrollo, por sectores de ejecución: sector empresas, considerado como el porcentaje del total de empleo (en el denominador, equivalentes a tiempo completo)	Bajo: (0-1]; Alto: (1-3]
<i>PURDSY</i>	(45) Sistema público de I+D	El conjunto de instituciones implicado en las actividades de I+D.	Apropiado=1; Inapropiado=0
<i>SBNETW</i>	(46) Redes sociales y de negocios	Las redes sociales suponen la práctica de expandir el número los contactos del negocio o sociales mediante la realización de conexiones a través de los individuos. El potencial de Internet para promover tales conexiones está ahora siendo completamente reconocido y explotado.	Bajo=0; Alto=1
<i>SECEST</i>	(47) Estructura de sectores del país	Categorización de la economía de un país de acuerdo a la prevalencia o no del sector industrial.	Industrial=1; Otros sectores=0
<i>SIZEENT</i>	(48) Tamaño de las empresas	Tamaño de las empresas predominante en el país.	PYMEs=0; Grandes Empresas=1
<i>SPILLO</i>	(49) Spillovers	Presencia de spillovers fuertes o débiles.	Bajo=0; Alto=1
<i>TECACQ</i>	(50) Otras formas de adquisición de tecnología	Diferentes formas de adquisición de tecnología medida a través de la balanza tecnológica de pagos, etc.	Positivo=1; Negativo=0
<i>TECCHA</i>	(51) Cambio Tecnológico	El cambio tecnológico es un término que es usado en este estudio para describir un cambio en el conjunto de posibilidades de producción viables.	Fuerte=1; Débil=0
<i>WELFAR</i>	(52) Bienestar de las sociedades europeas	Se refiere a la mejora de la calidad de vida de las sociedades europeas.	Alto, Medio, Bajo

(a) Fuente: Eurostat, excepto: (a) Publicaciones científicas (SCImago), (b) Banco Mundial, (c) WIPO y (d) Informe Global de Competitividad por el Foro Económico Mundial.

(b) Interpretación de los intervalos de las variables definidos por los expertos consultados.

BLOQUE 3

ANÁLISIS

EMPÍRICO

*Si buscas resultados distintos,
no hagas siempre lo mismo.*

Albert **Einstein**

CAPÍTULO VII. EFICIENCIA DE LAS POLÍTICAS PÚBLICAS DE I+D+i DE LA UE EN EL MARCO DE UN DESARROLLO SOSTENIBLE

1. INTRODUCCIÓN

Una vez abordado el objetivo de diseñar un modelo conceptual causal, dicho modelo va a ser empleado para analizar la eficiencia del uso de los recursos públicos de I+D+i (primer objetivo) desde tres perspectivas diferentes e interrelacionadas: aspectos de la propia I+D+i, la economía del conocimiento (EC) y el desarrollo sostenible (DS), que fueron presentados en el capítulo II.

La principal razón para evaluar los impactos de las políticas públicas de I+D+i es conocer su contribución a los objetivos públicos de tipo general evitando dedicar los recursos escasos a programas ineficientes y/o inefectivos. El diseño de un modelo conceptual para identificar las variables clave y sus relaciones en los procesos de I+D+i que sirva a la consecución de un DS puede ser una útil herramienta de apoyo a la toma de decisiones para los responsables del diseño de políticas públicas (ver capítulos del IV al VI).

La evaluación de la consecución de un DS utiliza como principal herramienta indicadores, índices compuestos, una gran variedad de variables y adopta perspectivas diversas pero complementarias (Singha et al., 2012) (desde la innovación, el conocimiento, las tecnologías de la información, etc). La perspectiva que adoptamos, en este caso concreto, para evaluar la eficiencia de las políticas públicas de I+D+i en la UE es la de la consecución de un desarrollo sostenible (DS). Si estos recursos públicos dedicados a la I+D+i son empleados de forma eficiente respecto a varios escenarios socioeconómicos de DS en una economía basada en el conocimiento, ello redundará en el bienestar de la sociedad europea (ver figura 2.1, capítulo II, que refleja la interconexión entre los conceptos empleados).

Respecto a la metodología seleccionada, el Análisis Envolvente de Datos (DEA) consiste en una técnica no paramétrica que evalúa la eficiencia técnica relativa de un conjunto comparable de Unidades de Decisión (*Decision-Making Units* o DMUs), las cuales usan múltiples inputs para producir múltiples outputs (Cooper et al., 2004). Este análisis identifica las DMUs más

BLOQUE 3- ANÁLISIS EMPÍRICO

eficientes que son usadas como estándares (por ello se denomina eficiencia relativa) para la comparación con el resto de DMUs.

Rousseau y Rousseau (1997) fueron los primeros autores que evaluaron la eficiencia técnica relativa de los procesos de I+D en 18 países desarrollados. Posteriormente, la técnica DEA ha sido empleada también para estudiar la eficiencia de la Ciencia y la Tecnología (Meng et al. 2006), de las universidades (Bonaccorsi et al., 2006), de centros de investigación (Korhonen et al. 2001), de grupos de investigación (Guan and Wang, 2004) y de actividades de I+D (Wang and Huang, 2007).

En concreto, la eficiencia de las políticas públicas de los 27 Estados Miembros de la UE va a ser evaluada a través de un DEA en dos pasos enlazados (2LS-DEA). Este método propone tener en cuenta la fase de difusión de las actividades de I+D+i a través de las innovaciones introducidas con éxito en el mercado, de modo que los outputs tradicionales de la I+D+i se convierten en los inputs de varios escenarios socioeconómicos correspondientes a un DS. Todo ello en base al modelo diseñado (capítulo VI).

Los inputs y outputs correspondientes para cada escenario se eligieron de acuerdo a las relaciones representadas en el modelo conceptual causal. Los valores de dichos inputs y outputs fueron interpretados mediante la base de reglas (RU_i) construida que recogió el conocimiento experto explícito (K_i), tanto incorporado en la RB como el proveniente de la revisión de la literatura y de la opinión de los expertos, obteniéndose así la base de datos transformados (\mathbf{X}) empleados en el método guiado 2LS-DEA (ver los pasos seguidos por la metodología EbCA, capítulo V, figura 5.1).

Por tanto, los objetivos concretos del presente análisis empírico son:

- i) Mejorar el diseño del modelo conceptual que identifica las variables clave y sus relaciones causales para describir tanto los procesos de I+D+i como sus impactos socioeconómicos a través de la metodología EbCA. Toda la información relativa a la metodología para su diseño, sus asunciones, variables y relaciones y el resultado final se recogen en los capítulos V y VI del Bloque 2.
- ii) Evaluar la eficiencia técnica relativa de las políticas públicas de I+D+i en los 27 EEMM de la UE a través del enfoque 2LS-DEA, guiado por dicha metodología de adquisición de

conocimiento experto. Iterativamente, analizar las puntuaciones de eficiencia y obtener la versión final del modelo.

2. EVALUACIÓN DEL MODELO CONCEPTUAL CAUSAL A TRAVÉS DE UN MÉTODO DE ANÁLISIS CLÁSICO: LA EFICIENCIA DE LAS POLÍTICAS PÚBLICAS DE I+D+I DE LA UE

2.1. EL ANÁLISIS ENVOLVENTE DE DATOS (DEA)

El DEA, como técnica no paramétrica de programación matemática para la medición y la evaluación de la eficiencia técnica relativa de unas unidades de decisión o DMUs¹⁴, resulta muy apropiada en la comparación de la eficiencia de las políticas públicas de I+D+i en términos de sus múltiples outputs o resultados. Charnes, Cooper y Rhodes promocionaron ampliamente el uso de esta técnica cuyas raíces se encuentran en el trabajo seminal de Farrell de 1957.

Los conceptos centrales son los de frontera de producción y de eficiencia (técnica). El primero hace referencia al máximo output teórico alcanzable dada una combinación de inputs y la tecnología. El segundo, es definido por Farrell (1957) como la capacidad que tiene una entidad para obtener el máximo output a partir de un conjunto dado de inputs. Por tanto, evaluar la eficiencia (o ineficiencia) técnica de un conjunto de entidades pasa, en primer lugar, por estimar la frontera de producción, puesto que ésta no es conocida en la práctica.

Los métodos de estimación para construir la frontera de producción pueden clasificarse, en función de que se requiera o no especificar una forma funcional que relacione los inputs con los outputs, en métodos paramétricos y no-paramétricos. A su vez, pueden emplearse métodos estadísticos o no para estimar la frontera que, en última instancia, puede ser especificada como estocástica (aleatoria) o determinista.

En relación al modelo específico de DEA empleado, el modelo BCC (Banker, Charnes y Cooper, 1984) de retornos variables a escala fue el seleccionado dado que las actividades de I+D+i muestran retornos a escala crecientes o decrecientes (Bound et al., 1984). Se asume implícitamente, además, que el énfasis de las políticas públicas se encuentra más en lo que se

¹⁴ En el análisis, las DMUs son cada una de políticas de I+D+i de los países miembros de la UE. Por razones de brevedad, en ocasiones se hará referencia indistintamente tanto a estas políticas como a los países.

BLOQUE 3- ANÁLISIS EMPÍRICO

produce con unos recursos dados. De esta forma, el modelo adoptado ha sido el BCC (también conocido como VRS, Variable Returns to Scale) de retornos variables a escala orientado a los outputs. La libre imputación de los valores de los inputs y los outputs fue aplicada para la identificación de ineficiencias (Saen, 2010).

Considerando un conjunto de n DMUs ($DMU_j; j=1,2, \dots, n$), que cada DMU_j consume m inputs (x_{ij}) en diferentes cantidades ($x_{ij}; i=1,2, \dots, m$) y que produce s outputs (y_{rj}) ($y_{rj}; r=1,2, \dots, s$), la eficiencia técnica relativa de una DMU (orientada al output) puede ser evaluada mediante el siguiente modelo lineal (Cooper et al., 2004):

$$\max \phi + \varepsilon \left(\sum_{i=1}^m s_i^- + \sum_{r=1}^s s_r^+ \right)$$

Sujeto a:

$$\sum_{j=1}^m x_{ij} \lambda_j + s_i^- = x_{i0} \quad i=1,2, \dots, m$$

$$\sum_{j=1}^n y_{rj} \lambda_j + s_r^+ = \phi y_{r0} \quad r=1,2, \dots, s$$

$$\lambda_j \geq 0 \quad j=1,2, \dots, n$$

$$\sum_{j=1}^n \lambda_j = 1 \tag{7.1}$$

donde ϕ representa la puntuación de eficiencia, x_{ij} e y_{rj} son el i input y el r output, respectivamente, de la DMU j . De forma similar, x_{i0} e y_{r0} corresponden a los inputs y los outputs, respectivamente, de la DMU bajo evaluación (DMU_0). Además, λ_j denota escalares no negativos y s_i^- y s_r^+ son las holguras de los inputs y outputs, respectivamente.

Para aplicar DEA de forma efectiva se requiere (Cooper et al., 2007) que cumpla el criterio:

$$n \geq \max \{m \times s, 3 \times (m + s)\} \tag{7.2}$$

siendo n el número de DMUS, m el número de inputs y s el número de outputs. Esto es debido a que la técnica DEA se basa en la comparación de cada DMU con el resto de DMUs,

BLOQUE 3- ANÁLISIS EMPÍRICO

de forma que se necesita un conjunto considerablemente importante de ellas para que el análisis sea discriminante. Esta regla se cumple para cada uno de los escenarios seleccionados.

2.2. ESTRUCTURA DE LOS ESCENARIOS PARA EL DEA: INPUTS Y OUTPUTS SELECCIONADOS

Como ya se comentó anteriormente, para evaluar la eficiencia técnica relativa de las políticas públicas de I+D+i en los 27 EEMM de la UE y mejorar (iterativamente, a través de los resultados obtenidos) el diseño del modelo conceptual causal (metodología EbCA), el método de análisis fue el DEA, aplicado a la serie de escenarios representados en dicho modelo. Todos los inputs y todos los outputs seleccionados provienen del modelo conceptual (tabla 6.1) pero no todas las variables en el modelo son empleadas como inputs y outputs de los escenarios socioeconómicos seleccionados (tabla 6.2), sino que su influencia se recoge a través de la base de reglas (RU) que es aplicada a los datos originales (\mathbf{X}_0).

Los escenarios del modelo seleccionados para evaluar la eficiencia fueron¹⁵:

a. Escenario Tradicional. La eficiencia técnica relativa de las políticas públicas de I+D+i en la UE fue evaluada teniendo en cuenta dos inputs (Gasto público en I+D+i, $ERDFUN_{t,2}$, y personal del sector público dedicado a actividades de I+D+i, $PUBPER_{t,2}$) y dos outputs (publicaciones científicas, $SPUBLI_t$ y solicitudes de patentes, $TPATRE_t$) considerados comúnmente en la literatura (de ahí la denominación de “tradicional” de este escenario).

El número de publicaciones científicas es un indicador utilizado como medida de la productividad académica, pero no es el único. Informes de proyectos, monografías, tesis y otra literatura conocida como *grey literature* son igualmente importantes (Thomas et al., 2011). Sin embargo, dado que los valores de estos outputs no están disponibles, se usa el número de publicaciones científicas, $SPUBLI_t$. El número de patentes es el indicador, por otra parte, más ampliamente aceptado como medida de la innovación debido a la ausencia de indicadores más consistentes (Griliches, 1990). Este indicador tampoco está exento de desventajas: no todas las invenciones son técnicamente patentables, no todas las invenciones son patentadas y,

¹⁵ Las variables representadas en el modelo (Figura 6.6) están rodeadas por un círculo de forma individual (cada nodo un círculo), excepto aquellas que se corresponden con los escenarios socioeconómicos, que incluyen más de una variable (un círculo incluye más de una variable) al objeto de una mejor visualización.

BLOQUE 3- ANÁLISIS EMPÍRICO

finalmente, las empresas muestran diferentes propensiones a patentar en sus mercados internos y en los países extranjeros.

b. Escenarios socioeconómicos. Estos escenarios son combinaciones de inputs y outputs que describen diferentes facetas de un desarrollo económico sostenible. Todos ellos cuentan con los mismos inputs pero diferentes outputs. Los inputs son el número de publicaciones científicas $SPUBLI_t$ y el número de patentes $TPATRE_t$ (que son, precisamente, los outputs del escenario Tradicional).

De esta forma, podremos analizar la contribución de estos inputs en la ejecución de cada DMU (las políticas públicas de I+D+i de la UE) en relación a los escenarios de Salud, Medio Ambiente, de Sociedad de la Información, Educación, Económico y de Innovación. La idea de encadenar los outputs del escenario tradicional convirtiéndolos en los inputs de los escenarios socioeconómicos es que, si un país es eficiente en dicho escenario tradicional, se espera que sea eficiente también en los escenarios socioeconómicos que caracterizan un desarrollo sostenible (hipótesis H4).

De forma detallada, los escenarios contemplados se encuentran en el apartado 2 del capítulo VI donde se explica la estructura de la red Bayesiana, así como los inputs y outputs por escenarios, se encuentran recogidos en la tabla 6.1 de dicho capítulo VI.

2.3. EL ENFOQUE DEA EN DOS PASOS ENLAZADOS (2LS-DEA)

Basándonos en el modelo conceptual causal diseñado, la eficiencia de las políticas públicas de I+D+i en la UE no sólo fue evaluada usando el escenario Tradicional – lo que podemos denominar paso 1 –, sino que también fue analizada en seis escenarios socioeconómicos – lo que denominaremos paso 2 –, enlazado con el primero.

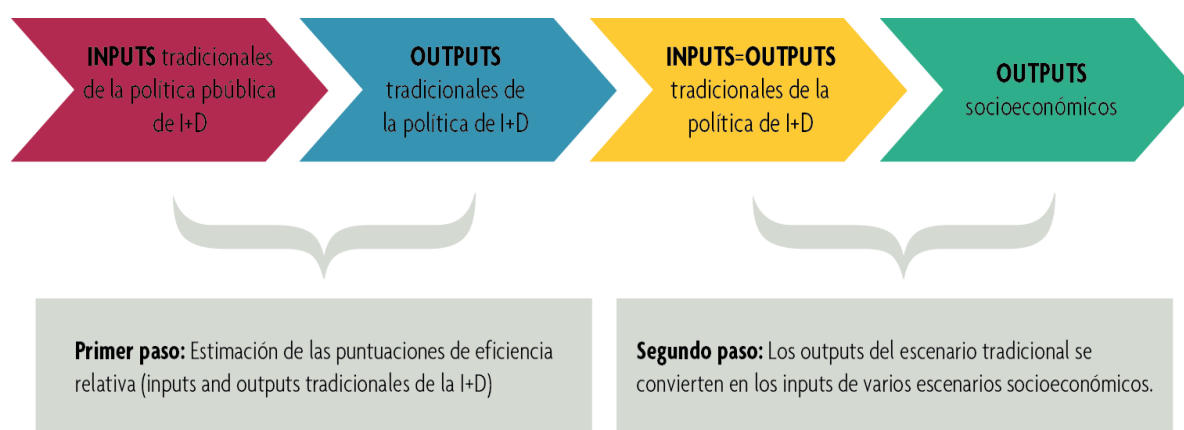
Al centrarnos, por otro lado, en la vertiente pública del gasto en I+D+i y los recursos públicos, estas variables influyen (de media) en un determinado porcentaje del valor que alcanzan las patentes solicitadas y las publicaciones científicas. Así, sólo el 10% del valor del total de las solicitudes de patentes ($TPATRE_t$) se considera que provienen de los recursos públicos $PUBPER_{t,2}$ y $ERDFUN_{t,2}$ y sólo el 90% del valor de las publicaciones científicas, $SPUBLI_t$, proviene de $PUBPER_{t,2}$ y $ERDFUN_{t,2}$. En la tabla 6.1 (columna del %) se especifica el porcentaje del valor original del output que será utilizado en cada escenario. El grupo de

BLOQUE 3- ANÁLISIS EMPÍRICO

expertos consultado para el diseño del modelo conceptual causal determinó este porcentaje de forma individual y, posteriormente, se empleó la media aritmética (este porcentaje constituye un parámetro que se puede modificar mediante un análisis de sensibilidad).

Los pasos propuestos por el enfoque 2LS-DEA, donde los outputs tradicionales se convierten en los inputs de dichos escenarios, de forma que ambos pasos se encuentran “vinculados” o “enlazados”, quedan resumidos en la figura 7.1.

Figura 7.1. Esquema del enfoque DEA en dos pasos enlazados (2LS-DEA)



Fuente: Elaboración propia.

2.4. CONJUNTO DE DATOS

El conjunto de datos de este análisis empírico está formada por los datos correspondientes a los 27 Estados Miembros de la UE. Los inputs y outputs cuantitativos (variables continuas y aleatorias) son previamente seleccionados, como ya hemos indicado, del conjunto de variables aleatorias del modelo conceptual causal para cada uno de los escenarios y posteriormente recopilados y procesados. Para el análisis, se han asumido los valores que ofrecen las bases de datos oficiales como estimadores modales de estas variables aleatorias.

La base de datos proviene de las siguientes fuentes oficiales: indicadores de Ciencia y Tecnología de la página oficial de Eurostat para todas las variables, excepto la Organización Mundial Internacional de Patentes (World Intellectual Property Organization o WIPO para las

BLOQUE 3- ANÁLISIS EMPÍRICO

solicitudes de patentes, SCImago Journal & Country Rank en el caso del número de publicaciones científicas por país y la base de datos del Banco Mundial para las variables índice de dependencia (*AGEDEP*), suscriptores de banda ancha de internet (*BROSUS*), exportaciones de alta tecnología (*HITECEX*) y número de suscripciones de telefonía móvil (*MOBILE*). Las variables se normalizaron a 1000 habitantes cuando fue necesario. La tabla 7.1 muestra los estadísticos descriptivos para estos inputs y outputs.

Tabla 7.1. Estadísticos descriptivos de los inputs y outputs empleados

ETIQUETA DEL NODO	Rango	Min.	Max.	Media	Desv. Típica	Varianza
<i>AGEDEP</i>	15,73	38,15	53,87	46,88	4,15	17,23
<i>BROSUS</i>	322,93	88,26	411,19	217,01	85,03	7230,25
<i>CARBDI</i>	0,79	0,08	0,87	0,27	0,19	0,04
<i>ELEAV</i>	34,00	5,00	39,00	14,37	8,46	71,62
<i>ENINT</i>	841,03	103,13	944,16	285,07	201,53	40612,65
<i>ERDFUN</i>	0,72	0,16	0,88	0,50	0,21	0,04
<i>GASEM</i>	20,31	5,25	25,57	10,70	3,96	15,66
<i>GDPGRO</i>	12,70	-5,20	7,50	,88	3,18	10,14
<i>HITECEX</i>	45,16	5,03	50,19	14,42	10,28	105,61
<i>INNEWP</i>	12,79	5,88	18,67	12,89	3,58	12,79
<i>INNSME</i>	36,79	16,82	53,61	32,39	10,48	109,88
<i>INTUSE</i>	58,90	28,79	87,70	60,82	16,49	271,77
<i>LIFEXP</i>	12,55	71,96	84,51	78,33	3,32	11,02
<i>LLEARN</i>	28,60	1,40	30,00	9,57	7,18	51,55
<i>MOBILE</i>	42,25	2,63	44,88	9,33	9,01	81,14
<i>MORRAT</i>	24,60	1,80	26,40	5,29	4,71	22,16
<i>PHD06</i>	0,30	0,02	0,31	0,17	0,09	0,01
<i>PUBPER</i>	0,83	0,18	1,01	0,47	0,16	0,03
<i>RLPGH</i>	138,40	39,30	177,70	93,50	30,30	917,84
<i>SPUBLI</i>	2,27	0,28	2,54	1,27	0,68	0,46
<i>TPATRE</i>	1,89	0,04	1,92	0,61	0,67	0,44
<i>TRADEM</i>	16,54	1,03	17,56	5,09	3,52	12,36
<i>UNEMPL</i>	8,50	2,80	11,30	6,21	1,87	3,52

Estas fuentes fueron empleadas, así mismo, para asignar los intervalos de valores de cada variable. La tabla 6.1 muestra la lista de variables cuantitativas (inputs y outputs) representadas en el modelo causal de RB por escenario, sus definiciones, sus unidades de medida, la fuente de la que provienen y los intervalos (última columna). Esta columna servirá a la construcción de la base de conocimiento, entendida como el conjunto de reglas que transforma los valores

BLOQUE 3- ANÁLISIS EMPÍRICO

originales de las variables penalizando y bonificándolos (inputs/outputs) y que permite la transferencia del conocimiento explícito disponible al modelo 2LS-DEA. La tabla 6.2 recoge los mismos aspectos del resto de variables no empleadas en los escenarios. Finalmente, como vimos en la descripción de los dos pasos del DEA, la columna “%” representa el porcentaje del valor de cada output que, de media, proviene de los correspondientes inputs.

El periodo de referencia de los datos fue el año 2008 debido a la disponibilidad de los mismos en la Encuesta Comunitaria de la Innovación realizada por Eurostat (Community Innovation Survey) que es llevada a cabo cada dos años. Por tanto, considerando el retardo presente en el proceso de innovación, tal y como se recogió en la red, los inputs representan datos del año 2006 ($t-2$) y los outputs del año 2008 (t). Así, queda representado el hecho de que si un país promueve inversiones en el año $t-2$, se espera obtener resultados en el año t .

2.5. LA TRANSFORMACIÓN DE LOS INPUTS Y LOS OUTPUTS: LA BASE DE REGLAS

Con el objeto de diseñar los modelos BCC-DEA, los valores originales de las variables cuantitativas en el conjunto de datos fue transformado mediante la aplicación de la base de conocimiento experto estructurado en forma de reglas. Así, el conjunto de datos fue transformado aplicando esta base de reglas (tabla 7.2) que incluye todas las variables, cualitativas y cuantitativas; las principales fuentes para las variables cualitativas empleadas fueron el Índice de Competitividad Global del Foro Económico Mundial (WEF, 2010), los perfiles de los países europeos recogidos en la página ERAWATCH¹⁶ perteneciente al Espacio Europeo de Investigación y, finalmente, de la opinión del panel de expertos consultado.

Los modelos clásicos DEA interpretan que, a mayor output, mayor será la eficiencia de una DMU (si los valores de los inputs se mantuvieran constantes) y que, cuanto más bajos los inputs, mayor la eficiencia (si los valores de los outputs se mantuvieran constantes). El DEA orientado a los outputs trata de maximizar los valores de los outputs una vez se han fijado los valores de los inputs. Sin embargo, en algunas situaciones, los inputs necesitan aumentarse y algunos outputs necesitan reducirse para mejorar la eficiencia de una DMU (se trata de los inputs y outputs no estándar). En estos casos, (por ejemplo, el desempleo y el índice de

¹⁶ Disponible en <http://erawatch.jrc.ec.europa.eu> (último acceso, 15 de marzo de 2013).

BLOQUE 3- ANÁLISIS EMPÍRICO

dependencia de las personas mayores son outputs cuyo valor sería tanto mejor cuanto menor fuera) realizamos previamente una transformación monótona lineal creciente o decreciente sobre rangos específicos a los valores originales basada en la citada base de reglas de conocimiento experto.

Esta base se compone de un conjunto de reglas del tipo *if-then* inferidas de las relaciones causales en el modelo, de las características de las variables, y de las consultas bibliográficas y a expertos adicionales (tabla 7.2), siempre que finalmente se interpreten como restricciones. Por ejemplo, una regla de transformación lineal simple transforma $GDPGRO_t$ de acuerdo a la influencia de $RLPGH_t$. Si $GDPGRO_t \leq 1$ (*bajo*) y $RLPGH_t > 3$ (*alto*), entonces el valor original de $GDPGRO_t$ es penalizado dividiendo su valor original por dos (R15). La razón de ser de esta regla descansa en el hecho de que elevados niveles de productividad deben conducir a incrementar las tasas de crecimiento del PIB; cuando $GDPGRO_t$ es *bajo* y $RLPGH_t$ es *alto*, el país analizado no está alcanzando su objetivo.

Por otro lado, si $GDPGRO_t > 3$ (*alto*) y $RLPGH_t \leq 1$ (*bajo*) entonces el valor original de $GDPGRO_t$ se bonifica multiplicándolo por dos (R14 en la tabla 7.2) para hacer “entender” a la metodología DEA que esta situación refleja unos muy buenos resultados por parte del país. Una regla especial surge en el caso de las economías del Este, las economías de transición de los Nuevos Estados Miembros, debido a las altas tasas de crecimiento del PIB que presentan. Como nuevas economías, presentan esta elevada tasa de crecimiento del PIB ($GDPGRO_t$) y del crecimiento de la productividad real del trabajo por hora trabajada ($RLPGH_t$) debido, entre otras razones, a la abundancia de recursos naturales, a los bajos niveles de población y a una gran demanda interna (pero una baja tasa de empleo después de un periodo de socialismo). Para conseguir que el método DEA adquiriera este conocimiento, el valor real de $GDPGRO_t$ es transformado mediante la regla R16:

$$\text{IF } GDPGRO_t > 5 \text{ THEN } GDPGRO_t' = GDPGRO_t / 2 \quad (7.3)$$

Resulta de interés resaltar la transformación de $ERDFUN_{t,2}$ llevada a cabo gracias a las reglas R1, R2 y R3. $ERDFUN_{t,2}$ depende de la Estrategia Europea ($EUSTRAT_{t,2}$ en el modelo conceptual). Debido a los objetivos de la Estrategia de Lisboa aplicables en el periodo bajo estudio (alcanzar al menos el 3% del PIB de gasto en I+D+i y que el 60% del total del gasto en I+D+i debe ser financiado por el sector privado), $ERDFUN_{t,2}$ se transformó de acuerdo a las reglas R1, R2 y R3:

BLOQUE 3- ANÁLISIS EMPÍRICO

$$\begin{aligned} &\text{IF } ERDFUN_{t,2}/GERD_{t,2} \leq 0,4 \text{ and } GERD_{t,2} >= 3 \\ &\text{THEN } ERDFUN_{t,2}' = ERDFUN_{t,2} * 2 \end{aligned} \quad (7.4)$$

$$\begin{aligned} &\text{IF } ERDFUN_{t,2}/GERD_{t,2} > 0,55 \text{ AND } GERD_{t,2} < 2 \\ &\text{THEN } ERDFUN_{t,2}' = ERDFUN_{t,2} / 2 \end{aligned} \quad (7.5)$$

$$\text{IF } ERDFUN_{t,2} \leq 0,2 \text{ THEN } ERDFUN_{t,2}' = ERDFUN_{t,2} / 2 \quad (7.6)$$

A pesar del valor concreto seleccionado para transformar los datos “penalizando” o “bonificando”, los parámetros empleados se pueden modificar para realizar análisis de sensibilidad sobre ellos.

3. RESULTADOS DEL ANÁLISIS DE LA EFICIENCIA DE LAS POLÍTICAS PÚBLICAS DE I+D+I

El enfoque 2LS-DEA arrojó como resultado un ranking de eficiencia de las 27 políticas públicas de I+D+i de la UE en cada escenario de desarrollo sostenible (Tabla 7.3). Una puntuación igual a 100 significa que la política pública de I+D+i es eficiente (eficiencia técnica relativa); valores menores que 100 deben ser interpretados como diferentes grados de ineficiencia. En ausencia de restricciones en las ponderaciones, el índice de eficiencia se interpreta como una medida radial que nos informa sobre la expansión necesaria del output sobre el nivel actual para que una unidad productiva sea eficiente. Así, una unidad será eficiente cuando $\phi=1$ y además todas las variables de holgura (tanto asociadas a los inputs como a los outputs) sean nulas. Por el contrario, si $\phi=1$ coincide con la unidad pero las variables de holgura no son todas nulas, significa que puede experimentarse un incremento individual en alguno de los outputs superior al radial representado por el índice de eficiencia, y/o puede admitirse una reducción en alguno de sus inputs.

Los países más eficientes, atendiendo a la puntuación total obtenida en todos los escenarios, serían Suecia (el único país eficiente en todos los escenarios), Finlandia, Luxemburgo y Bélgica. Los países menos eficientes en total son Letonia, Rumania, Malta y Estonia (de menos a más eficiencia).

Tabla 7.2. Base de reglas para la transformación de los inputs/outputs de acuerdo a los valores de las variables padre en la red Bayesiana.

RELACIONES DE DEPENDENCIA EN LA RED BAYESIANA (tes el año de referencia 2008)	CONJUNTO DE REGLAS (Ri) ⁽⁶⁾
ERDFUN _{t,2} =(EUSTRAT _{t,2})	<p>R1: IF ERDFUN_{t,2}/GERD_{t,2}<=0.4 AND GERD_{t,2}=3 THEN ERDFUN_{t,2}=ERDFUN_{t,2}*2</p> <p>R2: IF ERDFUN_{t,2}/GERD_{t,2}>0.55 AND GERD_{t,2}<2 THEN ERDFUN_{t,2}=GERD_{t,2}/2</p> <p>R3: IF ERDFUN_{t,2}<=0.2 THEN ERDFUN_{t,2}=ERDFUN_{t,2}/2</p>
PUBPER _{t,2} =(MOBIL _{t,2} , MARKRE _{t,2})	<p>R4: IF PUBPER_{t,2}=LOW AND MOBIL_{t,2}=HIGH AND MARKRE_{t,2}=ATTRACTIVE THEN PUBPER_{t,2}=PUBPER_{t,2}/2</p> <p>R5: IF PUBPER_{t,2}=HIGH AND MOBIL_{t,2}=LOW AND MARKRE_{t,2}=NOT ATTRACTIVE THEN PUBPER_{t,2}=PUBPER_{t,2}*2</p>
TPATRE _t =(ERDFUN _{t,3} , PUBPER _{t,2} , ERDPRI _{t,3} , PERPRI _{t,2} , SPUBLI _t , AGECOL _t , INTECAP _t , ENTREP _t , SBNETW _t , SPILLO _t , INFRAE _t)	<p>R6: IF TPATRE_t=LOW AND ERDFUN_{t,3}=HIGH AND PUBPER_{t,2}=HIGH AND ERDPRI_{t,3}=HIGH AND PERPRI_{t,2}=HIGH AND SPUBLI_t=HIGH AND AGECOL_t=HIGH AND INTECAP_t=HIGH AND ENTREP_t=HIGH AND SBNETW_t=HIGH AND SPILLO_t=HIGH AND INFRAE_t=GOOD AND INSTFR_t=ESTABLE THEN TPATRE_t=TPATRE_t/2</p> <p>R7: IF TPATRE_t=HIGH AND ERDFUN_{t,3}=LOW AND PUBPER_{t,2}=LOW AND ERDPRI_{t,3}=LOW AND PERPRI_{t,2}=LOW AND SPUBLI_t=LOW AND AGECOL_t=LOW AND INTECAP_t=LOW AND ENTREP_t=LOW AND SBNETW_t=LOW AND SPILLO_t=LOW AND INFRAE_t=NOT GOOD AND INSTFR_t=UNSTABLE THEN TPATRE_t=TPATRE_t*2</p> <p>R8: IF TPATRE_t=HIGH AND SPUBLI_t=LOW AND PUBPER_{t,2}=LOW AND ERDFUN_{t,3}=LOW AND ERDPRI_{t,3}=LOW AND PERPRI_{t,2}=LOW THEN TPATRE_t=TPATRE_t*2</p> <p>R9: IF TPATRE_t=HIGH AND ERDPRI_{t,3}=LOW AND PERPRI_{t,2}=LOW AND TPATRE_t=TPATRE_t*2</p> <p>R10: IF TPATRE_t=LOW AND SPUBLI_t=HIGH AND PUBPER_{t,2}=HIGH AND ERDFUN_{t,3}=HIGH AND TPATRE_t=TPATRE_t*2</p>
SPUBLI _t =(ERDFUN _{t,3} , PUBPER _{t,2} , ERDPRI _{t,3} , PERPRI _{t,2} , INFRAE _t , AGECOL _t)	<p>R11: IF SPUBLI_t=LOW AND ERDFUN_{t,3}=HIGH AND PUBPER_{t,2}=HIGH AND ERDPRI_{t,3}=HIGH AND PERPRI_{t,2}=HIGH AND INFRAE_t=GOOD AND AGECOL_t=HIGH THEN SPUBLI_t=SPUBLI_t/2</p> <p>R12: IF SPUBLI_t=HIGH AND ERDFUN_{t,3}=LOW AND PUBPER_{t,2}=LOW AND ERDPRI_{t,3}=LOW AND PERPRI_{t,2}=LOW AND INFRAE_t=NOT GOOD AND AGECOL_t=LOW THEN SPUBLI_t=SPUBLI_t*2</p> <p>R13: IF SPUBLI_t=LOW AND AGECOL_t=1 AND ERDFUN_{t,3}=HIGH AND INFRAE_t=SI AND ERDPRI_{t,3}=HIGH AND SBNETW_t=1 THEN SPUBLI_t=SPUBLI_t/2</p>
GDGRO _t =(RLPGH _t)	<p>R14: IF GDGRO_t=HIGH AND RLPGH_t=LOW THEN GDGRO_t=GDGRO_t*2</p> <p>R15: IF GDGRO_t=LOW AND RLPGH_t=HIGH THEN GDGRO_t=GDGRO_t/2</p> <p>R16: IF GDGRO_t>5 THEN GDGRO_t=GDGRO_t/2</p>

Tabla 7.2. Base de reglas para la transformación de los inputs/outputs de acuerdo a los valores de las variables padre en la red Bayesiana (cont.)

RELACIONES DE DEPENDENCIA EN LA RED BAYESIANA (res el año de referencia 2008)	CONJUNTO DE REGLAS (Ri) ⁽⁶⁾
<p>RLPGH_i = f(UNEMPL_i, LABCLL_i, MOTIV_i, LLEARN_i, PHD06_i, ELEAV_i, INFR4E_i, TECCHA_i)</p>	<p>R17: IF RLPGH_i=HIGH AND UNEMPL_i=HIGH AND LABCLL_i=BAD AND MOTIV_i=UNMOTIVATED AND LLEARN_i=LOW AND PHD06_i=LOW AND ELEAV_i=HIGH AND INFR4E_i=NOT GOOD AND TECCHA_i=WEAK THEN RLPGH_i=RLPGH_i*2</p> <p>R18: IF RLPGH_i=LOW AND UNEMPL_i=LOW AND LABCLL_i=GOOD AND MOTIV_i=MOTIVATED AND LLEARN_i=HIGH AND PHD06_i=HIGH AND ELEAV_i=LOW AND INFR4E_i=GOOD AND TECCHA_i=STRONG THEN RLPGH_i=RLPGH_i/2</p> <p>R19: IF RLPGH_i=HIGH AND TECCHA_i=WEAK AND LLEARN_i=LOW THEN RLPGH_i=RLPGH_i/2</p> <p>R20: IF RLPGH_i=LOW AND MOTIV_i=1 AND LABOCLLM=1 AND TECCHA_i=STRONG AND PHD06_i=HIGH AND ELEAV_i=HIGH AND LLEARN_i=HIGH THEN RLPGH_i=RLPGH_i/2</p> <p>R21: IF RLPGH_i=LOW AND PHD06_i=HIGH AND ELEAV_i=HIGH THEN RLPGH_i=RLPGH_i/2</p>
<p>UNEMPL_i = f(PHD06_i, LLEARN_i, ELEAV_i, GDPGRO_i)</p>	<p>R22: IF UNEMPL_i=LOW AND AND PHD06_i=LOW AND LLEARN_i=LOW AND ELEAV_i=LOW AND GDPGRO_i=LOW THEN UNEMPL_i=UNEMPL_i*2</p> <p>R23: IF UNEMPL_i=HIGH AND PHD06_i=HIGH AND LLEARN_i=HIGH AND ELEAV_i=HIGH AND GDPGRO_i=HIGH THEN UNEMPL_i=UNEMPL_i/2</p>
<p>HITCEX_i = f(COMPET_i)</p>	<p>R24: IF HITCEX_i=HIGH AND COMPET_i=LOW THEN HITCEX_i=HITCEX_i*2</p> <p>R25: IF HITCEX_i=LOW AND COMPET_i=HIGH THEN HITCEX_i=HITCEX_i/2</p> <p>R26: IF HICETEX_i>=30 AND COMPET_i=LOW THEN HICETEX_i=HICETEX_i/3</p>
<p>ELEAV_i = f(EDUEXP_i, SPUBL_i)</p>	<p>R27: IF ELEAV_i=HIGH AND EDUEXP_i=HIGH AND SPUBL_i=HIGH THEN ELEAV_i=ELEAV_i/2</p> <p>R28: IF ELEAV_i=LOW AND EDUEXP_i=LOW AND SPUBL_i=LOW THEN ELEAV_i=ELEAV_i*2</p>
<p>PHD06_i = f(EDUEXP_i, SPUBL_i)</p>	<p>R29: IF PHD06_i=HIGH AND EDUEXP_i=LOW AND SPUBL_i=LOW THEN PHD06_i=PHD06_i*2</p> <p>R30: IF PHD06_i=LOW AND EDUEXP_i=HIGH AND SPUBL_i=HIGH THEN PHD06_i=PHD06_i/2</p>
<p>LLEARN_i = f(EDUEXP_i, SPUBL_i)</p>	<p>R31: IF LLEARN_i=HIGH AND EDUEXP_i=LOW AND SPUBL_i=LOW THEN LLEARN_i=LLEARN_i*2</p> <p>R32: IF LLEARN_i=LOW AND EDUEXP_i=HIGH AND SPUBL_i=HIGH THEN LLEARN_i=LLEARN_i/2</p>
<p>INNEWP_i = f(INTECAP_i, ENTREP_i, SBNETW_i, SPILLOV_i, INSTTR_i, LEARDO_i, BENCHM_i)</p>	<p>R35: IF INNEWP_i=HIGH AND INTECAP_i=LOW AND ENTREP_i=NO AND SBNETW_i=LOW AND SPILLOV_i=LOW AND INSTTR_i=NOT GOOD AND LEARDO_i=LOW AND BENCHM_i=LOW THEN INNEWP_i=INNEWP_i*2</p> <p>R36: IF INNEWP_i=LOW AND INTECAP_i=HIGH AND ENTREP_i=YES AND SBNETW_i=HIGH AND SPILLOV_i=HIGH AND INSTTR_i=GOOD AND LEARDO_i=HIGH AND BENCHM_i=HIGH THEN INNEWP_i=INNEWP_i/2</p>

Tabla 7.2. Base de reglas para la transformación de los inputs/outputs de acuerdo a los valores de las variables padre en la red Bayesiana (cont.)

RELACIONES DE DEPENDENCIA EN LA RED BAYESIANA (res el año de referencia 2008)	CONJUNTO DE REGLAS (R) ^(a)
TR-ADEM _i = f(INTECAP _i , ENTREP _i , SBNETW _i , SPILLOV _i , INSTFR _i , LE-ARDO _i , BENCHM _i)	<p>R37: IF TR-ADEM=HIGH AND INTECAP=LOW AND ENTREP=NO AND SBNETW=LOW AND SPILLOV=LOW AND INSTFR=NOT GOOD AND LE-ARDO=LOW AND BENCHM=LOW THEN TR-ADEM_i=TR-ADEM*2</p> <p>R38: IF TR-ADEM=LOW AND INTECAP=HIGH AND ENTREP=YES AND SBNETW=HIGH AND SPILLOV=HIGH AND INSTFR=GOOD AND LE-ARDO=HIGH AND BENCHM=HIGH THEN TR-ADEM_i=TR-ADEM_i/2</p>
INTUSE _i = f(SPUBL _i , TP-ATRE _i)	<p>R39: IF INTUSE=HIGH AND SPUBL=LOW AND TP-ATRE=LOW THEN INTUSE_i=INTUSE*2</p> <p>R40: IF INTUSE=LOW AND SPUBL=HIGH AND TP-ATRE=HIGH THEN INTUSE_i=INTUSE_i/2</p>
BROSUS _i = f(SPUBL _i , TP-ATRE _i)	<p>R41: IF BROSUS=HIGH AND SPUBL=LOW AND TP-ATRE=LOW THEN BROSUS_i=BROSUS*2</p> <p>R39: IF BROSUS=LOW AND SPUBL=HIGH AND TP-ATRE=HIGH THEN BROSUS_i=BROSUS_i/2</p>
MOBILE _i = f(SPUBL _i , TP-ATRE _i)	<p>R40: IF MOBILE=HIGH AND SPUBL=LOW AND TP-ATRE=LOW THEN MOBILE_i=MOBILE*2</p> <p>R42: IF MOBILE=LOW AND SPUBL=HIGH AND TP-ATRE=HIGH THEN MOBILE_i=MOBILE_i/2</p>
AGEDEP _i = f(SPUBL _i , TP-ATRE _i , HE-AEXP _i)	<p>R43: IF AGEDEP=HIGH AND SPUBL=HIGH AND TP-ATRE=HIGH AND HE-AEXP=HIGH THEN AGEDEP_i=AGEDEP*2</p> <p>R44: IF AGEDEP=LOW AND SPUBL=LOW AND TP-ATRE=LOW AND HE-AEXP=LOW THEN AGEDEP_i=AGEDEP/2</p>
LIFEXP _i = f(SPUBL _i , TP-ATRE _i , HE-AEXP _i)	<p>R45: IF LIFEXP=HIGH AND SPUBL=LOW AND TP-ATRE=LOW AND HE-AEXP=LOW THEN LIFEXP_i=LIFEXP*2</p> <p>R46: IF LIFEXP=LOW AND SPUBL=HIGH AND TP-ATRE=HIGH AND HE-AEXP=HIGH THEN LIFEXP_i=LIFEXP/2</p>
MORR-AT _i = f(SPUBL _i , TP-ATRE _i , HE-AEXP _i)	<p>R47: IF MORR-AT=HIGH AND SPUBL=HIGH AND TP-ATRE=HIGH AND HE-AEXP=HIGH THEN MORR-AT_i=MORR-AT*2</p> <p>R48: IF MORR-AT=LOW AND SPUBL=LOW AND TP-ATRE=LOW AND HE-AEXP=LOW THEN MORR-AT_i=MORR-AT/2</p>
ENINT _i = f(SPUBL _i , TP-ATRE _i)	<p>R49: IF ENINT=HIGH AND SPUBL=HIGH AND TP-ATRE=HIGH THEN ENINT_i=ENINT*2</p> <p>R50: IF ENINT=LOW AND SPUBL=LOW AND TP-ATRE=LOW THEN ENINT_i=ENINT/2</p>
G-ASEMI _i = f(SPUBL _i , TP-ATRE _i)	<p>R51: IF G-ASEMI=HIGH AND SPUBL=HIGH AND TP-ATRE=HIGH THEN G-ASEMI_i=G-ASEMI*2</p> <p>R52: IF G-ASEMI=LOW AND SPUBL=LOW AND TP-ATRE=LOW THEN G-ASEMI_i=G-ASEMI/2</p>
C-ARBDI _i = f(SPUBL _i , TP-ATRE _i)	<p>R53: IF C-ARBDI=HIGH AND SPUBL=HIGH AND TP-ATRE=HIGH THEN C-ARBDI_i=C-ARBDI*2</p> <p>R54: IF C-ARBDI=LOW AND SPUBL=LOW AND TP-ATRE=LOW THEN C-ARBDI_i=C-ARBDI/2</p>

(a) Ver tablas 6.1 y 6.2 para una información detallada sobre las variables.

BLOQUE 3- ANÁLISIS EMPÍRICO

En el caso de Estonia, llama la atención la eficiencia que presenta en el escenario de Sociedad de la Información. De Letonia, destacar su eficiencia en el escenario de Medio Ambiente y, respecto a Rumania, que es eficiente en el escenario Económico. España se encuentra en el lugar 19 respecto a la eficiencia total obtenida, por encima de países como Portugal e Italia.

De las 27 DMUs evaluadas, cuatro políticas son eficientes en el escenario Tradicional (las de Suecia, Finlandia, Bélgica y Lituania), lo que representa aproximadamente el 14,81% de la muestra. A la cola se encuentran Rumania (53,20%), seguida de Letonia (58,50%), Bulgaria (64,80%) y Polonia (69,80%). La eficiencia media es del 51,62% (es decir, por término medio, las DMUs podrían aumentar su producción científica y sus patentes en un 48,38% a partir del consumo de inputs) y, si sólo se consideran las DMUs ineficientes, se reduce dicha media a un 45,17%, por lo que existe un gran margen potencial de mejora en este escenario. Con ineficiencias muy importantes, por tanto, se encuentran en este escenario las políticas de Rumania, Letonia, Bulgaria y Polonia.

El escenario Económico presenta eficiencias en las políticas de 13 países y 8 próximas a la eficiencia, por lo que no resulta un escenario discriminante. Los buenos resultados en este escenario llaman la atención en países como Eslovenia, Chipre, Hungría, Malta y Rumania. El menos eficiente de todos ellos es Letonia. Recordar que el período se refiere a los años 2006-2008, por lo que aún los efectos de la crisis no se han reflejado en los datos con toda su crudeza.

En el escenario de la Innovación, 11 países aparecen como eficientes. De ellos, en relación a su eficiencia en el escenario Tradicional, destacan España y Chipre (con puntuaciones de eficiencia muy bajas en el primero pero eficientes en este escenario concreto). Dado que el escenario centra su atención en las innovaciones no tecnológicas, cuyo origen directo no tiene por qué necesariamente proceder de la inversión en I+D+i, sino que pueden proceder de cualquier otro tipo de actividad y de la propia creatividad humana, abre el debate de qué resulta de mayor interés en un país: ¿ser eficiente en el escenario Tradicional o en escenarios como el de Innovación? En este caso, las empresas parece que encuentran los medios propios para innovar y para incorporar dichas innovaciones al mercado con éxito. El país menos eficiente es de nuevo Letonia, siguiéndole Italia. Esta circunstancia se debe a que, siguiendo las reglas aplicadas a los datos originales, Italia cuenta con un valor “bajo” de marcas comerciales ($TRADEM_{2008}$) en relación al valor de las variables que le influyen, por lo que la regla R38 se aplicó al valor original y lo penalizó reduciéndolo a la mitad (tabla 7.2).

BLOQUE 3- ANÁLISIS EMPÍRICO

En el escenario de Educación, 6 políticas alcanzan la eficiencia, destacando el caso de Eslovaquia. Muy de cerca les siguen Polonia, Austria y Eslovenia. La menos eficiente, alejada de las demás, es la política de Malta (29,79). Eslovaquia cuenta con un nivel muy bajo de abandonos escolares en edades tempranas (*ELEAV*₂₀₀₈), lo que unido a que su gasto total en educación y de que el número de publicaciones es bajo, fue de aplicación la regla R28 (que duplicó su valor “bonificándolo” para transferir la información a la metodología, haciéndole entender que se trata de un dato muy bueno, más de lo que el valor en sí puede representar). Puede deberse a que por la creciente armonización entre los sistemas de Educación Superior en los países europeos, algunos (entre ellos Eslovaquia) han visto crecer rápidamente sus tasas de graduación.

La eficiencia alcanzada por las políticas de I+D+i europeas en el escenario de Sociedad de la Información es clara: 9 países de los 27 alcanzan la eficiencia total, 3 se encuentran muy próximos y 5 se encuentran en un intervalo de puntuación de eficiencia entre 80-90. Las políticas menos eficientes son las de Eslovenia y la de Letonia. Irlanda aparece como uno de los países menos eficientes porque fue bonificado el valor del número de publicaciones científicas, *SPUBLI*₂₀₀₈ (aplicación de la regla R12, dado que si el valor de las variables de las que depende es considerado “bajo”, según el intervalo fijado por los expertos, y el valor de *SPUBLI*₂₀₀₈ es “alto”, es doblado su valor para hacer entender lo importante de este valor a la metodología DEA). Con el valor de este input incrementado, el valor de los outputs obtenidos conlleva dicha ineficiencia.

El uso de las TICs es uno de los pilares de la sociedad del conocimiento, por lo que cabría afirmar que los recursos dedicados a las actividades de I+D+i están cumpliendo su cometido de fomentar el uso de este tipo de tecnologías, con las implicaciones que conlleva (relaciones sociales, incremento de la productividad de los trabajadores, etc.), aunque sigue quedando camino por recorrer en algunos países.

Resultados similares encontramos en el escenario de la Salud. Respecto a los outputs seleccionados, encontramos que 10 países son eficientes y 6 están muy próximos a la eficiencia total. No se encuentran en este escenario valores tan bajos como en el resto de escenarios, siendo la menor puntuación la de Dinamarca (74,83). Este dato sorprende en un país que es eficiente en 4 de los 7 escenarios y que casi es eficiente en los restantes. Nuevamente, nos encontramos con la aplicación de la base de reglas que penalizó el valor de *AGEDEP*₂₀₀₈ (ratio de dependencia de las personas mayores) y de *LIFEXP*₂₀₀₈ (esperanza de vida) en el caso de

BLOQUE 3- ANÁLISIS EMPÍRICO

Dinamarca (multiplicando los valores por dos, dado que son outputs que deben interpretarse que cuanto mayor es su valor, es peor resultado).

Finalmente, respecto al escenario de Medio Ambiente, 8 países son eficientes (entre los que se encuentran Letonia y Portugal), 13 se encuentran muy próximos a ella y los menos eficientes son Estonia y República Checa.

Para resumir y analizar las puntuaciones de eficiencia, además de la breve descripción realizada en los párrafos anteriores, se realizó un agrupamiento de las políticas utilizando el algoritmo k -medias y se consideró $k=4$ de forma experimental (ver apartado 2 del siguiente capítulo IX para un mayor detalle sobre este algoritmo):

- El grupo 1 incluye Bulgaria, Estonia, Francia, Italia, Letonia, Polonia y Rumania. Sus políticas de I+D+i en el escenario Tradicional son las menos eficientes. También aparecen como los menos eficientes en los escenarios de Innovación y en el de la sociedad de la Información, además de obtener puntuaciones bajas en el escenario de Educación. Resulta llamativa la pertenencia a este grupo de las políticas de Francia e Italia. Esto se debe a la pobre relación que presentan entre los esfuerzos públicos de inversión en I+D+i y los outputs tradicionales (patentes y publicaciones científicas) además de por los bajos resultados en los escenarios de Educación e Innovación.
- El grupo 2 incluye las políticas públicas de I+D+i más eficientes. Agrupa las políticas de Austria, Bélgica, Dinamarca, Finlandia, Alemania, Irlanda, Lituania, Luxemburgo, Holanda, Suecia y Reino Unido. Como vimos, Suecia es eficiente en todos los escenarios bajo estudio seguida de Bélgica (seis escenarios) y Finlandia (cinco escenarios). Irlanda y Luxemburgo aparecen entre las cinco primeras puntuaciones del ranking en cuatro de los siete escenarios. Lituania aparece en este grupo, principalmente, por la eficiencia presentada en los escenarios Tradicional y Económico. A pesar de la relativamente baja cantidad de recursos dedicados a I+D+i por el sector público, este país muestra un elevado número de publicaciones científicas, circunstancia que sería de interés estudiar detenidamente. De los once países que pertenecen a este grupo, cuatro de ellos son eficientes en el escenario Tradicional (puntuación=100). La mayoría de políticas restantes, aunque no alcanzan

BLOQUE 3- ANÁLISIS EMPÍRICO

la máxima eficiencia, sí que se encuentran cerca del valor 100 (a excepción de Reino Unido). Este país, curiosamente, sólo aparece como eficiente en el escenario de Salud.

- El grupo 3 incluye las políticas de Chipre, República Checa, Hungría, Eslovaquia, Eslovenia y España. Muestran bajas eficiencias en los escenarios Tradicional y de Medioambiente pero tienen muy buenos resultados en el escenario de Innovación, que incluye las innovaciones no tecnológicas.
- Finalmente, el grupo 4 incluye a las políticas de I+D+i de Grecia, Malta y Portugal. Agrupa las políticas menos eficientes mostrando bajas puntuaciones en el escenario Tradicional y obteniendo las eficiencias más bajas en los escenarios de Educación y de Salud. Sin embargo, son eficientes en el escenario de Sociedad de la Información debido al elevado número de suscriptores de teléfonos móviles (países que muestran las mayores tasas de aumento en el uso de las TICs).

Para discriminar entre los países más eficientes, se usó una medida complementaria consistente en analizar el número de veces que las DMUs fueron elegidas como *peers* o referencia en su grupo, para ser comparadas con otras DMUs en la evaluación DEA. Suecia, Bélgica, Finlandia e Irlanda fueron los seleccionados con mayor frecuencia por las DMUs ineficientes como *benchmarks* o referencia. Lógicamente, estos países pertenecen al grupo 2 y 3 y son eficientes en el escenario tradicional (Irlanda, casi eficiente con una puntuación de 96,16%). Como *peers*, por escenarios concretos, los que aparecen con mayor frecuencia son España en el de Innovación, Finlandia en el de Educación, Italia en el de Salud, y Grecia en el de Sociedad de la Información.

El DEA proporciona los objetivos óptimos de outputs y de inputs que las DMUs ineficientes deben alcanzar para ser catalogadas como eficientes. Estos objetivos se calculan aplicando directamente el índice de eficiencia y las correspondientes variables de holgura a los valores reales de los inputs y los outputs de cada entidad ineficiente. Por tanto, para que una DMU pase a ser eficiente, no sólo debe corregir las ineficiencias radiales (representadas por el índice de eficiencia) sino también las no radiales (representadas por las variables de holgura).

La tabla 7.4 recoge los incrementos en los valores de los inputs y los outputs sobre los objetivos (*targets*) de los mismos, por escenario, para alcanzar la eficiencia (en términos de aumento o disminución porcentual). Debe señalarse que, al ser un modelo orientado a la maximización de los outputs, para que un país fuese eficiente, el incremento en todos sus

BLOQUE 3- ANÁLISIS EMPÍRICO

outputs debería ser igual a cero (o lo que es igual, el valor de la variable y el valor objetivo coinciden).

Las holguras son las cantidades de los inputs y los outputs que son excesivas (o deficitarias) y que deberían reducir el valor de los inputs (o incrementar el valor de los outputs) para que se conviertan en óptimas. Hay dos tipos de reducciones o incrementos: (1) la reducción proporcional de inputs (θ) (orientando a inputs) o el incremento proporcional de outputs (ϕ) (enfoque orientado a los outputs) y (2) el movimiento de holguras.

El valor objetivo que arroja el análisis DEA incorpora tanto el incremento proporcional como el movimiento de holguras. El porcentaje de ganancia se obtiene del siguiente ratio: (valor objetivo-valor de la variable)/(valor de la variable)*100. En otras palabras, para que un país sea eficiente, el porcentaje de ganancia para todas sus variables (p.e., outputs en el DEA orientado a los outputs) debería ser cero. Si la ganancia (%) es mayor que cero, esto significa que el valor original de la variable es menor al óptimo que debería usar el país (política de I+D+i). El significado de la ganancia es el incremento de la producción, en términos de outputs, para que la política analizada se torne eficiente y ese es en concreto el valor que se ha recogido en la tabla 7.4, detallado por escenarios¹⁷.

En el escenario Tradicional, la variable que en una mayor proporción habría que incrementar para obtener la eficiencia técnica relativa correspondiente es el número de patentes ($TPATRE_{2008}$). Los que mayor incremento deberían realizar serían Eslovaquia, Bulgaria y Rumania (los países menos eficientes en todos los escenarios).

En el escenario Económico, la variable que debe incrementarse en mayor medida, en general, es $HICETEX_{2008}$ (exportaciones de alta tecnología). El país que más debe incrementar los valores de las variables seleccionadas para conseguir la eficiencia es precisamente España y, de forma muy especial, el valor del nivel de desempleo ($UNEMPL_{2008}$), con un nivel no comparable a ninguno del resto de países. Otros países que se encuentran en una situación parecida de incrementos importantes en los outputs de este escenario son Italia, Estonia y Letonia, Portugal y Grecia.

En el escenario de Innovación, Letonia debería incrementar las tres variables representativas de las innovaciones no tecnológicas ($INNEWP_{2008}$, ventas de innovaciones nuevas para el

¹⁷ Los valores anormalmente elevados que aparecen en la tabla se deben a que el incremento porcentual se aplica a un valor muy pequeño y cualquier aumento da lugar a un gran porcentaje.

BLOQUE 3- ANÁLISIS EMPÍRICO

mercado y nuevas para la empresa como % de los ingresos, $INNSME_{2008}$, PYMEs que introducen innovaciones de marketing u organizacionales como porcentaje del total de PYMEs, y $TRADEM_{2008}$, marcas comunitarias por billón de euros del PIB), siguiéndole de cerca Italia y Lituania. Llama la atención, en el caso de $INNEW P_{2008}$, que el mayor incremento debe realizarlo Reino Unido y, además, muy alejado del incremento que debe realizar el siguiente país (Letonia).

Para el escenario de Educación encontramos situaciones con que los países no eficientes deben incrementar las tres variables seleccionadas, en general. Malta y Luxemburgo presentan la necesidad de aumentar de una forma muy importante los abandonos tempranos en la educación ($ELEAV_{2008}$) y el número de graduados doctores ($PHD06_{2008}$), respectivamente. En la variable $LLEARN_t$ (aprendizaje a lo largo de la vida) no aparecen valores tan altos como los anteriores. Los países que deberían incrementar este valor, junto a los anteriores, para alcanzar la eficiencia son Bulgaria, Grecia, Rumania y Malta.

En el caso del escenario de Medio Ambiente, no destaca ninguna variable para que sea disminuida (no incrementada, como aparece en la tabla, dado que se trata de un output que cuanto menor sea su valor, mejor, y así debe ser interpretado el término de la ganancia al haberse transformado linealmente dicho output). En cada una de las variables, sí destaca un país que tiene que disminuir el valor del indicador de forma casi exponencial: $ENINT_{2008}$ (intensidad energética de la economía de un país) y $CARDBI_{2008}$ (daños procedentes del dióxido de carbono), en Bulgaria y, la variable $GASEM_{2008}$ (emisiones totales de gases invernadero), en Finlandia. Lo comentado no significa de forma categórica que estos países sean los que más contaminan sino que, respecto al número de patentes y publicaciones que cuentan, deberían mejorar los valores de estas variables para usar dichos recursos de forma eficiente.

Resultados muy similares al escenario de Medio Ambiente se encuentran en el escenario de la Salud. Es un escenario en el que no deben realizarse grandes incrementos en ninguno de los outputs de forma especial, salvo en el caso de Finlandia ($AGEDEP_{2008}$, índice de dependencia) y en Rumania ($MORRAT_{2008}$ o tasa de mortalidad de niños menores de un año, por cada 1000 habitantes). Este escenario sería, junto al de Medio Ambiente, el que menos variaciones porcentuales de los outputs requeriría para alcanzar la eficiencia técnica relativa.

Para terminar con este análisis, en el escenario de Sociedad de la Información o de las TICs, Eslovenia, Letonia e Irlanda son los tres países que más deben incrementar las tres variables

BLOQUE 3- ANÁLISIS EMPÍRICO

seleccionadas para analizar la eficiencia de este escenario ($INTUSE_{2008}$, $BROSUS_{2008}$ y $MOBILE_{2008}$, número de usuarios de internet, número de suscriptores de banda ancha y número de usuarios de móvil, respectivamente).

En todo caso, estos porcentajes de ganancia deben interpretarse con cautela. Resulta arriesgado suponer que las ineficiencias de las respectivas DMUs evaluadas podrían eliminarse automáticamente adaptando sus outputs a estos incrementos/decrementos.

Finalmente, los rankings obtenidos a través del enfoque 2LS-DEA se compararon con las puntuaciones del Cuadro de Indicadores de Innovación de la UE (IUS) por dimensiones (European Commission, 2010), en la medida que son comparables con los escenarios Tradicional, de Innovación y Económico. Además del Índice Sintético de Innovación (SII), que se construye en base a los datos de 24 indicadores normalizados, se calcula otro índice por temática concreta, que es el empleado para la comparación. En el informe de 2010, se refleja la ejecución de los datos de 2008/2009 y el informe de 2011 se refiere a los años 2009/2010 debido al retraso en la disponibilidad de los datos; por ello, se utiliza el informe del año 2010 en esta comparativa.

Los resultados del test no paramétrico U de Mann–Whitney Whitney (p -valor menor que 0,05 para todas las comparaciones) mostraron algunas discrepancias relevantes con los resultados obtenidos en el análisis que ponen de manifiesto la utilidad de incorporar conocimiento experto gracias a la aplicación de la base de reglas. Por ejemplo, la puntuación de eficiencia de Lituania puede diferir de la clasificación de IUS porque los valores de algunas variables fueron transformados al aplicar las reglas. De esta forma, en el escenario Tradicional, el valor de $ERDFUN_{2006}$ de Lituania se penalizó por encontrarse lejos de cumplir los objetivos de la Estrategia de Lisboa y se bonificó el valor de $PUBPER_{2006}$ multiplicándolo por dos debido al elevado número de personal público dedicado a la I+D+i que presenta (aplicación de las reglas R1, R2 y R3 en el conjunto de datos original; ver tabla 7.2). Este país es eficiente en el escenario Tradicional pero no lo es en los escenarios socioeconómicos, excepto el Económico.

Otro país donde los resultados de eficiencia difieren de las puntuaciones del IUS es Francia. En este caso, la variable $SPUBLI_{2008}$ se penalizó mediante la regla R11, principalmente porque el número de publicaciones científicas era *bajo* (<1.5 por cada 1.000 habitantes) respecto a los esfuerzos públicos realizados ($ERDFUN_{2006}$ y $PUBPER_{2006}$).

Tabla 7.3. Eficiencia técnica relativa de las políticas públicas de I+D+I de los países de la UE (%) (DMUs) por escenario (100 = eficiente). Periodo 2006-2008.

DMUs	Escenario Tradicional	Escenario Económico	Escenario de la Innovación	Escenario Educación	Escenario Sociedad Información	Escenario Salud	Escenario Medio Amb.
Alemania	84,04	99,47	100	86,83	84,2	88,04	97,13
Austria	64,01	97,09	70,39	99,55	100	84,1	94,83
Bélgica	100	100	100	100	61,01	100	100
Bulgaria	11,4	98,77	43,66	70,18	83,81	98,27	92,69
Chipre	26,39	100	100	73,49	78,23	98,85	83,72
Dinamarca	85,79	100	92,04	100	95,84	74,83	100
Eslovaquia	21,11	96,58	96,84	100	100	76,10	94,74
Eslovenia	55,52	100	100	99,4	43,93	100	80,98
España	27,87	81,83	100	71,54	74,98	85,2	97,18
Estonia	30,77	87,69	40,91	93,01	100	91,44	50,63
Finlandia	100	100	100	100	100	98,3	85,36
Francia	38,55	96,66	45,53	86,47	100	100	100
Grecia	38,52	83,92	100	70,18	100	88,92	98,73
Holanda	86,81	100	73,24	80,65	80,06	100	93,93
Hungría	20,61	100	90,16	79,52	57,3	86,08	98,38
Irlanda	96,16	100	100	100	100	98,29	100
Italia	32,57	82,75	38,9	58,26	45,2	100	100
Letonia	9,01	78,87	38,57	68,07	94,06	77,83	100
Lituania	100	100	76,94	70,18	100	79,14	98
Luxemburgo	98,68	100	100	74,4	79,55	100	99,3
Malta	18,07	100	61,03	29,79	98,55	100	99,46
Polonia	16,48	96,78	43,7	99,7	75,31	100	71,02
Portugal	28,08	84,61	96,64	33,92	88,22	83,84	100
Reino Unido	63,02	90,12	79,7	89,95	63,93	84,74	97,06
Rep. Checa	31,65	97,96	100	93,44	87,1	100	56,96
Rumania	8,57	100	45,81	66,87	69,8	97,4	96,38
Suecia	100	100	100	100	100	100	100

Tabla 7.4. Porcentaje de variación de los inputs y los outputs para lograr la eficiencia, por escenario

DMUs	TRADICIONAL						ECONÓMICO					
	ERDFUN	SUMPER	TPATRE	SPUBLI	TPATRE	SPUBLI	TPATRE	SPUBLI	GDPGROW	RPLGH	HITECEX	UNEMPL
Alemania	-99,89	0,00	18,98	91,91	0,00	0,00	-35,04	0,00	0,53	3,43	44,56	29,02
Austria	0,00	-3,36	97,09	56,23	-40,26	-8,78	0,00	0,00	2,99	8,63	56,13	2,99
Bélgica	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Bulgaria	-40,95	0,00	1978,12	777,38	-1,94	-11,20	0,00	0,00	1,25	9,05	1,25	1,25
Chipre	-39,39	-1,72	278,92	278,92	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Dinamarca	-37,53	0,00	16,56	16,56	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Eslovaquia	-45,57	0,00	2179,35	373,67	-1,52	0,00	0,00	0,00	3,55	8,49	22,81	30,98
Eslovenia	-33,42	0,00	260,41	80,12	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
España	-78,68	0,00	801,52	258,84	0,00	0,00	0,00	0,00	22,20	23,32	122,42	16837781,39
Estonia	-24,92	0,00	1054,01	225,04	-83,52	-40,98	0,00	0,00	30,49	43,75	256,36	14,04
Finlandia	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Francia	-99,31	0,00	159,40	276,93	-58,95	-58,13	0,00	0,00	3,46	24,84	3,46	27,04
Grecia	-41,73	0,00	1246,84	159,64	-14,41	-4,52	0,00	0,00	19,16	39,49	19,16	19,16
Hungría	-24,23	0,00	670,05	385,19	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Irlanda	-11,45	-7,18	4,00	4,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Italia	-27,94	0,00	207,04	207,04	-39,58	-18,09	0,00	0,00	20,85	34,63	133,17	20,85
Letonia	-62,80	0,00	1010,12	1010,12	-38,62	-32,88	0,00	0,00	27,86	26,79	99,37	26,79
Lituania	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Luxemburgo	-99,98	0,00	1,34	166,13	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Malta	-73,04	-0,75	453,35	453,35	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Países Bajos	-94,96	0,00	15,20	15,20	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Polonia	-33,24	0,00	1218,31	506,91	0,00	-1,22	0,00	0,00	3,33	3,33	140,43	10,40
Portugal	-9,07	0,00	1134,14	256,06	-5,50	0,00	0,00	0,00	18,19	20,31	18,19	22,89
Reino Unido	-43,39	0,00	92,34	58,69	-70,43	-26,62	0,00	0,00	10,96	15,91	10,96	10,99
República Checa	-25,32	0,00	868,05	215,91	-41,28	-25,97	0,00	0,00	2,09	2,09	2,09	2,09
Rumania	-19,86	-7,33	1432,16	1067,10	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
Suecia	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00

Tabla 7.4. Porcentaje de incremento/disminución de los inputs y los outputs para lograr la eficiencia por escenario (cont.)

DMUs	INNOVACIÓN						EDUCACIÓN					
	TPATRE	SPUBLI	INNEWP	INNSME	TRADEM	TPATRE	SPUBLI	ELEAV	PHD06	LLEARN		
Alemania	0,00	0,00	0,00	0,00	0,00	0,00	-36,60	17,96	15,17	84,31		
Austria	-9,86	0,00	42,06	198,16	42,06	-98,60	-32,28	0,45	8,59	61,39		
Bélgica	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00		
Bulgaria	-9,63	-13,19	129,02	129,02	129,02	-99,23	-58,39	42,49	264,31	339,66		
Chipre	0,00	0,00	0,00	0,00	0,00	-99,13	-50,76	36,07	784,60	136,11		
Dinamarca	0,00	0,00	8,65	89,73	8,65	0,00	0,00	0,00	0,00	0,00		
Eslovaquia	-14,60	-35,54	3,26	3,26	162,97	0,00	0,00	0,00	0,00	0,00		
Eslovenia	0,00	0,00	0,00	0,00	0,00	-99,10	-28,44	0,61	43,34	36,36		
España	0,00	0,00	0,00	0,00	0,00	-24,35	-24,01	361,79	39,77	39,77		
Estonia	-32,15	0,00	144,43	144,43	144,43	-73,49	-46,52	34,65	67,04	7,52		
Finlandia	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00		
Francia	0,00	-17,00	119,63	154,88	119,63	-34,60	-33,28	15,64	15,64	60,68		
Grecia	0,00	0,00	0,00	0,00	0,00	-99,22	-43,80	42,49	129,54	259,15		
Hungría	-1,33	-15,58	10,92	10,92	52,79	-99,20	-54,01	25,76	152,67	203,57		
Irlanda	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00		
Italia	0,00	-3,01	157,07	157,07	157,07	-71,95	-33,59	71,65	71,65	119,86		
Letonia	-6,99	-9,20	420,47	159,25	159,25	-99,20	-59,79	46,90	368,18	112,50		
Lituania	-7,82	0,00	232,12	29,97	175,29	-99,24	-55,38	42,49	161,38	105,65		
Luxemburgo	0,00	0,00	0,00	0,00	0,00	-51,11	-50,61	34,41	1651,22	91,73		
Malta	-21,87	-6,81	63,86	63,86	63,86	-70,52	-55,03	16225000,00	651,31	235,71		
Países Bajos	0,00	0,00	93,63	36,54	36,54	-76,03	-12,22	24,00	41,50	24,00		
Polonia	-12,69	-0,70	202,59	128,83	128,83	-99,23	-55,88	0,30	94,98	79,58		
Portugal	-8,78	0,00	3,48	3,48	3,48	-45,41	-24,85	1012,63	194,80	194,80		
Reino Unido	0,00	0,00	624,07	25,47	101,24	-92,69	-19,01	11,17	11,17	15,45		
República Checa	0,00	0,00	0,00	0,00	0,00	-35,27	-10,44	7,02	7,02	27,60		
Rumania	-13,18	-35,97	118,28	118,28	365,90	-99,24	-59,18	49,55	88,89	254,17		
Suecia	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00		

Tabla 7.4. Porcentaje de incremento/disminución de los inputs y los outputs para lograr la eficiencia por escenario (cont.)

DMUs	ESCENARIO SALUD						ESCENARIO MEDIO AMBIENTE					
	TPATRE	SPUBLI	AGEDEP	MORRAT	LIFEXP	TPATRE	SPUBLI	ENINT	GASEMI	CARBDI		
Alemania	0,00	0,00	13,58	13,58	13,58	0,00	-36,20	2,95	2,95	13,36		
Austria	0,00	0,00	18,91	18,91	18,91	0,00	0,00	5,45	5,45	7,08		
Bélgica	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00		
Bulgaria	-1,13	-5,69	1,76	124,99	10,64	-3,63	0,00	327739770,58	7,89	340905,39		
Chipre	-5,17	-5,90	1,16	1,16	2,45	-5,63	-5,93	19,44	49,51	33,26		
Dinamarca	0,00	0,00	179,93	166,25	33,64	0,00	0,00	0,00	0,00	0,00		
Eslovaquia	-6,50	-6,72	31,41	31,41	31,41	-2,41	0,00	67,03	5,55	53,91		
Eslovenia	0,00	0,00	0,00	0,00	0,00	-30,02	0,00	23,48	23,48	24,62		
España	-50,52	-20,77	17,38	17,38	17,38	-19,01	0,00	2,90	2,90	3,02		
Estonia	-4,32	0,00	9,37	9,37	9,37	-0,47	0,00	104,23	97,51	798,77		
Finlandia	0,00	0,00	4740925,57	1,73	1,73	0,00	-7,00	55,90	9820884949,53	17,15		
Francia	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00		
Grecia	-6,76	0,00	12,46	12,46	12,46	-16,78	0,00	1,28	40,51	25,41		
Hungría	-0,48	-4,66	16,17	16,17	16,17	0,00	0,00	30,09	1,65	45,26		
Irlanda	0,00	0,00	1,74	2,17	1,74	0,00	0,00	0,00	0,00	0,00		
Italia	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00		
Letonia	0,00	-9,86	28,49	28,49	28,49	0,00	0,00	0,00	0,00	0,00		
Lituania	-11,49	-23,92	26,36	26,36	26,36	-3,49	0,00	31,84	2,04	32,15		
Luxemburgo	0,00	0,00	0,00	0,00	0,00	0,00	-54,46	0,70	1213,55	21,38		
Malta	0,00	0,00	0,00	0,00	0,00	-7,70	0,00	0,54	0,54	33,77		
Países Bajos	0,00	0,00	0,00	0,00	0,00	0,00	-13,45	6,46	485,29	9,16		
Polonia	0,00	0,00	0,00	0,00	0,00	-23,39	0,00	40,81	40,81	139,15		
Portugal	-76,27	0,00	19,27	19,27	23,55	0,00	0,00	0,00	0,00	0,00		
Reino Unido	0,00	0,00	18,01	18,01	18,01	0,00	0,00	3,03	3,03	8,30		
República Checa	0,00	0,00	0,00	0,00	0,00	0,00	0,00	81,97	75,57	139,60		
Rumania	-1,25	-7,47	2,67	5400000,00	7,77	-4,05	0,00	95,51	3,76	70,70		
Suecia	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00		

Tabla 7.4. Porcentaje de incremento/disminución de los inputs y los outputs para lograr la eficiencia por escenario (cont.)

País	ESCENARIO SOCIEDAD DE LA INFORMACIÓN					
	TPATRE	SPUBLI	INTUSE	BROSUS	MOBILE	
Alemania	0,00	0,00	0,00	0,00	0,00	0,00
Austria	0,00	-61,73	63,89	63,89	75,65	75,65
Bélgica	-1,87	-7,34	19,32	50,11	19,32	19,32
Bulgaria	0,00	-13,41	32,08	27,82	27,82	27,82
Chipre	0,00	-11,48	14,81	45,32	14,81	14,81
Dinamarca	0,00	-8,25	5,67	4,34	4,34	4,34
Eslovaquia	0,00	0,00	0,00	0,00	0,00	0,00
Eslovenia	0,00	0,00	0,00	0,00	0,00	0,00
España	0,00	0,00	0,00	0,00	0,00	0,00
Estonia	0,00	-22,49	18,76	19,60	18,76	18,76
Finlandia	0,00	0,00	0,00	0,00	0,00	0,00
Francia	0,00	-15,89	24,91	41,54	24,91	24,91
Grecia	0,00	-60,42	74,52	91,46	74,52	74,52
Hungría	0,00	0,00	0,00	0,00	0,00	0,00
Irlanda	0,00	-96,32	121,22	121,22	131,47	131,47
Italia	-2,73	-3,95	6,32	13,78	6,32	6,32
Letonia	0,00	0,00	0,00	0,00	0,00	0,00
Lituania	-21,64	-32,30	59,83	25,71	25,71	25,71
Luxemburgo	0,00	-6,15	3,92	1,48	1,48	1,48
Malta	-0,86	-19,34	32,78	134,47	32,78	32,78
Países Bajos	-0,92	-13,48	30,31	13,35	13,35	13,35
Polonia	-2,03	-15,91	54,28	43,28	43,28	43,28
Portugal	0,00	0,00	0,00	0,00	0,00	0,00
Reino Unido	0,00	-86,33	127,65	130,35	127,65	127,65
República Checa	0,00	-23,33	33,37	33,37	33,37	33,37
Rumania	0,00	0,00	0,00	0,00	0,00	0,00
Suecia	0,00	-67,60	56,43	56,43	56,43	56,97

4. CONCLUSIONES DEL ANÁLISIS DE LA EFICIENCIA TÉCNICA RELATIVA DE LAS POLÍTICAS DE I+D+I EN EL MARCO DE UN DESARROLLO SOSTENIBLE

1. A través del modelo conceptual causal propuesto en el Bloque 2 y el enfoque del DEA en dos pasos encadenados (2LS-DEA), se han abordado los principales retos para la evaluación de los impactos públicos de la I+D+i en la sociedad: i) la identificación de relaciones entre los inputs de I+D+i, sus outputs e impactos socioeconómicos; ii) la existencia de relaciones directas o unidireccionales entre los inputs y los outputs (nuestro modelo sólo incluye relaciones directas causales); iii) la evaluación de qué proporción de los beneficios sociales debería ser atribuida a los inputs de I+D+i; iv) el efecto del tiempo cuando se miden estos impactos (el denominado *retardo*) y, finalmente, v) el uso de indicadores apropiados (selección de variables).

2. Los puntos cruciales en este diseño fueron i) el seguimiento de la metodología de incorporación de adquisición del conocimiento experto, EbCA; ii) el pre-procesamiento de los datos utilizando la base de reglas obtenida en los pasos de la metodología EbCA ; iii) la interpretación de los resultados del enfoque DEA en dos pasos (2LS-DEA). Iterativamente, el modelo conceptual causal y la base de reglas fueron mejorados hasta que no se encontraron inconsistencias en los resultados del enfoque 2LS-DEA.

3. Una de la ventajas que ha presentado esta metodología es transferir el conocimiento experto al análisis de los datos para la elicitación de conocimiento implícito, siendo una herramienta operacional en lo que se refiere a la provisión de información fundamental para los procesos de toma de decisiones. Así, los resultados no dependen de asunciones subjetivas en los pesos del modelo matemático DEA, ni de la selección arbitraria de inputs y outputs, sino de la incorporación del conocimiento experto en la base de datos mediante la base de reglas que se construyó en base al modelo conceptual causal presentado.

4. Otras metodologías estudian los resultados de la I+D+i o, de forma más amplia, de la competitividad y de los logros respecto a una economía basada en el conocimiento arrojando resultados comparables, pero no desde este enfoque que pretende ser multidimensional y global. Además, no tienen en cuenta las relaciones de causalidad entre las variables y no interpretan las combinaciones especiales de variables (escenarios) que resultan de políticas

BLOQUE 3- ANÁLISIS EMPÍRICO

públicas específicas (algunas veces ineficientes). Esto fortalece la utilidad del modelo propuesto.

5. Con el empleo del citado enfoque, se evaluó la eficiencia técnica relativa (final) de las políticas públicas de I+D+i en los 27 países de la UE. En relación a los resultados concretos, las políticas de Suecia, Finlandia, Luxemburgo, Bélgica, Irlanda y Dinamarca se encuentran en las posiciones del ranking de eficiencia más altas. Los que menores valores de eficiencia presentaron fueron Bulgaria (este país en concreto no es eficiente en ningún escenario), Estonia, Letonia y Rumania.

6. En el momento en que los escenarios socioeconómicos examinados están relacionados con políticas públicas, los resultados empíricos deberían ser de interés para los responsables políticos a la hora de monitorizar la realización de las estrategias dirigidas hacia un crecimiento sostenible e inteligente.

7. Con el modelo conceptual y el enfoque 2LS-DEA es posible identificar las políticas públicas de I+D+i más eficientes en los escenarios socioeconómicos diseñados, lo que abre una puerta para la práctica conocida como *benchmarking*. Por otro lado, el modelo puede proveer a los principales grupos de interés de las dimensiones que necesitan mejorar, lo que puede guiar una política racional para aumentar o disminuir la cantidad de fondos dedicados a I+D+i.

8. El modelo conceptual empleado sería diferente si se hubieran asumido otras teorías y concepciones y se hubiera diseñado para países con un nivel de desarrollo distinto. Sin embargo, el principal objetivo de analizar la eficiencia de las políticas públicas de I+D+i en la UE mediante el uso del enfoque 2LS-DEA, apoyándonos en dicho modelo, fue alcanzado satisfactoriamente a la vista de los resultados congruentes con otros estudios como el del IUS.

9. El modelo puede ser fácilmente ajustado en términos de las variables seleccionadas, de las reglas empleadas según la base de conocimiento experto, de acuerdo con otras perspectivas, para otros países y para diferentes escenarios, considerando indicadores *proxy* alternativos.

CAPÍTULO VIII. ANÁLISIS DE LOS RESULTADOS DE LA I+D+I EN LA UE.

1. INTRODUCCIÓN

Como se ha establecido en anteriores apartados de esta investigación, las actividades de I+D+i se han convertido en claves para un crecimiento económico inteligente, en el contexto de una sociedad basada en el conocimiento, así como la fuerza motora para obtener ventajas competitivas nacionales (Watts y Porter, 1997).

Los estudios previos generalmente evalúan la eficiencia/ejecución de las inversiones en I+D+i usando principalmente indicadores estadísticos y técnicas no paramétricas, como el Análisis Envoltente de Datos empleado en el capítulo VII anterior, tanto a nivel macroeconómico (Lokkainen et al., 2009; Conte, 2009) como a nivel microeconómico en relación a programas, proyectos y empresas (Ojanen y Vuola, 2003)¹⁸. Resulta evidente que los propósitos de la medición y el contexto en el que estos conceptos resultarían adecuados puede ser muy variado (Kerssens-van Drongelen y A. Cook).

Los llamados indicadores compuestos han sido aplicados para medir la ejecución de la ciencia, la tecnología e innovación nacionales, siendo las comparaciones entre países un amplio campo de aplicación de tales indicadores¹⁹. Los académicos, los gobiernos y las organizaciones internacionales, con sus bien conocidos esfuerzos de normalización de los indicadores y conceptos (la OECD, la National Science Foundation o NSF y la UE), adoptan ampliamente este enfoque para elaborar recomendaciones (Arcibugui y Coco, 2005). Estos indicadores presentan dos facetas complementarias: i) están formados por otros sub-indicadores, y (ii) proveen de una síntesis de sub-indicadores mediante su agregación a través de medias aritméticas (ponderadas y no ponderadas) (Cerulli y Filipetti, 2012).

En cuanto a la revisión de los métodos, a nivel de Europa hay que volver a mencionar el especial estudio de la ejecución de I+D+i o, más ampliamente, de la competitividad, que

¹⁸ Para una revisión de los indicadores compuestos, ver Ojanen, V., y Vuola, O. (2003).

¹⁹ Las revisiones y contribuciones acerca de esta metodología se encuentran en el Centro de Investigación Conjunta de la Comisión Europea, <http://composite-indicators.jrc.ec.europa.eu/>.

BLOQUE 3- ANÁLISIS EMPÍRICO

contiene el Cuadro de Indicadores de Innovación de la Unión respecto a los países de la UE y otros países adicionales (ver apartado 3 del capítulo VII).

Otro amplio abanico de metodologías e indicadores para evaluar los impactos de la I+D+i ha sido presentado como un conjunto de instrumentos o herramientas en un informe auspiciado por la Comisión Europea (Fahrenkrog, et al., 2002). El informe presenta la evaluación desde la perspectiva del usuario y destaca las expectativas de los diferentes participantes describiendo aspectos de la evaluación en cuatro áreas políticas, p.e., la financiación de la I+D+i, la provisión de infraestructuras para la misma, la transferencia de tecnología y su marco legal. Se revisan once metodologías de evaluación con una somera descripción de cada una, sus requerimientos, aplicaciones y ejemplos de buenas prácticas, así como resalta las sinergias entre estos métodos de evaluación.

El objetivo de plantear este segundo análisis empírico se centra en estudiar el uso de los recursos de I+D+i en general desde otra perspectiva de análisis de la eficiencia: se trata de clasificar los EEMM de la UE según su grado de ejecución o resultados de la I+D+i en grupos de similares características y predecir dicha clasificación en estos grupos, lo que ayudará a monitorizar las estrategias Europeas de I+D+i y algunos aspectos claves relacionados con la política científico-tecnológica de la UE.

Al enfocar el análisis sobre el concepto de la I+D+i, se han introducido las variables que se consideran más directamente relacionadas con los recursos empleados y sus impactos más inmediatos. Además, se ha introducido el total del gasto en I+D+i pero separando ambas variables de forma que, por un lado, aparecerá el indicador del gasto privado en I+D+i y, por otro, el del gasto público. No se han diseñado escenarios desde el enfoque inputs/outputs sino que no se realiza ninguna hipótesis sobre la forma de la relación entre las variables para estudiar las características de los países según su esfuerzo en I+D+i y los impactos que son más utilizados en la literatura, para detectar las variables más relevantes una vez obtenido el mejor clasificador.

En cuanto a la metodología en concreto que se va a aplicar, se propone el uso del algoritmo de agrupamiento k -medias para detectar los patrones de comportamiento de los EEMM de la UE seleccionados en el periodo 2005-2008. Como resultado del agrupamiento k -medias, se fijará un número de clases y se describirán las características de cada una de ellas. Seguidamente, se propone la construcción de un clasificador nominal multiclase para asignar cada observación

BLOQUE 3- ANÁLISIS EMPÍRICO

país-año a su correspondiente grupo/clase de acuerdo al conjunto de rasgos específicos que caracterizan a cada país europeo.

Para la construcción del clasificador, se han desarrollado modelos de redes neuronales empleando para ello Redes Neuronales Artificiales (RNAs) de Unidad Sigmoide (SUNN) y Redes Neuronales Artificiales de Unidad Producto (PUNN) entrenadas con Algoritmos de Programación Evolutiva (APEs). Estos APEs usan operaciones de mutación estructural que permiten construir modelos de redes SUNN (*Sigmoid Unit Neural Networks*) y PUNN (*Product Unit Neural Networks*), con un número reducido de funciones de base y de conexiones, lo que determina un número más reducido de coeficientes en el modelo.

Centrándonos en la metodología propuesta en este análisis empírico y que acabamos de citar, en las dos últimas décadas las RNAs han sido empleadas con éxito en el campo de la Economía para la detección y predicción de crisis bancarias (Gutiérrez et al., 2010), para la modelización de series temporales financieras (Trippi y Turban, 1994), para determinar predicciones de bancarrota (Charalambaus et al., 2000) y para identificar la estructura de un determinado mercado (Gruca y Klemz, 1998), entre otros muchos ejemplos. Las RNAs presentan muchas ventajas sobre otros métodos convencionales. Comparadas con un modelo econométrico, se podría decir que es más sencillo usar RNAs cuando se necesita obtener una predicción en un periodo de tiempo más corto²⁰. Como la metodología DEA, la RNA es una técnica no paramétrica que adopta hipótesis menos estrictas y, desde el punto de vista de la aplicación en el análisis de la ejecución de I+D+i, es capaz tanto de detectar relaciones funcionales no lineales que se encuentren ocultas en los datos analizados como de aplicar esas relaciones a nuevos datos; finalmente, el problema puede ser descrito mediante unas cuantas variables independientes consiguiendo a veces, y dependiendo del tipo de RNA, un modelo con una menor dimensionalidad o una menor complejidad (Bode, 1998).

A pesar de ello, al día de hoy no son muchos los trabajos relacionados con la ejecución de los recursos en I+D+i que han empleado RNAs para su análisis y éstos se centran, básicamente, en la gestión de actividades de I+D+i (Becker y Prischmann, 1993). Nos encontramos que las técnicas de Data Mining y de agrupamiento han sido usadas de forma similar cuando se trataba de Sistemas Regionales de Innovación (Hájková y Hájek, 2010) y para el estudio de la competitividad de las naciones (Zanak y Becerra-Fernández, 2005). En Hájková y Hájek

²⁰ Para una revisión de las redes neuronales ver S. Haykin (1995) y R.P. Lippman (1989).

BLOQUE 3- ANÁLISIS EMPÍRICO

(2010), p.e., los autores escogieron los Sistemas Regionales de Innovación en lugar de países. De esta forma, un amplio rango de 265 regiones de la UE se agruparon en 9 grupos mediante 11 variables relacionadas con la economía, la educación y la I+D+i.

En relación con los modelos disponibles de redes neuronales, el más popular es el Perceptrón Multicapa (MLP, *Multilayer Perceptron*) debido a su arquitectura simple pero con una gran capacidad de resolución de problemas. Se trata de un modelo con funciones de transferencia sigmoideas donde la información va desde la capa de entrada a las capas ocultas (en general una sola) y de ésta a la capa de salida (Smith y Gupta, 2000). Sin embargo, han surgido alternativas en los últimos años al modelo MLP: modelos de Redes Neuronales de Unidades Producto (PUNN) que se basan en neuronas multiplicativas en lugar de aditivas. En este tipo de redes las funciones de base son productos de funciones de potencias, lo que las hace muy flexibles, siendo también del tipo de paso de información hacia delante (Durbin y Rumelhart, 1989).

En muchos casos, las redes neuronales que usan funciones de base de unidad sigmoide (SUNNs)²¹ y las que usan funciones de base de unidad producto (PUNNs) se entrenan mediante el empleo de Algoritmos de Programación Evolutiva, esto es, algoritmos evolutivos que no utilizan el cruce como operador de recombinación debido a los problemas de permutación que se presentan en la representación en grafo de los modelos de red; con este método se obtienen ventajas significativas respecto a otros enfoques tradicionales de entrenamiento basados en descenso gradiente (Becker y Prischmann, 1993). Uno de los aspectos más importantes del empleo de Algoritmos de Programación Evolutiva (APEs) como metodología de modelización, es su habilidad de hacer más flexible la complejidad de la red durante el entrenamiento a través del uso de operadores estructurales de mutación que añaden y eliminan nodos y añaden y eliminan conexiones.

Tras esta breve introducción, este capítulo presenta una sección metodológica asociada a la estructura dinámica seguida en este análisis: primero, un análisis de agrupamiento en grupos similares de los países-año considerados en este trabajo para, a continuación, determinar el método de clasificación de nuevos países año en las clases o grupos determinados en el primer paso, mediante RNAs; en la siguiente sección, se lleva a cabo el diseño experimental con la descripción de la base de datos seleccionada (basándonos en las variables y relaciones de nuestro modelo conceptual causal), la presentación de los resultados y el análisis del mejor

²¹ Llamadas MLP o SUNN en este trabajo.

modelo obtenido para la clasificación. Las conclusiones acerca de los resultados obtenidos cierran el presente análisis.

2. METODOLOGÍA

2.1. AGRUPAMIENTO K -MEDIAS Y DESCRIPCIÓN DE LAS CLASES

En muchas situaciones, los decisores tienen que agrupar el objeto de su análisis en clases homogéneas. La técnica de agrupamiento que se va a emplear se basa en el algoritmo k -medias, que es una de las técnicas de aprendizaje no supervisado más ampliamente utilizadas (es decir, a priori no se conoce la clase o grupo al que pertenece un determinado país-año). Este método permite agrupar los países-año en grupos o clases de forma que se mantengan las distancias más cortas posibles entre ellos dentro de la misma categoría, mientras se mantienen las distancias más largas entre los países-año de categorías diferentes (De Smet, 2004).

Nuestra tarea inicial es emplear la técnica de agrupamiento para detectar patrones entre los países europeos en relación a su desempeño de I+D+i. Una de las ventajas de esta técnica es que descubre patrones ocultos dentro de los datos en un análisis multidimensional.

Para la aplicación de esta técnica, el número de clases deseado debe ser especificado. Como esta información no es necesariamente obvia, se ejecutaron experimentos para un número de clases que van de la 1 a la 9. Las particiones en tres o cuatro clases aparecen como prometedoras: la suma de las distancias intra-clase dividida por la suma total de las distancias es relativamente baja mientras que el número de clases también es pequeño. El considerar más clases no conduce a un incremento significativo en la calidad de la partición (De Smet, 2004). Por otro lado, si se considera una partición en cuatro clases, en lugar de tres, permite que la calidad del agrupamiento se incremente.

El algoritmo simple de k -medias (Hartigan y Wong, 1979) fue el empleado debido a su fácil implementación y su bajo coste computacional, encontrándose disponible en el software WEKA (Hall et al., 2009). Como se ha puntualizado, dado que hubo un decremento significativo en la suma de los errores al cuadrado de los grupos o clases cuando se incrementa k , se escogió un valor de $k=4$. Seguidamente, los valores de las variables para cada país-año en cada grupo se compararon con los valores de los centroides para comprobar cuáles eran los más cercanos y los más lejanos de acuerdo a la distancia Euclídea. Usando un análisis

BLOQUE 3- ANÁLISIS EMPÍRICO

económico de estos patrones y de los centroides de los grupos, el perfil de cada grupo fue descrito y etiquetado²² en cuatro clases como sigue (para la descripción de las variables, los centroides de los grupos y sus desviaciones estándar, ver la tabla 8.1):

Grupo 1: países de desempeño moderado en innovación. Este grupo comprende los patrones (países-año) en los que la I+D+i juega un papel menor dado que la intensidad de I+D+i no está bien implementada ni en el sector público ni en el sector privado o, de media, en el número de investigadores y de recursos humanos dedicados a ciencia y tecnología. Estos países se encuentran rezagados en los resultados de la I+D+i mostrando una baja intensidad en relación a sus solicitudes de patentes y en el número de publicaciones científicas. La cantidad de inversión en I+D+i sólo puede describirse como un conductor complementario del crecimiento. Estos países no son económicamente fuertes en términos de tasa de empleo, y de crecimiento del PIB y de la productividad real del trabajo, no destacando tampoco por el número de graduados doctores ni el aprendizaje a lo largo de toda la vida, aunque presentan buenos resultados en el número de marcas comerciales comunitarias y los ingresos por las licencias y las patentes.

Grupo 2: países orientados a la innovación. Esta clase incluye los países-año más desarrollados. Son económicamente prósperos (con las mejores tasas de empleo, aunque el crecimiento del PIB y de la productividad real del trabajo son menores debido a su ya favorable situación de partida) y presentan la mayor proporción de inversión en I+D+i, de personal investigador, de recursos humanos en ciencia y tecnología y en el número de graduados doctores. También destaca la educación de nivel terciario y el aprendizaje a lo largo de toda la vida, variables muy representativas de sus fuertes sistemas educativos. Este hecho les proporciona la fuerza de trabajo con las habilidades necesarias en un ambiente laboral cambiante y sienta las bases para altos niveles de adopción tecnológica e innovación. Finalmente, también presentan los mejores resultados en cuanto al número de publicaciones científicas y en los ingresos procedentes de las patentes.

Grupo 3: países de baja innovación. Estos países-año representan principalmente las economías de transición de los Nuevos Estados Miembros (excepto Grecia y Portugal). Como

²² De forma similar, y en base a su ejecución media en innovación, el Cuadro de Indicadores de la Unión (IUS) clasifica a todos los EEMM de la UE y algunos países más en grupos que denomina: líderes de innovación, seguidores, de innovación moderada e innovadores modestos. En el caso del grupo 2 (países dirigidos por la innovación), por ejemplo, se debe decir que es un término comúnmente usado y que se puede encontrar en numerosos estudios e informes relacionados con la competitividad de los países.

BLOQUE 3- ANÁLISIS EMPÍRICO

nuevas economías, presentan una alta tasa de crecimiento del PIB y de la productividad real del trabajo debido a la abundancia de recursos naturales, la escasa población y la gran demanda interna después de un periodo de socialismo, entre otras razones. Pero, también se caracterizan por una baja tasa de desempleo. Sus economías no parecen basarse ni en la educación ni en la inversión en I+D+i y esto podría ser un obstáculo para su futuro desarrollo socioeconómico.

Grupo 4: países de gran innovación o “seguidores”. Los países en este grupo se ajustan bien al modelo de crecimiento basado en el cambio tecnológico, presentando altos valores medios en casi todas las variables. Se encuentran muy próximos al grupo 2 de los países orientados a la innovación en el número de graduados doctores, en la tasa de empleo, en el número de patentes y en las exportaciones de media y alta tecnología (estos dos últimos casos, en la posición más alta). Sin embargo, presentan valores medios bajos en términos de beneficios procedentes de sus marcas comerciales y de los ingresos de licencias y patentes así como en el número de personas matriculadas en el nivel terciario de educación. El nivel de aprendizaje a lo largo de toda la vida no parece tan bueno como una podría suponer de este grupo de países. De esta forma, sus resultados económicos de la innovación y los educativos, a pesar las variables arriba mencionadas son peores que en el grupo 2 y por ello se han considerado en un grupo diferenciado.

Sólo dos observaciones (países-año) varían de grupo a lo largo del periodo de tiempo seleccionado: Malta y Estonia. Ambos cambian de un país de baja innovación a innovación moderada (del grupo 3 al 1), principalmente por el incremento de las solicitudes de patentes de residentes tramitadas a través del Tratado de Cooperación en Materia de Patentes o a través de la oficina de patentes nacional (*TPATRE*) y los ingresos procedentes de licencias y patentes del exterior (*PATREV*). Pero mientras que el primero cambia en el año 2006, el segundo lo hace en el 2008.

2.2. PROBLEMAS DE CLASIFICACIÓN Y REGRESIÓN

Uno de los temas fundamentales tratados en el campo de la estadística consiste en cuantificar la relación entre una variable dependiente que resulta de interés y una serie de variables independientes que se encuentran relacionadas con dicha variable dependiente. Este tipo de problema es considerado habitualmente como dos tareas distintas: *regresión* (cuando se trata de

Tabla 8.1. Descripción de las variables, centroides (media) de los grupos \pm desviaciones típicas

CÓDIGO	VARIABLE ^(e)	UNIDAD	#PATRONES/ TOTAL	GRUPO				
				GRUPO 1 ^(a)	GRUPO 2 ^(b)	GRUPO 3 ^(c)	GRUPO 4 ^(d)	
Recursos de la I+D+i				100	24	16	40	20
<i>HUMST</i>	Recursos Humanos totales relacionados con la Ciencia y la Tecnología	% de la fuerza de trabajo	31,39 \pm 6,08	30,09 \pm 6,03	40,52 \pm 6,08	27,70 \pm 6,08	33,04 \pm 6,11	
<i>RDPERS</i>	Personal dedicado a la I+D+i.	% de la fuerza de trabajo	0,99 \pm 0,48	1,00 \pm 0,48	1,59 \pm 0,49	0,66 \pm 0,48	1,17 \pm 0,47	
<i>ERDPRI</i>	Gasto interno bruto en I+D: sector empresas	% del PIB	0,94 \pm 0,72	0,78 \pm 0,69	2,00 \pm 0,73	0,40 \pm 0,72	1,38 \pm 0,71	
<i>GERDGO</i>	Gasto interno bruto en I+D: gobierno y educación superior	% del PIB	0,56 \pm 0,21	0,41 \pm 0,21	0,89 \pm 0,21	0,45 \pm 0,21	0,70 \pm 0,21	
Resultados de la I+D+i								
<i>SPUBLI</i>	Número de publicaciones científicas	Por cada 1000 habitantes	1,23 \pm 0,65	1,09 \pm 0,63	2,29 \pm 0,65	0,74 \pm 0,65	1,53 \pm 0,64	
<i>TPATRE</i>	Solicitudes de patentes cumplimentadas a través del Tratado de Cooperación de Patentes o una oficina nacional (residentes)	Por cada 1000 habitantes	0,14 \pm 0,14	0,07 \pm 0,14	0,26 \pm 0,14	0,06 \pm 0,14	0,31 \pm 0,14	
Educación								
<i>PHD0%</i>	Número de graduados Doctores	Por cada 1000 habitantes	0,17 \pm 0,09	0,10 \pm 0,09	0,25 \pm 0,09	0,14 \pm 0,09	0,24 \pm 0,09	
<i>TERITT</i>	Población con nivel de educación terciaria	% de la población entre 30-34 años	31,09 \pm 10,29	38,20 \pm 10,39	41,58 \pm 10,40	22,79 \pm 10,29	30,76 \pm 10,56	
<i>LLEARN</i>	Aprendizaje a lo largo de la vida	% de la población entre 18-64 años	15,90 \pm 6,71	12,76 \pm 6,54	27,26 \pm 6,80	11,71 \pm 6,71	18,94 \pm 6,92	

Tabla 8.1. Descripción de las variables, centroides (media) de los grupos \pm desviaciones típicas (cont.)

CÓDIGO	VARIABLE ^(e)	UNIDAD	#PATRONES/ TOTAL	GRUPO 1 ^(a)	GRUPO 2 ^(b)	GRUPO 3 ^(c)	GRUPO 4 ^(d)
				24	16	40	20
Economía							
<i>GDPGRO</i>	Tasa de crecimiento del PIB	% de cambio sobre el año anterior	3,01 \pm 3,17	1,63 \pm 2,94	2,17 \pm 2,99	4,50 \pm 3,17	2,35 \pm 3,04
<i>RLPGH</i>	Crecimiento de la productividad real del trabajo por hora trabajada	Output real por unidad de trabajo (medido por el total de horas trabajadas).	1,89 \pm 2,36	0,80 \pm 2,21	1,08 \pm 2,29	2,84 \pm 2,36	1,93 \pm 2,30
<i>EMPLO</i>	Tasa total de empleo	% de personas entre 20-64 años con empleo	70,74 \pm 5,59	69,45 \pm 5,43	77,47 \pm 5,63	67,90 \pm 5,59	72,59 \pm 5,43
<i>PATREV</i>	Ingresos de licencias y patentes del exterior	% del PIB	0,39 \pm 0,51	0,55 \pm 0,53	1,00 \pm 0,52	0,12 \pm 0,51	0,27 \pm 0,52
<i>TRADEM</i>	Marcas comerciales comunitarias	Por cada billón del PIB (en PPS€)	4,59 \pm 3,22	7,65 \pm 3,34	5,38 \pm 3,28	2,39 \pm 3,22	4,69 \pm 3,24
<i>MHTEXP</i>	Exportaciones de media y alta tecnología	% del total de las exportaciones manufactureras	49,16 \pm 12,00	48,65 \pm 12,08	47,80 \pm 11,94	45,98 \pm 12,00	57,21 \pm 11,92

a. Bélgica05-08, Chipre05-08, Irlanda05-08, Luxemburgo05-08, Malta06-08, España05-08, Estonia08.

b. Finlandia05-08, Suecia05-08, Dinamarca05-08, Países Bajos05-08.

c. República Checa 05-08, Estonia05-07 Grecia05-08, Hungría05-08, Italia05-08, Letonia05-08, Lituania05-08, Polonia05-08, Portugal05-08, Eslovaquia05-08.

d. Austria05-08, Francia05-08, Eslovenia05-08, Reino Unido05-08, Alemania05-08.

e. Fuente: Eurostat para todas las variables excepto *SPUBLI* (SCImago JCR), *TPATRE* (Banco Mundial), *TRADEM* y *MHTEXP* (Innometrics/Eurostat).

BLOQUE 3- ANÁLISIS EMPÍRICO

una variable dependiente continua) y *clasificación* (cuando la variable dependiente es categórica, nominal u ordinal). El resultado del proceso es la obtención de modelos, es decir, de abstracciones de la realidad que pueden ser aplicadas tanto para predecir valores de la variable dependiente dados nuevos valores de las variables independientes como para mejorar nuestra comprensión de la misma.

En ambos tipos de problemas, el objetivo se centra en determinar una relación funcional entre las variables predictoras y la variable (o variables) de respuesta. Tradicionalmente, estos problemas se han resuelto empleando técnicas de optimización que minimicen una determinada función de error, previo establecimiento del tipo de modelo a aplicar. En general, no será posible determinar la relación funcional si no es mediante el uso de un conjunto de datos de ejemplo. De esta forma, la relación se modela en términos de alguna función que contiene parámetros ajustables, cuyos valores se determinan precisamente con la ayuda de los datos.

Una forma de aproximación es considerar funciones lineales en la relación causa-efecto entre el vector de entradas \mathbf{x} y la variable dependiente y pero, a pesar de su amplio uso, los modelos lineales son habitualmente demasiado restrictivos para capturar de forma precisa la relación que subyace en el problema. Esta necesidad de modelos más avanzados fue reconocida hace años por varios investigadores pioneros (por ejemplo, Whittaker, 1922) sugiriendo modelos más complejos (modelos no lineales).

Existen dos cuestiones fundamentales a las que dar respuesta:

1. ¿Qué tipo de funciones de aproximación utilizar o cuáles se tienen disponibles?
2. ¿Cómo saber que se ha llegado a la mejor aproximación posible?

Respecto a la primera pregunta, actualmente existen una enorme cantidad de métodos y modelos y ninguno se puede considerar mejor que el resto para todo tipo de problema. Entre los modelos más utilizados se encuentran los modelos lineales y los modelos lineales generalizados, Redes Neuronales Artificiales (RNAs), bases de Fourier, árboles de decisión, modelos de tipo ventana o “kernel”, etc.

Dentro del campo de la clasificación, Lippmann (1996) propone una taxonomía de clasificadores considerando cuatro grupos diferenciados:

- Las funciones de densidad de probabilidad.
- Los clasificadores globales basados en probabilidad a posteriori.

BLOQUE 3- ANÁLISIS EMPÍRICO

- Los clasificadores locales basados en probabilidad a posteriori.
- Los métodos basados en el vecino más cercano.

Respecto a la segunda pregunta, un modelo es sólo una aproximación a la relación causa-efecto real. El modelo que se considerara el “mejor”, entre un conjunto de modelos alternativos, sería el que mejor capturase esa relación.

En general, existen dos fases claramente diferenciadas en la obtención de los modelos: entrenamiento y generalización. La adquisición del modelo y de los posibles valores de los parámetros asociados al mismo se realiza durante la fase de entrenamiento o fase de aprendizaje en base a los datos de entrenamiento. Una vez que el modelo ha sido entrenado puede predecir valores asociados a nuevos datos de la misma naturaleza de los que se han utilizado en el entrenamiento. Estos datos forman parte del conjunto de test o de generalización y a la habilidad de determinar los valores de la variable objetivo en dichos datos se denomina capacidad de generalización. De este modo, uno de los principales objetivos es encontrar un modelo que sea capaz de generalizar de la mejor forma posible.

Es una cuestión generalmente aceptada que medidas de predicción robustas deben usar datos independientes. Existen varias estrategias para obtener los datos de prueba y los datos de generalización. Una práctica común es dividir los datos disponibles en sendos conjuntos. Algunas de las técnicas que existen son resustitución, aleatorización, muestreo prospectivo, *leave-one-out*, etc., siendo las más habituales las llamadas *holdout*, consistente en usar una porción de patrones de la muestra como conjunto de entrenamiento (en general, un 75%) para construir el modelo y usando el resto como conjunto de prueba y la validación cruzada o *cross validation*, en la que los datos se dividen en k subconjuntos mutuamente excluyentes, del mismo tamaño aproximadamente, se generaliza cada vez en uno de los k subconjuntos y se construye el modelo con los subconjuntos restantes.

2.2.1. Redes Neuronales

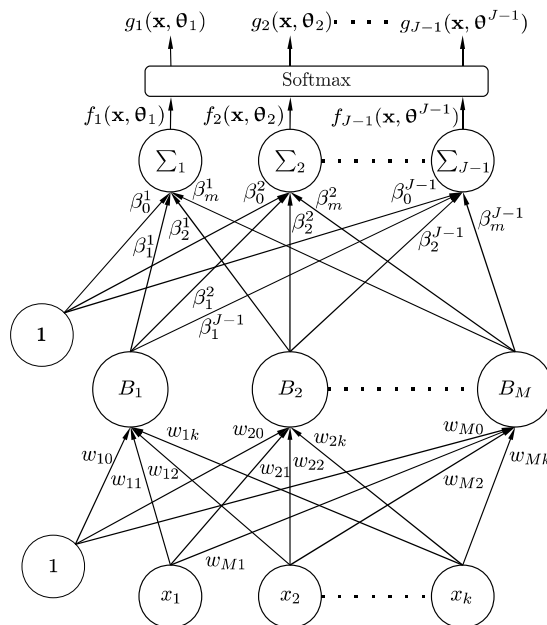
Las Redes Neuronales Artificiales o RNAs (Bishop, 1995) son una técnica de modelado enormemente flexible fundamentada en imitar los sistemas nerviosos biológicos y cuya capacidad de cómputo se desarrolla mediante un proceso adaptativo de aprendizaje. La diferencia clave entre las redes neuronales y otros métodos de resolución de problemas es que las redes neuronales aprenden mediante ejemplos, utilizando procesamiento en paralelo de la

información, más que siendo programadas con unas reglas específicas y predeterminadas; pueden modificarse a sí mismas para ajustarse a los datos presentados y, más aún, también son capaces de aprender de muestras muy ruidosas, desordenadas o incompletas.

La estructura de una RNA consiste en una capa de entrada, una o varias capas ocultas, y una capa de salida. La capa de entrada toma la información (p.e. de variables independientes) y ésta se transmite hacia las neuronas inferiores vía “conexiones”. La unidad base de cualquier red neuronal es la neurona (procesador). Cada unidad o neurona tiene un número de entradas que se combinan para producir una única salida. Cada neurona de la capa oculta o de la capa de salida es capaz de agregar los valores que proporcionan las neuronas de la capa de entrada, asociados a las variables independientes del problema, tanto si esas entradas provienen de una base de datos como de otras neuronas, modificando cada input mediante un peso ajustable w_j . La suma de estas entradas ponderadas (β) se añade a un umbral ajustable para la neurona y luego pasan a través de una función que los modifica (función de transferencia, f) que determina la salida o resultado final y .

La capa oculta es calculada en el proceso de entrenamiento de la red. No existe un límite teórico en cuanto al número de este tipo de capas pero típicamente será suficiente con utilizar una capa. Para los problemas de clasificación, el utilizado se representa en la figura 8.1.

Figura 8.1. Modelo de red utilizado en problemas de multclasificación



Desde un punto de vista formal, las RNAs de transmisión de la información hacia delante (*feedforward*) y una sola capa oculta no son más que una forma de un modelo de regresión lineal generalizado y formada por una combinación lineal de transformaciones no lineales $B_j(\mathbf{x}, \mathbf{w}_j)$ de las variables de entrada, llamadas funciones de base, con la siguiente expresión:

$$y(\mathbf{x}, \boldsymbol{\theta}) = \beta_0 + \sum_{j=1}^M \beta_j B_j(\mathbf{x}, \mathbf{w}_j) \quad (8.1)$$

donde $\boldsymbol{\theta}$ y \mathbf{w}_j recogen los parámetros de la red. Este tipo de modelos es conocido como modelo lineal de funciones de base. Existe la posibilidad de elección de las funciones de base entre, por ejemplo, las *Funciones de Base Radial* (*Radial Basis Functions*, RBFs), que dan lugar a las Redes Neuronales de Función de Base Radial (RNRBFs), las *funciones de base de Unidad Sigmoide* (US) que dan lugar al conocido como Perceptrón Multicapa (MLP) o las *Unidades Producto* (UP) que dan lugar a las Redes Neuronales de UPs (RNUPs) (Durbin y Rumelhart, 1989), donde un nodo multiplicativo es dado por

$$\beta_j(x, w_j) = \prod_{i=1}^k x_i^{w_{ji}} \quad (8.2)$$

Por otro lado, se pueden considerar dos tipos fundamentales de funciones de base:

- **Las funciones locales.** Presentan un valor elevado distinto de cero sólo sobre una región localizada del espacio de entrada (p.e. las redes RBFs).
- **Las funciones de proyección o globales.** Funciones de entorno global (es decir, presentan un valor distinto de cero sobre una región amplia del espacio de entrada) como las Unidades Sigmoide o las Unidades Producto.

El aprendizaje de las RNAs consiste en estimar un valor para el conjunto de parámetros $\boldsymbol{\theta}$ y una arquitectura para la red (es decir, el número de transformaciones no lineales M y de conexiones entre los distintos nodos de la red). Suponiendo una arquitectura fija, el aprendizaje de los parámetros ha sido tradicionalmente llevado a cabo mediante algoritmos de optimización basados en el gradiente descendente de la función de error, tales como el Algoritmo de Retropropagación del error (*Back Propagation*, BP). Una alternativa a estos algoritmos, como se verá, son los algoritmos evolutivos.

En una RNA, cada función de base tiene como salida una función de transferencia cuya entrada es el valor obtenido al aplicar una función de activación sobre las entradas de la red. En general, se considerarán dos tipos de funciones de activación de neuronas o nodos, las aditivas (de unidad sigmoide) y las multiplicativas (de unidades producto). Estos dos tipos de funciones dan lugar a dos modelos de RNAs que se emplearán en este trabajo: el modelo aditivo y el modelo multiplicativo.

Las funciones típicas de activación en las redes neuronales son la Logística, la Logística Simétrica, la Sigmoide, la Tangente Hiperbólica, la Gausiana y la Sinusoidal.

Aunque existe un modelo funcional para el caso de la regresión y otro para el caso de la clasificación, sólo se verá el segundo dado que los análisis planteados son de clasificación. Cuando el problema es de este tipo, se considera un número de nodos de salida igual al número de clases o categorías de la variable a predecir menos uno. De este modo, si Q es el número de clases, $\theta = \{\theta_1, \dots, \theta_{Q-1}\}$ es el conjunto de coeficientes asociados a la RNA y $\theta_j = \{B_0^j, \dots, B_M^j, \mathbf{w}_1, \dots, \mathbf{w}_{Q-1}\}$ es el conjunto de parámetros asociados al nodo de salida j , el modelo funcional sería el siguiente:

$$f_j(\mathbf{x}, \theta_j) = \beta_0^j + \sum_{i=1}^M \beta_i^j B_i(x, w_i), \quad 1 \leq j \leq Q-1 \quad (8.3)$$

Para resolver el problema de seleccionar el tamaño apropiado de la capa oculta de la red, es decir, el número de nodos M en la capa oculta, se ha introducido una nueva categoría de algoritmos que determinan automáticamente la estructura de la red. Esta categoría de algoritmos incluye el algoritmo ortogonal de mínimos-cuadrados; el entrenamiento individual de cada unidad oculta basada en el análisis funcional; métodos constructivos, donde la estructura de la red se construye incrementalmente; métodos *pruning* que comienzan con una selección inicial de un gran número nodos en la capa oculta que se va reduciendo a medida que trabaja el algoritmo; y la selección simultánea de la estructura de la red y de los parámetros mediante el empleo de métodos de optimización basados en la programación de algoritmos evolutivos. Este último será el empleado para el presente análisis (ver sección 2.2.3).

2.2.2. Clasificación multiclase y métricas de evaluación

El reconocimiento de patrones multiclase es un problema de construcción de un sistema que

mapee exactamente un espacio de características de entrada a un espacio de salida de más de dos clases de patrones. Mientras que el problema de clasificación en dos clases se comprende bien, existe una investigación relativamente escasa acerca de la clasificación multiclase. Así, muchos sistemas de clasificación de patrones han sido desarrollados para los problemas de clasificación en dos clases y los estudios teóricos de aprendizaje se han centrado casi completamente en el aprendizaje de funciones binarias incluyendo las Máquinas de Vectores Soporte (SVM), las RNAs y los correspondientes algoritmos de retropropagación del error (BP).

En general, un sistema de red neural F puede ser un sistema de redes neuronales o de una sola red neuronal cuyos pesos son determinados mediante un algoritmo de aprendizaje (un algoritmo evolutivo en nuestro caso).

Sea una muestra de entrenamiento $D = \{(\mathbf{x}_n, y_n); n = 1, 2, \dots, N\}$ donde $\mathbf{x}_n = (x_{1n}, \dots, x_{kn})$ es el vector de medidas que toman valores en $\Omega = \mathbb{R}^k$ e y_n es la etiqueta de clase del n -ésimo individuo. Se adopta la técnica común de representar las etiquetas de clase usando un vector de codificación “1-de- Q ” $y = (y^{(1)}, y^{(2)}, \dots, y^{(Q)})$ tal que $y^{(l)} = 1$ si \mathbf{x} corresponde a un ejemplo que pertenezca a la clase l e $y^{(l)} = 0$ en caso contrario. Basándonos en la muestra de entrenamiento se desea encontrar un clasificador $RNA(\hat{\theta})$ que llegue a una buena clasificación de los patrones en el conjunto de generalización. Esto sería igual a obtener una función de decisión $C: (\Omega, \theta) \rightarrow \{1, 2, \dots, Q\}$ para clasificar a los individuos (una red neuronal en nuestro caso). Una clasificación errónea ocurre cuando C asigna a un individuo (basado en un vector de medidas) a la clase j cuando realmente proviene de una clase $l \neq j$.

Para un clasificador g se considera un problema de clasificación de Q clases y N patrones de entrenamiento o generalización, obteniendo una matriz de contingencia o matriz de confusión de dimensiones $Q \times Q$ y que se representa con $M(g)$:

$$M(g) = \left\{ n_{ij}; \sum_{i,j=1}^Q n_{ij} = N \right\} \quad (8.4)$$

donde n_{ij} representa el número de veces que los patrones han sido clasificados por g en la clase j cuando realmente pertenecen a la clase i . Se representa el número de patrones asociados con

la clase i como $f_i = \sum_{j=1}^Q n_{ij}, i=1, \dots, Q$.

Entonces, para evaluar la precisión de los resultados de los clasificadores que se obtengan, se definen tres medidas escalares que utilizan los elementos de la matriz de confusión (tabla 8.2) desde diferentes puntos de vista.

Tabla 8.2. Matriz de confusión de un clasificador

CLASE	1	2	...	Q	
1	n_{11}	n_{12}	...	n_{1Q}	f_1
2	n_{21}	n_{22}	...	n_{2Q}	f_2
...
Q	n_{Q1}	n_{Q2}	...	n_{QQ}	f_Q

La primera medida sería la llamada sensibilidad. Sea $S_i = n_{ii}/f_i$ el ratio de patrones correctamente clasificados en la clase i respecto al número total de patrones en la clase i . De esta forma, la sensibilidad de la clase i estima la probabilidad predecir correctamente la clase i . La primera medida de comparación de un clasificador multiclase define la sensibilidad S del clasificador como el mínimo valor de sensibilidades para cada clase: $S = \min \{S_i; i=1, \dots, Q\}$ o MS .

La segunda medida es la tasa de clasificación correcta (*Correct Classification Rate, CCR*) o precisión,

$$CCR = \left(\frac{1}{N} \right) \sum_{i=1}^Q n_{ii} \tag{8.5}$$

donde N es el número total de patrones, n_{ii} el número de patrones de la clase i bien clasificados y Q representa el número total de clases.

La tercera es el coeficiente Kappa de Cohen (K) (Cohen, 1960), una medida de asociación entre la clase a la que *a priori* pertenece un patrón y la que *a posteriori* asigna el clasificador. Así, se considerarán las tres medidas de evaluación (CCR, MS, K) asociadas al clasificador g .

Las medidas tratan de evaluar tres características de un clasificador a través del conjunto de generalización: la ejecución en cada clase, los resultados globales en el conjunto de datos completo y el grado de asociación entre las distribuciones de los patrones en las clases antes y después de la aplicación del clasificador.

En cuanto al modelo de red, se considera la RNA estándar sigmoide, también conocida como Perceptrón Multicapa (MLP), como la base del modelo de clasificación (SUNN). Para determinar el modelo de RNA, se aplica un algoritmo evolutivo (ver sección 2.2.3) para encontrar las funciones de base

$$B(\mathbf{x}, \mathbf{W}) = \{B_1(\mathbf{x}, \mathbf{w}_1), B_2(\mathbf{x}, \mathbf{w}_2), \dots, B_M(\mathbf{x}, \mathbf{w}_M)\} \quad (8.6)$$

correspondientes a la parte no lineal de las funciones discriminantes, $f_i(\mathbf{x}, \boldsymbol{\theta}_i)$. Hay que determinar el número de funciones de base m y la matriz de pesos \mathbf{W} . Para la aplicación de las técnicas de redes neuronales evolutivas, se considera una RNA con outputs *softmax* definidos en la ecuación (8.7) y una estructura estándar: una capa de entrada con un nodo por cada variable de entrada; una capa oculta con varios nodos sigmoides; y una capa de salida con $Q-1$ nodos, donde j es el número de clases.

El enfoque adoptado a la hora de interpretar las salidas de los nodos de la capa de salida del clasificador $RNA(\hat{\boldsymbol{\theta}})$ es probabilístico, considerando la función de activación *softmax* en cada uno de dichos nodos, la cual viene dada por:

$$g_l(x, \boldsymbol{\theta}_l) = \frac{\exp f_l(x, \boldsymbol{\theta}_l)}{\sum_{l=1}^Q \exp f_l(x, \boldsymbol{\theta}_l)} \quad \text{para } l = 1, \dots, Q \quad (8.7)$$

donde $g_l(\mathbf{x}, \boldsymbol{\theta}_l)$ es la probabilidad que un patrón \mathbf{x} tiene de pertenecer a la clase l ; $\boldsymbol{\theta}_l = (\boldsymbol{\beta}_l, \mathbf{w}_1, \dots, \mathbf{w}_M)$, $\boldsymbol{\beta}_l = \beta_0^l, \dots, \beta_M^l$ es el vector l -ésimo de pesos del nodo de salida; M es el número de nodos ocultos; $\mathbf{w}_j = w_{j0}, \dots, w_{jk}$ para $j=1, \dots, M$, es el vector de pesos de entrada del nodo oculto j ; y $f_l = (\mathbf{x}, \boldsymbol{\theta}_l)$ es la salida del nodo de salida l para los patrones \mathbf{x} dado por:

$$f(\mathbf{x}, \boldsymbol{\theta}_l) = \beta_0^l + \sum_{j=1}^M \beta_j^l \sigma_j \left(w_{j0} + \sum_{i=1}^k w_{jk} x_i \right) \quad \text{para } l = 1, \dots, Q-1$$

$$f_Q(x, \boldsymbol{\theta}_Q) = 0 \quad (8.8)$$

donde $\sigma(\cdot)$ es la función de activación sigmoide de los nodos en la capa oculta. No obstante, las funciones de base PU de la ecuación (8.2) también se considerarán para dar lugar a modelos PUNN.

La transformación *softmax* produce estimaciones positivas en todas las salidas, siendo la suma total de 1, lo que permite que puedan ser interpretadas como la probabilidad de pertenencia a la clase correspondiente.

Al tratarse de probabilidades de pertenencia a una clase, no es necesario calcularlas todas ya que la probabilidad de la última salida se calcularía como la diferencia entre 1 y la suma del valor del resto de probabilidades de pertenencia. Con ello, no habría que entrenar ese nodo y se reduciría el número de coeficientes a estimar y la carga computacional del aprendizaje. Esto se consigue igualando la última salida a “0” (ver ecuación (8.8)).

La regla de clasificación coincide con la regla de optimización de Bayes. Así, la regla de clasificación asigna un individuo a la clase que tiene la máxima probabilidad, dado el vector de medidas \mathbf{x} :

$$C(x) = \hat{l} \quad \text{donde } \hat{l} = \arg \max_l g_l(\mathbf{x}, \hat{\theta}_l), \quad \text{para } l = 1, \dots, Q \quad (8.9)$$

2.2.3. Algoritmos de aprendizaje

El problema del aprendizaje en las RNAs se formula con la minimización de una superficie de error. La superficie de error es una función de los parámetros de la red $\theta = (\beta_j, w_{ij})$ para $j=0,1,\dots, k$ y para $i=1,\dots, M$ y, por tanto, con dimensión $m(k+1)$. Para una RNA basada en funciones de base, la función de error que se define es el error cuadrático medio (MSE), función continua y diferenciable, por lo que es posible aplicar los algoritmos basados en el gradiente de la función. Existe una gran variedad de algoritmos que utilizan el gradiente de la función objetivo, pero este tipo de algoritmos basados en gradiente suelen quedar atrapados a veces en óptimos locales cuando las funciones a optimizar son complejas, como es nuestro caso. Es por ello por lo que en este trabajo se utilizarán algoritmos Heurísticos (Computación Evolutiva) para optimizar tanto la topología como los coeficientes del modelo de red.

La Computación Evolutiva (CE) es una rama de la ciencia que interpreta la naturaleza como una inmensa máquina de resolver problemas (Holland, 1975). Los Algoritmos Evolutivos

BLOQUE 3- ANÁLISIS EMPÍRICO

(*Evolutionary Algorithms*, EAs) son uno de los paradigmas de la CE y son algoritmos de búsqueda basados en el concepto de la selección natural.

Los EAs son una clase de algoritmos de búsqueda estocástica y de heurísticas de optimización derivados de la teoría clásica de la evolución que mimetizan la metáfora natural de la evolución biológica operando sobre una población de soluciones potenciales y aplicando el principio de supervivencia de los mejores o los más capacitados. Estas dinámicas de la población siguen la regla básica de la teoría de la evolución de Darwin: la idea es que si sólo estos individuos que cumplen un determinado criterio de selección se reproducen y el resto de los individuos en la población mueren, la población converge hacia aquellos individuos que mejor cumplen los criterios de selección, como una adaptación natural (Goldberg, 2000).

Existen tres ramas principales de EAs y son los Algoritmos Genéticos de John Holland, uno de los pioneros de la computación evolutiva (Holland, 1975), las Estrategias de Evolución (Rechenberg, 1973) y, la rama que se sigue en este caso, la Programación Evolutiva (Fogel, 1966).

En resumen, las motivaciones principales que sugieren el uso de este tipo de técnicas vienen determinadas en gran medida por las limitaciones presentadas por los algoritmos clásicos de aprendizaje, como son:

- . Imposibilidad de calcular el gradiente cuando la función de activación de los nodos no es derivable.
- . Ausencia de convergencia de los algoritmos clásicos de entrenamiento, cuando el número de bits utilizados para representar la función de activación o los pesos de la red (precisión) no es suficientemente grande.
- . Tendencia, por parte de los algoritmos de entrenamiento, a obtener excesivas soluciones no óptimas en cada ejecución.
- . Si se tiene en cuenta la dificultad que implica el establecimiento de una arquitectura adecuada junto con el aprendizaje de los correspondientes pesos de la red neuronal, la aplicación de EAs está ampliamente justificada.

Existen tres formas de abordar el entrenamiento de una RNA con un EA:

- Evolucionar los pesos de la RNA.
- Evolucionar la arquitectura de la RNA.

BLOQUE 3- ANÁLISIS EMPÍRICO

- Evolucionar tanto los pesos como la arquitectura de la RNA simultáneamente.

La eficiencia de un EA dependerá de la elección de parámetros como el tamaño de la población, el esquema de selección, el número de generaciones o las probabilidades de cruce y de mutación. Por otro lado, es importante que todo EA sea capaz de alcanzar un equilibrio entre dos factores:

- **Explotación o Convergencia:** capacidad del algoritmo de concentrar los esfuerzos de la búsqueda en las áreas más prometedoras, aumentando la precisión de las soluciones obtenidas.
- **Exploración o Diversidad:** capacidad del algoritmo de evitar estancarse en un área del espacio de búsqueda que no contiene al óptimo global.

El problema es que ambos son objetivos contrapuestos: el aumento de la presión selectiva da lugar a una rápida pérdida de diversidad de la población, mientras que el hecho de mantener la diversidad puede llevar a una disminución de la precisión.

Un EA simple utiliza una población de tamaño constante y guía la evolución de dicha población durante un número de generaciones, aplicándole operadores que modifican a los individuos en cada generación. El tamaño de la población es uno de los parámetros principales que afectan a la robustez y a la eficiencia computacional del algoritmo: si son muy pequeños existe el riesgo de provocar una convergencia prematura del algoritmo, mientras que si el tamaño es demasiado grande, el coste computacional se dispara.

Para el cálculo óptimo, algunos trabajos se basan en utilizar un tamaño de población variable y, por otro lado, se han propuesto muchos métodos en la literatura para intentar incrementar la diversidad de la población y evitar una convergencia prematura.

Los EAs son aplicables a un amplio abanico de problemas de aprendizaje y optimización. Han sido aplicados a numerosos problemas en la optimización combinatoria, optimización de funciones, aprendizaje de redes neuronales artificiales, aprendizaje de sistemas de lógica difusa, etc.

Esta metodología de entrenamiento será aplicada en este análisis sobre los modelos de Redes Neuronales Evolutivas de Unidad Sigmoide (ESUNN) y de Redes Neuronales Evolutivas de Unidad Producto (EPUNN) con el objeto de comparar sus resultados con una batería de clasificadores conocidos en el campo del Aprendizaje Automático.

BLOQUE 3- ANÁLISIS EMPÍRICO

El marco básico del EA empleado en este trabajo es el siguiente: la búsqueda comienza con una población inicial de RNAs y la población es actualizada en cada iteración usando un algoritmo de actualización de la población, que incluye tanto a la estructura como a los pesos. La población está sujeta a operaciones de réplica y mutación. No se emplea el cruce debido a sus desventajas en la optimización de RNAs (Yao y Liu, 1997), en concreto, el ya mencionado problema de la permutación. El algoritmo desarrolla las arquitecturas y los pesos de las conexiones simultáneamente para cada individuo, que es una ESUNN o EPPUN. Las RNAs se representan usando un enfoque orientado a objetos y el algoritmo trata directamente con el fenotipo de la red. Cada conexión se especifica mediante un valor binario indicando si la conexión existe, y un valor real representando su peso. Como no se considera el cruce, esta representación orientada a objetos no asume un orden fijado entre los diferentes nodos ocultos.

La estructura básica del algoritmo es la siguiente:

1. Se genera una población inicial de tamaño Np .
 - Calcular la aptitud para cada individuo.
 - Ordenar de mayor a menor según la aptitud.
2. Se repite hasta un máximo número de generaciones.
 - Hacer una copia del mejor individuo.
 - Aplicar la mutación paramétrica al $r\%$ de los mejores individuos.
 - Aplicar la mutación estructural al restante $(100-r)\%$ de individuos.
 - Calcular el *fitness* de cada individuo en la población.
 - Añadir el mejor individuo ajustado de la última generación (algoritmo elitista).
 - Ordenar los individuos con respecto a su *fitness*.
 - El mejor 10% de los individuos se replica y se sustituye por el peor 10% de los individuos.
3. Seleccionar el mejor individuo de la población en la última generación y convertirlo en la solución final.

BLOQUE 3- ANÁLISIS EMPÍRICO

Los pesos se inicializan de la misma forma, con independencia del nodo oculto (nodos producto, PU, o nodos sigmoides, SU) y se asignan usando una distribución uniforme definida por dos intervalos: $[-5,5]$ para conexiones entre la capa de entrada y la capa oculta y $[-10,10]$ para las conexiones entre la capa oculta y la capa de salida.

Se realiza una mutación paramétrica para cada peso de la RNA con ruido Gausiano $w_{ij}(t+1) = w_{ij}(t) + \xi(t)$, donde $\xi(t)$ representa una variable unidimensional normalmente distribuida. La varianza de la distribución normal se actualiza aplicando la regla heurística simple de 1/5 de éxito de Rechenberg (1973). Por otro lado, la mutación estructural implica una modificación en la estructura de la RNA y permite exploraciones de diversas regiones en el espacio de búsqueda mientras ayuda a mantener la diversidad de la población.

Existen cinco mutaciones estructurales diferentes: eliminación de nodos, eliminación de conexiones, añadir nodos, añadir conexiones y fusión de nodos. Estas cinco mutaciones se aplican secuencialmente a cada red. Para más detalles sobre la estructura general de un EA y de las mutaciones paramétricas y estructurales, consultar Hervás-Martínez y Martínez Estudillo (2007) y Martínez-Estudillo et al. (2006).

Los parámetros usados en el algoritmo evolutivo fueron $N_p=1000$, el número de generaciones fue de 350, mientras que el número máximo de nodos ocultos fue de 3. El número de nodos que podían ser añadidos o eliminados en la mutación estructural se encontraba en el intervalo $[1,2]$. El número de conexiones que podían ser añadidas o eliminadas en una mutación estructural estuvo en el intervalo $[1,4]$. La mutación paramétrica se aplicó al mejor ($r_m=10$)% de los individuos. El criterio de parada se alcanzó cuando se cumplió la siguiente condición: durante 20 generaciones no se producía una mejora en los resultados medios del mejor (r_m)% de la población o en el mejor individuo.

El algoritmo evolutivo utiliza como función de aptitud $A(g) = \frac{1}{1 + E(g, \theta_1)}$, donde $E(g, \theta_1)$ es

la función de error de entropía cruzada dada por:

$$E(g, \theta_1) = \frac{1}{N} \sum_{n=1}^N \sum_{l=1}^Q y_n^{(l)} \log g_l(x_n, \theta_1) \quad (11)$$

donde $g_l(x, \theta_1)$ es definida en la ecuación (8.7).

3. DISEÑO EXPERIMENTAL

3.1. DESCRIPCIÓN DE LA BASE DE DATOS

La metodología anterior se ha aplicado al conjunto de datos formado por 25 Estados Miembros de la UE para su clasificación de acuerdo a su ejecución o desempeño en el uso de los recursos de I+D+i. Debido a la disponibilidad de datos a la fecha de realización de este experimento (hasta 2008 en la mayoría de las bases de datos oficiales), se seleccionaron 100 ítems considerando cada país y cada año desde 2005 hasta 2008 como un único ítem (patrón país-año) dado que no se esperaban cambios sustanciales en estos patrones durante ese periodo. Por otro lado, las incorporaciones de Bulgaria y de Rumania se produjeron en el año 2007, por lo que la política y estrategia científica y tecnológica europea no fueron directamente aplicables a estos países y su inclusión no parecía apropiada para esta clasificación en este periodo de tiempo. Junto a la disponibilidad de los datos, otras razones complementarias para seleccionar el periodo 2005-2008 fueron evitar los efectos de los ciclos económicos y el que los datos pertenecientes a los años de la crisis económica mundial pudieran conducir a resultados ambiguos. Para evitar estos inconvenientes, se estima razonable seleccionar los datos hasta el año 2008, después del episodio de las hipotecas *sub-prime* en Estados Unidos el mes de agosto de 2007.

Las variables seleccionadas se basan en el modelo conceptual detallado en el capítulo VI que, como se vio, incluye el consenso alcanzado sobre el tipo de indicadores para medir los resultados de I+D+i de un país por los Manuales e informes auspiciados por la OECD y los estudios de PRO INNO Europe que elabora el Innovation Union Scoreboard (IUS), de modo que esta selección no se ha realizado de forma arbitraria sino bajo un marco teórico y conceptual elaborado.

Tanto la dificultad en encontrar datos sobre la I+D+i como la presencia de datos perdidos en un elevado número de países, fuerza a seleccionar variables de tipo *proxy* (de medición indirecta) en la mayoría de los casos. Así, se seleccionaron quince variables teniendo en cuenta sus relaciones y definiciones en el modelo conceptual causal, como se puede ver en la tabla 8.1, para la caracterización del desempeño en I+D+i a nivel de país (macro). Adicionalmente, en línea con el objetivo planteado con este análisis empírico, se añadió la variable que recoge el gasto en I+D+i del sector privado contemplado también en el modelo (*ERDPRI*), el total del personal dedicado a actividades relacionadas con la ciencia y la tecnología en general (*HUMST*), el personal dedicado directamente con la I+D+i (*RDPER*) y los aspectos de

BLOQUE 3- ANÁLISIS EMPÍRICO

rentabilidad de la innovación como son *PATREV* (licencias e ingresos de patentes provenientes del exterior) y *TRADEM* (marcas comerciales comunitarias). Respecto a las exportaciones, se ha escogido un indicador algo más amplio que el que aparece en el modelo (*HITECX*) como es el de las exportaciones de media y alta tecnología (*MHTEXP*).

Los datos se obtuvieron de las mismas fuentes oficiales que en el análisis anterior (página oficial de Eurostat, la base de datos del Banco Mundial y el ranking de países del SCImago Journal & Country Rank en el caso de publicaciones científicas por país), ajustándolo a cada 1000 habitantes.

Para el caso de los datos perdidos se ha llevado a cabo una imputación de datos cuando fue necesario (lo que ocurrió en un 1.3% de los casos); esta imputación se hizo mediante una regresión lineal (Buck, 1960). Se ajustaron seis modelos, uno para cada variable con datos perdidos (para las variables *PHD06*, *GERDGO*, *ERDPRI*, *TPATRE*, *PATREV* y *RDPER5*). La variable dependiente es el país-año para el que no hay datos. Las covariables o variables independientes son los restantes países-año que no tienen datos perdidos. Todos los modelos fueron ajustados a un nivel de significación de $\alpha=0,05$ y un coeficiente de determinación de $R^2 > 0,90$.

3.2. RESULTADOS

En este apartado se comparan las Redes Neuronales Evolutivas de Unidad Sigmoide (ESUNNs) y las Redes Neuronales Evolutivas de Unidad Producto (EPUNNs) con nueve métodos bien conocidos del estado del arte en la literatura del Aprendizaje Automático. Ocho de ellos han sido configurados y realizados en WEKA (Hall, et al., 2009) y el Análisis Discriminante Lineal Simple se encuentra disponible en SPSS (Norusis, 2007).

Los métodos usados para la comparación son:

MLP: clasificador de red neuronal que usa el algoritmo de retropropagación del error para calcular los pesos.

SLDA: un procedimiento estadístico multivariante que obtiene ecuaciones para clasificar las instancias (Landwehr et al., 2005).

C4.5: un algoritmo de construcción de clasificación que incluye un mecanismo de poda.

AdaBoost100: un *ensemble* de clasificación formado por árboles de decisión (en nuestro caso).

BLOQUE 3- ANÁLISIS EMPÍRICO

LMT: clasificador para construir árboles de modelo logístico, que son árboles de clasificación con funciones de regresión logística en las hojas (Sheskin, 2011).

NaiveBayes: clasificador cuyos valores de precisión de los estimadores numéricos son elegidos en base al análisis de los datos de entrenamiento y asados en frecuencias.

SLogistic: clasificador para la construcción de modelos de regresión logística lineal simplificados.

MLogistic: clasificador para la construcción de modelos multinomiales de regresión logística lineal con un parámetro para regularizar el modelo y evitar que exista sobreentrenamiento

SVMs (máquinas de soporte vectorial): clasificador para construir un modelo lineal (hiperplano de máximo margen) sobre un espacio kernel, resultando finalmente en un modelo no lineal. Una función kernel Gausiana se emplea para la selección de los hiper-parámetros de la SVM (parámetro de regularización C , y ancho de las funciones Gaussianas γ); un algoritmo de búsqueda de rejilla se aplicó con una validación cruzada de tipo *10-fold*, empleando los siguientes rangos: $C \in \{2^{-5}, 2^{-3}, \dots, 2^{-15}\}$ y $\gamma \in \{2^{-15}, 2^{-33}, \dots, 2^3\}$.

La tabla 8.3 muestra una comparativa de la precisión de los once métodos (tres estocásticos, de los que se elige el mejor modelo, y ocho deterministas) en la predicción del desempeño en I+D+i de cada país, basándonos en los cuatro grupos definidos por el algoritmo k medias. Desde un punto de vista descriptivo, el mejor modelo ESUNN, el mejor modelo EPUNN, el mejor modelo MLP, LMT, SLogistic, Mlogistic y SVM obtienen los mismos resultados en las medidas de $CCR_G(100)$, $MS_G(100)$ y $K_G(1)$ en el conjunto de datos de generalización. Las tres medidas revelan una clasificación perfecta. Respecto al número de conexiones para cada modelo (ver la última columna en la tabla 8.3), el mejor modelo ESUNN presenta el menor número de coeficientes (20); así que éste es el modelo seleccionado y que se explicará con detalle más adelante.

Con respecto a los resultados obtenidos en las 30 ejecuciones realizadas para los tres algoritmos estocásticos (ver tabla 8.4), es claro que los mejores resultados vienen dados por la red neuronal MLP, aunque la metodología ESUNN nos proporciona los segundos mejores resultados con una media alta (98% para el CCR_G) y una baja desviación típica ($SD=2,73\%$ para CCR_G). La principal diferencia entre las metodologías es el número de conexiones o coeficientes del modelo, porque para el modelo MLP la media es de 44 y para el modelo ESUNN la media es de 26,57 con una $SD= 3,47$. Esto es debido a las operaciones de poda de

BLOQUE 3- ANÁLISIS EMPÍRICO

conexiones y/o neuronas que realiza el algoritmo evolutivo. En el caso de estos algoritmos estocásticos, el diseño experimental se condujo usando un procedimiento *holdout* con 30 ejecuciones. Aproximadamente el 75% de los patrones fueron seleccionados aleatoriamente del conjunto de entrenamiento y el restante 25% para el conjunto de generalización.

Tabla 8.3. Resultados estadísticos del CCR_G , MS_G , K_G y número de conexiones (#conn.).

Método ^a	$CCR_G(\%)$	$MS_G(\%)$	K_G	#conn.
Mejor ESUNN	100	100	1	20
Mejor EPUNN	100	100	1	26
Mejor MLP	100	100	1	44
SLDA	96	80	0,94	24
C4.5	84	77	0,78	13
AdaBoost100	52	0	0,31	24
LMT	100	100	1	27
NaiveBayes	92	80	0,89	124
SLogistic	100	100	1	27
MLogistic	100	100	1	48
SVM	100	100	1	96

a. El mejor resultado se representa en negrita.

Tabla 8.4. Media y Desviación Típica (DT) de los métodos estocásticos.

	$CCR_G(\%)$	$MS_G(\%)$	K_G	# conn.
Método ^a	Media±DT	Media±DT	Media±DT	Media±DT
ESUNN	98,00±2,73	92,33±9,60	0,97±0,05	26,57±3,47
EPUNN	94,00±4,79	84,55±11,13	0,92±0,07	29,93±4,66
MLP	100,00±0,00	100,00±0,00	1,00±0,00	44,00±0,00

a. El mejor resultados se representa en negrita.

Consecuentemente, considerando la precisión o exactitud y el bajo número de conexiones, se recomienda elegir la metodología ESUNN para la clasificación de los patrones país-año según su desempeño en I+D+i.

3.3 EL MEJOR MODELO

Una vez el algoritmo fue ejecutado 30 veces, para satisfacer el teorema central del límite a la hora de considerar distribuciones normales en los test de comparaciones, se obtuvieron 30 mejores modelos, tanto para ESUNN como para EPUNN. Se escogió el mejor modelo ESSUN debido a su bajo número de conexiones. Este modelo consiste en nueve variables en la capa de entrada, dos nodos ocultos con funciones de transferencia de Unidad Sigmoide y tres nodos en la capa de salida que devuelven la probabilidad de que un patrón pertenezca a

cada clase. Al ser cuatro clases las utilizadas, de las probabilidades de tres de las clases se obtiene la probabilidad de pertenencia a la cuarta clase por diferencia. Las variables de entrada seleccionadas fueron *TERTIT*, *LLEARN*, *GERDGO*, *TRADEM*, *RLPGH*, *GPGROW*, *TPATRE*, *PATREV* y *HUMST*.

El interés en el mejor modelo ESUNN propuesto se encuentra en el hecho de que mejora los resultados (en cuanto a número de conexiones) de otras aproximaciones (tabla 8.4) y porque proporciona un modelo (no lineal) bastante simple. La tabla 8.5 muestra la expresión de la probabilidad de este mejor modelo ESUNN así como su desempeño.

Tabla 8.5. Expresión de la probabilidad y resultados del mejor modelo ESUNN

Modelo ESUNN
$f_1(x, \theta_1) = 4,69 - 18,69 * SU_1$
$f_2(x, \theta_2) = 9,3 - 13,62 * SU_2$
$f_3(x, \theta_3) = 0,49 + 13,49 * SU_1 - 23,59 * SU_2$
$f_4(x, \theta_4) = 0$
$SU_1 = \frac{1}{(1 + EXP - (-9,32 + 8,07 * TERTIT + 9,31 * TRADEM - 4,08 * RLPGH + 5,03 * PATREV + 3,68 * HUMST))}$
$SU_2 = \frac{1}{(1 + EXP - (-12,44 + 3,42 * TERTIT - 9,61 * LLEARN - 9,04 * GERDGO + 3,07 * GPGRO + 7,32 * TPATRE + 4,47 * HUMST))}$
$CCR_G = 100\% ; MS_G = 100\% ; K_G = 1$

La tabla 8.5 también nos muestra la construcción de la unidad sigmoide SU_1 como una función de las variables *TERTIT*, *HUMST* (lo que representa las habilidades y destrezas de la fuerza de trabajo), *PATREV*, *TRADEM* (variables que representan los beneficios de la innovación a través de las patentes y las marcas comerciales) y *RLPGH* (productividad real del trabajo por hora trabajada). La misma tabla nos muestra cómo se construye SU_2 como una función de las variables *TERTIT*, *HUMST* (como SU_1) y *LLEARN* (lo que denota la importancia de la educación para la adquisición de competencias por parte de la fuerza de trabajo), *GERDGO* (apoyo de la política pública a la inversión en I+D+i que estimula la inversión en el proceso de innovación en el sector privado), *TPATRE* (el principal indicador indirecto para la medición de las capacidades innovadoras de un país) y *GDPGRO* (el indicador mayormente utilizado para la medición del bienestar de una economía y de la sociedad en general).

BLOQUE 3- ANÁLISIS EMPÍRICO

Por otro lado, si se atiende a los coeficientes de la suma ponderada de las unidades sigmoide, junto a la importancia de las variables de acuerdo a la influencia sobre la probabilidad de que los patrones pertenezcan a una clase determinada (g_j) en la tabla 8.6, las variables que presentan los efectos más fuertes para medir la ejecución o desempeño de sus actividades de I+D+i a nivel macro son aquellas que presentan los signos (++) y (--). Las variables con el signo (+) o (-) tienen un efecto menor y aquellas en las que aparece el signo (=) no presentan efectos significativos.

Tabla 8.6. Influencia de cada variable en la probabilidad de pertenecer a un determinado grupo

Variable	Probabilidad Grupo 1 (a)	Probabilidad Grupo 2	Probabilidad Grupo 3	Probabilidad Grupo 4
TERTIT	(--)	(--)	(+)	(++)
LLEARN	(--)	(++)	(+)	(--)
GERDGO	(-)	(++)	(-)	(--)
TRADEM	(--)	(--)	(++)	(++)
RLPGH	(+)	(=)	(=)	(--)
GDPGRO	(=)	(=)	(=)	(=)
TPATRE	(--)	(++)	(+)	(--)
PATREV	(--)	(--)	(++)	(++)
HUMST	(-)	(=)	(-)	(=)

a. La probabilidad del grupo l es $g_l(x, \theta)$

Esto significa que un aumento en los valores de la variable implica principalmente un incremento (+), un descenso (-) o cambios no significativos (=) en la probabilidad de pertenecer a un determinado grupo.

4. DISCUSIÓN

Tras seleccionar nueve de las quince variables utilizadas para el análisis y que forman parte de nuestro mejor modelo ESUNN, (*TERTIT*, *LLEARN*, *GERDGO*, *TRADEM*, *RLPGH*, *GPGROW*, *TPATRE*, *PATREV* y *HUMST*), de nuestro análisis se puede destacar (ver tabla 8.6) las variables que podrían ejercer una mayor influencia sobre la probabilidad de pertenecer a la clase que representa un mejor desempeño en la I+D+i (países orientados a la innovación o clase 2). Estas variables son: el aprendizaje a lo largo de toda la vida (*LLEARN*), el total de

BLOQUE 3- ANÁLISIS EMPÍRICO

solicitudes de patentes de los residentes de un país (*TPATRE*) y el gasto bruto en I+D+i en el sector público, que incluye el sector gubernamental y el sector de la educación superior (*GERDGO*). Estas variables se corresponden con la educación, los resultados de la I+D+i y los facilitadores o impulsores de la ésta, respectivamente.

Mientras que estas variables mejoran la clasificación del país según su desempeño en I+D+i, las variables *TRADEM* y *PATREV* tienden a reducir la probabilidad de pertenecer a las clases de mejor desempeño. Estas variables representan los aspectos económicos de la I+D+i y podría interpretarse como que dichas variables no pueden ser aisladas para ser interpretadas sino que deben relacionarse con el sistema de innovación como un todo, tal y como se representa en nuestro modelo de red Bayesiana. Por otro lado, si se incrementan los valores del crecimiento del PIB y de la productividad real del trabajo (*RLPGH*) a su máximo valor normalizado (0,9), no varía la probabilidad de pertenencia de un país a la hora de ser clasificado en uno de los cuatro grupos.

Resulta muy interesante que el número de estudiantes de nivel terciario (*TERTII*) junto al aprendizaje a lo largo de la vida (*LLEARN*) tienden a incrementar la probabilidad de pertenecer a los grupos de mejor ejecución en I+D+i dado que la educación es básica para la mejora de las altas cualificaciones necesarias en una economía basada en el conocimiento y para alcanzar un desarrollo que se pueda llamar sostenible (Soudien, 2002).

Una de las contribuciones del trabajo a través de este diseño experimental es confirmar, de alguna forma, la relevancia de las variables relacionadas con las actividades de generación del conocimiento, específicamente el capital humano y la educación en general, así como la conveniencia de incrementar los recursos dedicados a las actividades de I+D+i (recursos financieros y humanos, fundamentalmente).

Al igual que estas últimas variables, la acumulación de capital humano ha sido considerada un importante motor económico en los modelos teóricos que se han ocupado de estudiar las fuentes del crecimiento económico. El análisis de nuestro modelo apoya la propuesta de que los países que invierten en capital humano tienen unos resultados mejores en I+D+i y en sus resultados económicos (Gemmell, 1995). También apoya el análisis académico que incluye los recursos en educación para explicar las variaciones en la productividad de la innovación entre países (Furman et al., 2002).

5. CONCLUSIONES DEL ANÁLISIS DE LOS RESULTADOS DE I+D+I

1. En este segundo análisis empírico enfocado desde el punto de vista de la ejecución de I+D+i, se han examinado nueve clasificadores pertenecientes al campo la Inteligencia Artificial y del Aprendizaje Automático y dos clasificadores de redes neuronales, para predecir la clasificación de 25 Estados Miembros de la UE. Esto se ha conseguido a través de la identificación del mejor modelo y analizando los principales factores de entre 15 variables que podrían influir en la probabilidad de pertenecer a una de las clases de desempeño de I+D+i, previamente establecidas mediante la técnica de agrupamiento *k*-medias.
2. Empleando este modelo se han identificado los principales factores que contribuyen en el desempeño en I+D+i (aprendizaje a lo largo de toda la vida, *LEARN*, el total de solicitudes de patentes de los residentes de un país, *TPATRE*, y el gasto bruto en I+D+i en el sector público, que incluye el sector gubernamental y el sector de la educación superior, *GERDGO*). En este sentido, el modelo propuesto podría ser útil para evaluar la efectividad y la eficiencia de las Estrategias de la UE relacionadas con la I+D+i.
3. El modelo presentado puede ser útil en términos de: i) una herramienta de gestión de apoyo a la toma de decisiones relativas a la I+D+i, dado que arroja información sobre las principales variables; ii) una herramienta de *benchmarking*, permitiendo a la UE comparar sus actividades propias de innovación con otros países de referencia y entre iniciativas nacionales de los Estados Miembros; y iii) una herramienta de monitorización y evaluación para los cuerpos institucionales y gubernamentales dado que permite que se realice un seguimiento del avance de los países según los cambios que se produzcan en algunos indicadores críticos que influyen en la predicción de su clasificación.
4. El análisis se muestra útil en el apoyo a los expertos y a los principales grupos de interés al clasificar los países y predecir sus resultados de I+D+i en años futuros en base a esa clasificación, de una forma similar a las puntuaciones de innovación y otras metodologías, con datos públicos disponibles de países de la UE.

CAPÍTULO IX. ANÁLISIS DESDE EL PRISMA DE UNA ECONOMÍA BASADA EN EL CONOCIMIENTO

1. INTRODUCCIÓN

El mundo académico, los responsables políticos, grupos de interés, consultores y medios de comunicación han mostrado un creciente interés en el papel de la creación del conocimiento como factor clave en el actual modelo de producción, permitiendo un aumento de las ventajas competitivas de las empresas y, consecuentemente, de las economías nacionales (von Krogh et al., 2000). Esto es especialmente crucial en el ambiente cambiante, de incertidumbre, ambiguo y complejo que presentan hoy las naciones (Johannessen and Olsen, 2010). Así, puede decirse que nos encontramos inmersos en la conocida como “economía del conocimiento” o “economía basada en el conocimiento” (EC) (para un marco teórico y conceptual ampliado, ver el capítulo 1).

En este contexto, los gobiernos tienen que planear sus inversiones y desarrollar sistemas educativos que formen trabajadores altamente cualificados para trabajos que requieren habilidades y competencias también altamente cualificadas si pretender alcanzar una sociedad basada en el conocimiento (Hsu et al, 2008).

Las herramientas de medición, los marcos, modelos y metodologías disponibles ayudan a los diferentes grupos de interés a analizar y referenciar las capacidades de los países como economías basadas en el conocimiento. Tales herramientas facilitan la adopción de políticas así como la creación de sistemas nacionales de conocimiento para un desarrollo holístico. Sin embargo, estas medidas arrojan diferentes calificaciones y rankings dependiendo de la naturaleza y el tipo de métodos de evaluación utilizados (Al Shami et al, 2012), rinden cuentas sobre resultados pasados, lo que implica responder cuestiones diversas que se plantean de forma subjetiva (Booyesen, 2002) y no predicen hacia dónde se dirige (o se podría dirigir) una determinada economía basada en el conocimiento si todas las variables empleadas en el modelo fueran conocidas.

En cuanto a la metodología empleada en este último análisis empírico, ya se vio en el capítulo 8 como la clasificación es una de las tareas de toma de decisiones más frecuentes de la actividad

BLOQUE 3- ANÁLISIS EMPÍRICO

humana. Se está ante un problema de clasificación cuando un objeto necesita ser asignado en una clase previamente definida (problema supervisado) o no (problema no supervisado) basado en el valor de los atributos observados relacionados con el objeto.

La mayoría de los algoritmos de clasificación se centran en predecir las etiquetas de clases que son consideradas nominales (no se les atribuye un orden) y los problemas reales son así tratados, como en el caso de nuestro anterior análisis. Sin embargo, muchos problemas de clasificación implican clasificar los datos en clases que tienen una naturaleza ordinal (problema de clasificación o regresión ordinal) (Zopounidis y Doumpos, 2002). El tratamiento de datos cuya naturaleza es ordinal es una situación con la que se encuentran frecuentemente los investigadores y la rama de las Ciencias Sociales no es una excepción. De hecho, un elevado número de problemas en esta rama son precisamente ordinales.

En este caso, se ha tratado de dar un paso más al abordarse el problema mediante la comparación de clasificadores nominales y ordinales con objeto de comprobar qué modelos realizan la mejor clasificación, ya que la naturaleza nominal u ordinal de los grupos obtenidos no es del todo obvia o natural y, aunque esté presente en la variable de salida, no se está seguro de encontrar ese orden en las variables de entrada.

La variable dependiente “a priori” (el grupo obtenido previamente con la técnica de agrupamiento jerárquico, como se verá) podría considerarse ordinal, del mismo modo que en la multitud de ejemplos que implican un ranking: las Agencias de Rating (p.e. Moody’s, Standard and Poor’s o Fitch)²³, el Índice de Competitividad Global del Fondo Monetario Internacional, el Índice de Economía del Conocimiento del Banco Mundial (KEI), el mencionado Cuadro de Mando de Indicadores de Innovación de la Unión (IUS) de la Comisión Europea, el Índice de Desarrollo Humano de las Naciones Unidas, los múltiples rankings de universidades, etc.

De acuerdo con todo lo anterior, el objetivo de este análisis empírico se centra en la obtención de un modelo que prediga la clasificación de 162 observaciones país-año (54 países en el periodo de tiempo 2007-2009) gracias a su asignación previa a un grupo. Los grupos han sido previamente definidos en relación a su progreso hacia una economía del conocimiento mediante el uso de 11 indicadores que representan los cuatro pilares de una economía del conocimiento, según la metodología del Banco Mundial, y aplicando el método de

²³ Por ejemplo, Standard and Poor’s clasifica el riesgo-país con una codificación alfanumérica ordinal que va desde el país menos arriesgado (AAA) al más arriesgado (C+).

agrupamiento jerárquico (un algoritmo no supervisado) para detectar los patrones de comportamiento de estos países. De esta forma, se ofrece una metodología alternativa al agrupamiento k -medias del caso anterior en cumplimiento de nuestro objetivo de ofrecer distintas visiones del problema y variadas metodologías para abordarlo a los diversos grupos de interés. En las categorías anteriores, se podría establecer un orden basado en su estado de transición hacia la consecución de una EC.

A continuación, se contruyeron tres clasificadores ordinales y tres clasificadores nominales para asignar cada país-año a cada uno de los grupos obtenidos. Los resultados de los clasificadores ordinales se compararon con los resultados de los nominales a la hora de evaluar cuáles de ellos clasifica mejor, determinando de esta forma la naturaleza nominal u ordinal del objeto de estudio.

Después de esta breve introducción, se revisará la literatura más relevante sobre la evaluación de la economía del conocimiento de los países y sobre las metodologías de clasificación (apartado 2). Seguidamente, se describirá la metodología concreta planteada en este estudio (apartado 3). En el apartado 4, se lleva a cabo el diseño experimental con la correspondiente descripción del conjunto de datos y las variables empleadas así como los principales resultados y su discusión. Finalmente, en el apartado 5 se extraen las principales conclusiones.

2. REVISIÓN DE METODOLOGÍAS

En relación a la literatura de evaluación de los logros de un país para convertirse en una EC, hasta la fecha, los estudios empíricos forman dos grupos bien destacados: técnicas estadísticas y los indicadores compuestos o índices. Los indicadores compuestos son reconocidos ampliamente como una herramienta útil para el diseño de políticas públicas a nivel nacional y para establecer prioridades políticas, referenciar y comparar o monitorizar resultados, al igual que en el caso de la evaluación de los resultados de las actividades de I+D+i. Un considerable número de ellos ha sido propuesto también por organismos internacionales (Nardo et al., 2005).

La modelización y la medida de una EC es una continuación de los esfuerzos previos por medir los componentes relacionados con la misma tales como el desarrollo humano regional y la construcción de estructuras o marcos generales del capital intelectual (Hanley and Malafsky, 2002). Varios modelos de medición de estos llamados “activos del conocimiento” han sido

BLOQUE 3- ANÁLISIS EMPÍRICO

propuestos por organizaciones de gran reputación como los modelos de la Comisión Económica de las Naciones Unidas para Europa o los modelos de la iniciativa eEurope de Medición de los Activos de Conocimiento Nacionales.

De entre estas metodologías, una de las principales iniciativas en el mundo es la desarrollada por el Banco Mundial conocida como Metodología de Evaluación del Conocimiento (Knowledge Assessment Methodology o KAM), que ofrece un Índice de Economía del Conocimiento para un gran número de países (KEI). Brevemente, la metodología consiste en un conjunto de 148 variables estructurales y cualitativas que puede ser usada para referenciarse (benchmarking) con otros países (World Bank, 2012; Chen and Dahlman, 2006). Un número reducido de variables sirve como medida indirecta para los cuatro pilares que son relevantes para el desarrollo de una economía del conocimiento (Régimen Económico, Sistema de Innovación, Educación y la Infraestructura de Información y Comunicación) ofreciendo así un denominado “indicador básico”.

En cuanto a la metodología propuesta concretamente, en general, se ha considerado una clasificación ordinal en vez de una nominal multiclase debido a que cada vez son más los trabajos que utilizan esta metodología cuando la variable de clasificación es ordinal. En los problemas de aprendizaje de clasificación nominal no se considera un orden entre las clases. El enfoque más popular es descomponer el problema en uno de clasificación mediante varios modelos bi-clase.

En los problemas de clasificación ordinal²⁴, las categorías objetivo presentan un orden. Esto distingue la clasificación ordinal de la convencional donde la variable dependiente no presenta un orden. Recientemente, se ha dedicado un gran esfuerzo al problema de la clasificación ordinal (Frank y Hall, 2001; Chu y Keerthi, 2005; Yu et al., 2006; Cardoso y da Costa, 2007) en los campos de investigación operativa y de inteligencia computacional, donde el aprendizaje de los modelos de clasificación ordinal ha consistido en una generalización de algunas técnicas nominales multicriterio y, más recientemente, ha conducido a nuevos desarrollos teóricos²⁵. El problema aparece como diferente a los de regresión estándar en la medida en que no puede establecerse una distancia entre las etiquetas.

Como ejemplos, las técnicas de clasificación ordinal tienen aplicaciones muy variadas donde la naturaleza de las instancias presenta un orden tales como la recuperación de información

²⁴ Esta clase de problemas es conocido también como *regresión ordinal*.

²⁵ Para una revisión de los modelos multicriterio para el aprendizaje de modelos ordinales, ver Sousa et al., 2013.

BLOQUE 3- ANÁLISIS EMPÍRICO

(Herbrich et al., 1999; Chu y Keerthi, 2007), los modelos econométricos (Mathieson, 1995), el riesgo crediticio (Doumpos, 2002; Xu et al., 2009) o el análisis de los genes (Pyon and Li, 2009), por nombrar sólo algunos.

Desde un punto de vista del aprendizaje, la estructura ordinal de la variable de clasificación, Y , se considera información adicional que se debe tratar de explotar y esto es lo que los métodos existentes para la clasificación ordinal tratan de realizar esencialmente (Frank y Hall, 2001; Cardoso y da Costa, 2007). A grandes rasgos, el valor de la información del orden es igual al aumento (esperado) en el rendimiento cuando se resuelve el problema de ordenar sin tener en cuenta esta información adicional.

Por supuesto, no es en absoluto evidente la suposición de que una estructura ordinal se mantenga en la práctica y, por lo tanto, que estas técnicas sean capaces de explotar la información ordinal, contenida en un problema. El propósito de este caso de estudio es investigar empíricamente si los clasificadores ordinales son efectivos en este sentido.

En este caso se va a emplear un grupo de algoritmos conocido por el término *threshold models* o métodos de umbral que agrupan la mayoría de las propuestas para regresión ordinal. Éstos se dedican a dos tareas principalmente (Fouad and Tino, 2012):

- i) cómo encontrar la línea óptima de proyección, representando el supuesto orden lineal de las clases, en el que los datos de entrada serán proyectados.
- ii) cómo posicionar óptimamente los umbrales definiendo los intervalos de la etiqueta de clase de forma que los márgenes de separación entre las clases vecinas se maximice.

Podría afirmarse que el grupo más conocido de modelos de respuesta ordinal se basa en la estimación de las probabilidades acumuladas (Fullerton and Xu, 2012). El modelo de Probabilidades Proporcionales (Proportional Odd Model, POM), es uno de los primeros modelos específicamente diseñados para la regresión ordinal. Se trata de un método para analizar datos binarios ampliado para modelos con respuestas ordenadas. A varios autores, especialmente a McCullagh (1980), se les ha atribuido la idea de usar la función de enlace logit para modelos de respuesta ordenada y, por tanto, el POM.

En el contexto de las Máquinas de Vectores Soporte (SVMs), ha sido desarrollada una clase de modelos bajo el nombre Regresión Ordinal de Vectores Soporte (SVOR). Shashua and Levin (2003) propusieron una formalización general de las SVMs aplicadas a datos ordinales basadas en dos principios: i) el principio de margen-fijo, en el que el margen del par de clases más

cercanas se maximiza, conduciendo a márgenes iguales entre dos clases vecinas y ii) el principio de la suma de márgenes, lo que permite diferentes márgenes y sólo la suma de todos los $(Q-1)$ márgenes se maximiza (asumiendo que hay Q categorías ordenadas). Sin embargo, el orden de los umbrales de las clases no se impone. Por lo tanto, se considera la formulación extendida de SVOR con restricciones explícitas de orden (SVOREX) donde el orden de los umbrales de clase se considera explícitamente (Chu y Keerthi, 2007). Más allá, también se emplea el modelo SVOR conocido como SVOR con restricciones de orden implícito (SVORIM), que no impone las restricciones de forma explícita, sino que se demuestra el cumplimiento de las mismas cuando los patrones de todas las clases son considerados para calcular el error de cada patrón. Tanto el modelo SVOREX como el modelo SVORIM son dos de los métodos adicionales seleccionados para la comparación entre clasificadores en el presente análisis.

3. METODOLOGÍA

3.1 AGRUPAMIENTO JERÁRQUICO Y DESCRIPCIÓN DE CLASES

Como se vio en el capítulo VIII, el análisis de agrupamiento permite asignar un conjunto de datos en grupos, de forma que los datos que se encuentran en un mismo grupo son más similares entre ellos que a aquellos que se encuentran en otro grupo y cada elemento pertenece sólo a un grupo. Para identificar un número apropiado de grupos, una amplia variedad de algoritmos de agrupamiento han aparecido en los últimos años, siendo la clasificación más importante la que distingue entre métodos *particionales* y *jerárquicos* (Lin, 2013).

Nuestra tarea inicial es emplear de nuevo una técnica de agrupamiento que detecte los patrones de comportamiento en los países seleccionados (países-año) con respecto a su estado de transición hacia una economía del conocimiento. En ausencia de etiquetas de clase o conocimiento sobre los grupos actuales, es muy difícil seleccionar un criterio para juzgar si un conjunto de características es mejor que otro. La técnica adoptada en este caso, una vez empleada la técnica de agrupamiento de k -medias en el caso anterior y en línea con tratar de ofrecer un abanico de metodologías, es el agrupamiento jerárquico. Se trata de una técnica no supervisada de aprendizaje automático empleada en muchos campos y a menudo preferida a un agrupamiento no jerárquico cuando el número de grupos Q no se conoce o cuando el

BLOQUE 3- ANÁLISIS EMPÍRICO

interés descansa en las relaciones entre los objetos, como sucede en los estudios taxonómicos (Questier et al., 2002).

El agrupamiento jerárquico fue ejecutado mediante el algoritmo COBWEB por su bajo coste computacional, al igual que el algoritmo k -medias, y disponible también en el software WEKA, (Hall et al., 2011). El algoritmo COBWEB fue desarrollado por investigadores del campo del aprendizaje automático para agrupar patrones en un conjunto de objetos-atributos dando lugar a un dendograma de agrupamiento o árbol de clasificación que caracteriza cada grupo con una descripción probabilística. Este algoritmo opera basado en la función categoría de utilidad (CU) que mide la calidad del agrupamiento. El agrupamiento jerárquico, en general, proporciona una mejor comprensión de los datos mediante el ensamblaje de todos los objetos en el dendograma de forma que cada sub-grupo es uno de sus nodos y las combinaciones de sub-grupos crea una jerarquía-estructura que es más informativa que el conjunto de grupos no estructurado que ofrece el agrupamiento particional (Wu et al, 2009).

El árbol de clasificación resultante del agrupamiento jerárquico o dendograma obtenido puede verse en la figura 9.1.

Para cada nodo, el identificador del grupo corresponde al número sin paréntesis y el número entre paréntesis es el número de patrones país-año incluidos en el correspondiente grupo.

Los grupos se refieren a clases en la figura 9.1 y se decidió la forma de agruparlos observando los patrones país-año e intentando reflejar el grado de transición hacia una EC. Se trató de obtener grupos representativos con un razonable número de elementos en cada clase o grupo. Un análisis adicional de estos grupos se realizó en base a sus centroides (ver tabla 9.1).

Mediante el empleo de un análisis económico llevado a cabo por expertos como una medida complementaria al presente agrupamiento jerárquico, se etiquetaron los grupos y se describió el perfil de cada uno. En relación al número de grupos seleccionado, cuatro grupos alcanzan una adecuada diferenciación; esto evita la necesidad de información masiva que una medida muy detallada requeriría si se eligiese un número mayor de grupos y ofrece simultáneamente la necesaria información que se diluiría si se escogiese un número menor.

Los principales hallazgos del análisis se pueden resumir en que, basándonos en los cuatro pilares de la EC (Régimen Económico, Innovación, Educación e Infraestructuras de TICs) y a través de los 11 indicadores seleccionados, los países se dividen en cuatro grupos de acuerdo al

BLOQUE 3- ANÁLISIS EMPÍRICO

correspondiente centroide, pero alejados del primero, en particular en lo que se refiere al nivel de patentes y de publicaciones científicas. Cuentan con un ambiente legal, institucional y comercial con valores cercanos a la media general y con tasas brutas de ingreso en los niveles de educación primaria, secundaria y terciaria que se encuentran cerca del primer grupo pero que, en el caso de la variable de educación primaria, es ligeramente superior.

Aunque se encuentran en segunda lugar, destacar que se encuentran muy lejos del nivel de ciencia y tecnología (patentes, publicaciones científicas y TICs) del primer grupo.

Grupo 3 (*EC moderadas*). Agrupa los patrones país-año que muestran unos resultados bajos en innovación (bajo número de solicitudes de patentes y la mitad aproximadamente del número de publicaciones científicas), y un nivel más bajo en la tasa de ingreso bruta en los tres niveles educativos (primaria, secundaria y terciaria). En relación a la variable que representa la medida en que los agentes confían y se comprometen a las reglas de la sociedad, muestra valores medios bajos, pero no son negativos. Estos países también presentan valores bajos en las variables que miden el pilar de las TICs. A pesar de ello, incluye los países con las mayores tasas de crecimiento de las conexiones de banda ancha de internet (p.e. Grecia, Rumania y Bulgaria tuvieron un crecimiento anual medio muy elevado en el periodo 2007-2009).

Grupo 4 (*fase incipiente de la EC*). Los patrones-año recogidos en este grupo muestran unos resultados muy bajos en cada uno de los pilares de la EC. Junto a Chile y México, este grupo representa principalmente las economías de transición de Europa (como el grupo 3 en el caso de estudio n° 2) y de China. Así, en relación al régimen de económico, debe decirse que se encuentran aún lejos de un marco favorable debido a la transición de un ambiente institucional de tipo soviético a uno capitalista. Por consiguiente, la especial situación socioeconómica explicaría el amplio rango de diferencia con los valores del resto de grupos. Con respecto al pilar de la innovación, el nivel de patentes y de publicaciones es muy bajo (casi inexistente en algunos casos) y, finalmente, el número de líneas de teléfono así como el número de suscriptores de internet están muy alejados del valor medio de las variables correspondientes.

3.2. CLASIFICADORES NOMINALES Y ORDINALES EVALUADOS

Con el propósito de realizar las comparaciones entre métodos, obtener el mejor clasificador y determinar la naturaleza de este problema, se han incluido en la experimentación clasificadores

nominales y ordinales bien conocidos en la literatura del Aprendizaje Automático (ver apartado 2).

La formalización general de este tipo de problemas, lógicamente, es similar a la clasificación nominal del análisis del capítulo anterior. Se trata de predecir la etiqueta y de un vector de entradas \mathbf{x} , donde $\mathbf{x} \in X \subseteq \mathbb{R}^K$ e $y \in Y = \{C_1, C_2, \dots, C_Q\}$, es decir, \mathbf{x} es un vector de entrada K -dimensional e y está en un espacio de etiquetas de Q categorías diferentes (grupos, en nuestro caso). El objetivo es encontrar una regla de clasificación o función $f: X \rightarrow Y$ para predecir las etiquetas de los nuevos patrones dado un conjunto de entrenamiento de N puntos, $D = \{(\mathbf{x}_i, y_i), 1 \leq i \leq N\}$. Si las etiquetas están ordenadas, de forma que $C_1 \prec C_2 \prec \dots \prec C_Q$, esto actúa como una restricción adicional en el tipo de problemas de naturaleza ordinal. El símbolo \prec representa la relación de orden entre las diferentes etiquetas. Muchas de las medidas y algoritmos de regresión ordinal consideran la posición de la etiqueta, o lo que es lo mismo, la posición de la etiqueta en escala ordinal, lo que se puede expresar mediante mediante la función $\mathcal{O}(\cdot)$, de forma tal que $\mathcal{O}(C_q) = q, 1 \leq q \leq Q$.

A) Clasificadores nominales.

Los clasificadores nominales empleados son:

- *SimpleLogistic*: El algoritmo SLogistic construye modelos de regresión logística multinomial mediante el uso del algoritmo conocido como Logiboot (un algoritmo de boosting que se ejecuta mediante un proceso iterativo en el cual se añaden las variables independientes de una en una, siendo el número de iteraciones guiado por validación cruzada) propuesto por Friedman et al. (2000) para ajustar modelos aditivos de regresión logística mediante máxima verosimilitud. Estos modelos son una generalización de los modelos de regresión logística (lineal). Esta versión de los algoritmos se basa en el control del número de variables en el modelo para evitar el sobre-entrenamiento.
- *MLogistic*: se trata de un algoritmo para la construcción de un modelo de regresión logística multinomial con un parámetro de regularización para evitar el sobre-entrenamiento penalizando los coeficientes grandes, basado en el trabajo de le Cessie y van Houwelingen (1992).
- *K-NN*: la regla del K vecino más cercano es una de las técnicas más exitosas y simples empleadas para resolver tareas de clasificación. Consiste en usar como clasificador los

Tabla 9.1. Estadísticos descriptivos de los grupos (media o centroide, desviación estándar y patrón país-año de referencia).

CÓDIGO		VARIABLE	Media±DT (observaciones país-año de referencia para cada for grupo y variable)			
Pilar 1: Régimen económico e institucional			Grupo 1 ^(a)	Grupo 2 ^(b)	Grupo 3 ^(c)	Grupo 4 ^(d)
<i>TNTBA</i>		Barreras arancelarias y no arancelarias	85,149±4,575 (Japón-09)	85,400±1,975 (Italia-07)	83,324±5,271 (Chipre-07)	75,638±9,605 (Chile-08)
<i>REGQU</i>		Calidad normativa	1,523±0,291 (Rep. Corea-08)	1,109±0,262 (Croacia-09)	0,721±0,550 (Macedonia-08)	0,009±0,539 (Túnez-09)
<i>RULEL</i>		Imperio de la Ley	1,640±0,298 (Letonia-09)	0,912±0,383 (Hungria-09)	0,300±0,568 (Grecia-07)	-0,394±0,502 (Chile-07)
Pilar 2: Educación y habilidades						
<i>SCHOO</i>		Matriculación en primaria (% bruto)	102,379±3,976 (Alemania-09)	103,585±6,220 (Hungria-09)	99,155±4,520 (Bulgaria-07)	103,323±7,803 (Rumania-07)
<i>SECON</i>		Matriculación en secundaria (% bruto)	105,982±11,667 (Reino Unido-07)	100,271±8,094 (Lituania-09)	92,691±5,255 (Chipre-07)	87,525±5,333 (Bielorrusia-07)
<i>TERTI</i>		Matriculación en nivel terciario (% bruto)	66,122±17,479 (Japón-07)	64,153±12,894 (Rep. Checa-08)	56,251±13,201 (Chile-09)	42,699±19,073 (Rumania-07)
Pilar 3: Infraestructuras de TICs						
<i>TELEP</i>		Número de líneas de teléfono	1.619,142±239,961 (Rep. Corea-07)	1.587,687±160,324 (Hungria-07)	1.387,418±207,976 (Bielorrusia-09)	1.035,833±288,396 (Ucrania-07)
<i>FIXBI</i>		Suscriptores de banda ancha de internet	280,205±49,153 (Estonia-07)	177,438±29,787 (Latvia-08)	107,771±16,798 (Chipre-08)	41,149±28,340 (Rusia-09)
<i>INTERN</i>		Usuarios de internet	768,438±82,969 (Malta-09)	534,872±73,289 (España-07)	457,583±114,678 (Poland-08)	235,764±94,148 (Albania-09)
Pilar 4: Sistema de innovación						
<i>TP-ATRE</i>		Solicitudes de patentes	1,630±1,224 (Reino Unido-09)	0,372±0,413 (Irlanda-07)	0,099±0,093 (Chipre-08)	0,075±0,079 (Bielorrusia-08)
<i>SPUBLI</i>		Artículos científicos y técnicos	1,864±0,706 (Estonia-08)	1,104±0,509 (Rep. Checa-08)	0,591±0,308 (Grecia-07)	0,169±0,100 (Turquia-09)

^(a) Australia, Austria, Bélgica, Canadá, Dinamarca, Estonia, Finlandia, Francia, Alemania, Hong Kong SAR China, Islandia, Irlanda-09, Japón, República de Corea, Letonia-09, Luxemburgo, Malta-09, Países Bajos, Nueva Zelanda, Noruega, Eslovenia-09, Suecia, Suiza, Reino Unido, Estados Unidos.

^(b) Croacia-09, Chipre-09, República Checa, Grecia-09, Hungría, Irlanda-07-08, Israel, Letonia-07-08, Lituania, Malta-07-08, Portugal, Eslovenia-07-08, España.

^(c) Bielorrusia-09, Bulgaria, Chile-09, Croacia-07-08, Chipre-07-08, Grecia-07-08, Macedonia FYR-08-09, Polonia, Rumania-08-09, República de Eslovaquia.

^(d) Albania, Armenia, Azerbaiyán, Bielorrusia-07-08, Bosnia y Herzegovina, Chile-07-08, China, Georgia, Macedonia FYR-07, Méjico, Moldavia, Rumania-07, Rusia, Túnez, Turquía, Ucrania.

BLOQUE 3- ANÁLISIS EMPÍRICO

patrones de entrenamiento y asignar el valor de clase mediante el voto mayoritario de entre los vecinos del patrón nuevo a clasificar (el patrón es asignado a la clase más común entre sus K vecinos más cercanos). Si $K=1$, entonces el patrón es simplemente asignado a la clase de su vecino más cercano (Cover y Hart, 1967).

Los dos primeros también se utilizaron en el análisis de los recursos de I+D+i del capítulo anterior.

B) Clasificadores ordinales.

En el caso de los clasificadores ordinales, los *métodos de umbral* constituyen el enfoque más común. La principal idea de estos métodos es modelar los problemas de regresión ordinal desde una perspectiva de la regresión en el sentido de que se asume que existen valores reales subyacentes, aunque son inobservables. De esta forma, se suelen estimar dos cosas distintas:

- Una función $f(\mathbf{x})$ que es capaz de predecir estos valores reales, revelando la naturaleza de la variable subyacente asumida.
- Un vector $\mathbf{b} \in \mathbb{R}^{Q-1}$ (donde Q es el número de clases) de parámetros libres (umbrales) donde $\mathbf{b} = (b_0, b_1, \dots, b_{Q-1}, b_Q)$, que representa un conjunto de intervalos en la recta, de forma que satisfacen las restricciones $b_1 \leq b_2 \leq \dots \leq b_{Q-1}$. Cada intervalo está asociado a una clase, permitiendo diferentes anchos para las diferentes clases.

Todos los clasificadores ordinales seleccionados para la experimentación siguen esta estructura y son brevemente resumidos de la siguiente forma:

- *Modelo de Probabilidades Proporcionalas* (Proportional Odd Model, POM). Este modelo se basa en la siguiente representación general: $g^{-1}(P(y \leq C_q | \mathbf{x})) = b_q - \mathbf{w}^T \mathbf{x}, 1 \leq q \leq Q$, que satisface el orden estocástico. De esta forma, $g^{-1}: [0, 1] \rightarrow (-\infty, +\infty)$ es una función monótona (la función de enlace inversa), y b_q es el umbral para la clase C_q . Una variable latente es la principal motivación del POM, que presenta la forma lineal $f(\mathbf{x}) = \mathbf{w}^T \mathbf{x}$. Entonces, se asume una función de densidad de la probabilidad sobre las etiquetas de clase para un vector de características \mathbf{x} . Así, la etiqueta C_q del conjunto de entrenamiento se observa si y sólo si $f(\mathbf{x}) \in [b_{q-1}, b_q]$, donde la función f (utilidad latente) y $\{b_0, b_1, \dots, b_{Q-1}, b_Q\}$ son los parámetros que deben aprenderse de los datos, con $b_0 = -\infty$ y $b_Q = +\infty$. Una vez que el modelo de la variable latente queda definido, $f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + \varepsilon$, donde ε es el componente aleatorio con esperanza $E[\varepsilon] = 0$ y distribuido

BLOQUE 3- ANÁLISIS EMPÍRICO

de acuerdo a $F\mathcal{E}$, la clave es seleccionar una asunción apropiada para la distribución $F\mathcal{E}$. El modelo acumulativo se obtiene eligiendo la inversa de la distribución $F\mathcal{E}^{-1}$ como la inversa de la función enlace g^{-1} . La elección más común para $F\mathcal{E}$ es la función logística estándar, donde el logit se modela de la forma siguiente:

$$\text{logit}(y \preceq C_q | \mathbf{x}) = \ln(\text{odds}(y \preceq C_q | \mathbf{x})) = \ln\left(\frac{P(y \preceq C_q | \mathbf{x})}{1 - P(y \preceq C_q | \mathbf{x})}\right) = \mathbf{w}^T \cdot \mathbf{x} + b_q \quad (9.1)$$

Esta forma del modelo permite la estimación directa de las probabilidades $P(y = C_q | \mathbf{x}) = P(y \preceq C_q | \mathbf{x}) - P(y \preceq C_{q-1} | \mathbf{x})$.

- *Support Vector Ordinal Regression* (SVOR).

Las Máquinas de Vectores Soporte (SVM) (Cortes and Vapnik, 1995) pueden considerarse el método de aprendizaje kernel más común para la clasificación. Se estructuran como un modelo de función de base con un cálculo de tipo *kernel* para el producto escalar aplicado a los vectores de inputs transformados $\phi(\mathbf{x})$. Aquí, $\phi(\mathbf{x})$ representa el vector de patrones de entrada \mathbf{x} en un espacio dimensional muy grande, que se relaciona con \mathbf{x} mediante una transformación específica (Cortes and Vapnik, 1995). Todos los cálculos son realizados empleando sólo la función kernel de reproducción, que se define como:

$$k(\mathbf{x}, \mathbf{x}') = \langle \phi(\mathbf{x}) \cdot \phi(\mathbf{x}') \rangle \quad (9.2)$$

donde \cdot representa el producto escalar en el gran espacio dimensional. La idea es que los modelos lineales en este espacio no serán lineales en el espacio original.

Las SVMs se basan en la idea de separar dos clases diferentes— son primero definidas para dos clases y luego extendidas al caso multiclase— usando un hiperplano especificado por su vector normal \mathbf{w} y el sesgo b , $\mathbf{w} \cdot \phi(\mathbf{x}) + b = 0$.

El hiperplano de separación óptimo es el que maximiza la distancia entre el hiperplano y los puntos más cercanos de ambas clases (llamada margen) para un problema bi-clase, resultando la mejor predicción para los datos inobservables.

Adicionalmente, los márgenes estrictos se reemplazan por márgenes suaves, lo que permite el manejo del ruido y de los errores de pre-etiquetado. Las variables de holgura ξ_i , son empleadas para relajar las restricciones de márgenes estrictos. El hiperplano de separación óptimo con el margen máximo es un problema de optimización de Programación Cuadrática. Para tratar el

BLOQUE 3- ANÁLISIS EMPÍRICO

caso multiclase, un enfoque “1-contra-1” es el que se considera habitualmente (Hsu and Lin, 2002).

La formulación de SVM ha sido adaptada a la definición de la regresión ordinal (Máquina Soporte Vectorial para regresión ordinal o SVOR), mediante la definición de un umbral distinto b_j para cada clase, y la adaptación del problema QP . En lugar de decidir la clase por el signo de proyección $\mathbf{w}^T \cdot \mathbf{x}$, la línea real correspondiente se divide en varios intervalos consecutivos con un vector de umbrales \mathbf{b} , dando lugar a hiperplanos paralelos con el mismo vector de proyección \mathbf{w} y diferentes umbrales b_j . En este caso, se consideran dos implementaciones distintas para esta idea (Chu y Keerthi, 2007):

- SVOR con restricciones explícitas (SVOREX). El problema QP incluye explícitamente el conjunto de restricciones asegurando el orden entre los umbrales, mientras las variables de holgura para el j -th hiperplano paralelo consideran patrones de clase j y $j + 1$.
- SVOR con restricciones implícitas (SVORIM). Se usa el siguiente principio: en lugar de considerar sólo los errores de las muestras de las categorías adyacentes, se permite que las muestras en todas las categorías contribuyan a los errores para cada hiperplano. Se demuestra que las desigualdades ordinales en los umbrales son satisfechas implícitamente en la solución óptima (Chu y Keerthi, 2007).

4. DISEÑO EXPERIMENTAL

4.1. DESCRIPCIÓN DE LA BASE DE DATOS

Para el presente análisis se emplearán 54 países que serán clasificados de acuerdo a su correspondiente progreso hacia la consecución de una economía del conocimiento y para la obtención de un clasificador (nominal u ordinal) que prediga dicha clasificación según los datos disponibles. En esta ocasión, los países seleccionados fueron los que pertenecen al continente europeo (excepto Mónaco, Andorra, Serbia and Montenegro)²⁶, añadiendo el resto de países miembros de la OECD, así como China y Hong Kong, por su relevancia política y socioeconómica en el panorama internacional, y Túnez como país emergente, por su proximidad geográfica y por su participación en el trabajo del comité de inversiones de la OECD.

²⁶ Mónaco y Andorra no se seleccionaron por su reducida población; Serbia y Montenegro por la falta de estadísticas tras su independencia en el año 2006.

BLOQUE 3- ANÁLISIS EMPÍRICO

Debido a la disponibilidad de los datos a la fecha de elaboración de este trabajo de investigación (hasta 2009 en la mayoría de las bases de datos oficiales y, especialmente, en las variables relacionadas con la Educación), se seleccionaron 162 patrones considerando cada país y cada año como un solo ítem compuesto (país-año) para el periodo 2007-2009.

Para la selección de las variables, se considera igualmente nuestro modelo conceptual causal (ver figura 5.1 del capítulo V), además de los indicadores propuestos por la Metodología de Evaluación del conocimiento del Banco Mundial (Banco Mundial, 2012) que encajan en los escenarios de nuestra red. Como se vio anteriormente, el Índice de la Economía del Conocimiento (KEI) del Banco Mundial se construye como una media simple de cuatro subíndices, los cuales representan los siguientes cuatro pilares de la economía del conocimiento:

Pilar 1: Económico y régimen institucional. Un régimen económico e institucional apropiado provee de incentivos para el uso eficiente de conocimiento ya existente o nuevo conocimiento y para el emprendimiento.

Pilar 2: Educación y competencias. Uno de los pilares más importantes de la EC es el llamado capital humano, empleado en algunas teorías de crecimiento económico como sinónimo de EC. La educación aparece como un factor crítico de la vida socioeconómica permitiendo a las personas crear y compartir conocimiento y adquirir las competencias necesarias para el desempeño de puestos de trabajo que requieren de personas altamente calificadas.

Pilar 3: Información e infraestructuras de comunicación. La economía del Conocimiento está fuertemente basada en el desarrollo y uso de las Tecnologías de la Información y de la Comunicación (TICs), siendo la característica que comúnmente se asocia al concepto de EC en el lenguaje cotidiano. Muchos de los cambios globales que han tenido lugar a principios de los años 70 están dirigidos por la fuerte explosión de las tecnologías de la comunicación. El uso global de las TICs ha cambiado radicalmente el panorama de los negocios, los gobiernos y la vida social en general, con nuevos patrones de comportamiento y de interrelación a todos los niveles.

Pilar 4: Sistema de Innovación. Este pilar trata de recoger la necesidad de un sistema de innovación eficiente de empresas, centros de investigación, institutos de investigación,

BLOQUE 3- ANÁLISIS EMPÍRICO

universidades y otro tipo de organizaciones y cuerpos gubernamentales donde la difusión del conocimiento y la creación del mismo son medidos típicamente por la habilidad que tiene un país de patentar y de publicar artículos científicos, a pesar de las desventajas de este tipo de indicadores.

Como puede observarse, los cuatro pilares se encuentran recogidos en nuestro modelo y de él se seleccionarán las variables. Para el primer pilar, se seleccionan las variables que propone la metodología del Banco Mundial (este pilar viene representado en nuestro modelo mediante la variable o nodo “Marco institucional” y puede medirse mediante las variables de tipo proxy que propone la metodología del Banco Mundial).

Estas variables tratan de representar la trayectoria de los resultados globales de la economía que “ilustran cómo una economía está en la actualidad empleando el conocimiento para su desarrollo económico y social general” (Banco Mundial, 2012). Los datos se obtuvieron de las fuentes que se han venido utilizando: base de datos del Banco Mundial, el ranking de artículos por países de SCImago Journal & Country Rank y, adicionalmente, de la Heritage Foundation (2012) para las variables del pilar del régimen institucional y económico. Las variables fueron ajustadas a 1000 habitantes cuando fue necesario así como la correspondiente imputación de datos perdidos (lo que ocurrió en 18 casos); esta imputación se realizó mediante regresión lineal como en el resto de casos de (Buck, 1960). Se ajustaron un total de cinco modelos, uno para cada variable con datos perdidos (*TPATRE*, *FIXBI*, *SCHOO*, *SECOND* y *TERTI*).

Para este caso de estudio, las variables seleccionadas y sus descriptivos son las que se recogen en la siguiente tabla 9.2:

Tabla 9.2. Descripción de las variables empleadas y sus estadísticos descriptivos.

CODE	VARIABLE	DESCRIPCIÓN	MEDIA±DT
Pilar 1: Régimen institucional y económico			
<i>TNTBA</i>	Barreras arancelarias y no arancelarias	Puntuación asignada a cada país basada en el análisis de sus barreras arancelarias y no arancelarias al comercio internacional, tales como cuotas de importación así como requerimientos de etiquetado y de licencias que sean muy restrictivos.	82,038±7,296
<i>REGQU</i>	Calidad normativa	Incidencia de políticas hostiles al mercado tales como los controles de precios o la inadecuada supervisión bancaria, y la percepción de las cargas impuestas por una regulación excesiva en áreas tales como el comercio exterior y el desarrollo de negocios.	0,943±0,734
<i>RULEL</i>	Imperio de la Ley	Este indicador mide el grado de confianza y acatamiento de las reglas de la sociedad por los distintos agentes. En particular se centra en la calidad del cumplimiento de los contratos, la policía y los tribunales, y la probabilidad de que se cometan delitos y actos de violencia.	0,796±0,933
Pilar 2: Educación y habilidades/competencias			
<i>SCHOO</i>	Matriculación en primaria (% bruto)	La tasa bruta de matriculación en primaria es el ratio de la matriculación total, independientemente de la edad, respecto a la población del grupo de edad que oficialmente corresponde al nivel de educación primaria según la clasificación ISCED de la OECD.	101,855±5,818
<i>SECON</i>	Matriculación en secundaria (% bruto)	La tasa bruta de matriculación en secundaria es el ratio de la matriculación total, independientemente de la edad, respecto a la población del grupo de edad que oficialmente corresponde al nivel de educación primaria según la clasificación ISCED de la OECD.	97,818±11,782
<i>TERTI</i>	Matriculación en nivel terciario (% bruto)	La tasa bruta de matriculación en terciaria es el ratio de la matriculación total, independientemente de la edad, respecto a la población del grupo de edad que oficialmente corresponde al nivel de educación primaria según la clasificación ISCED de la OECD.	58,142±19,179
Pilar 3: Infraestructuras de información y comunicación			
<i>TELEP</i>	Líneas de teléfono	Líneas principales de teléfono, por 1.000 habitantes. Incluye los canales de la red digital de servicios integrados (RDSI) y suscriptores de inalámbricos fijos.	1.425,006±340,150

Tabla 9.2. Descripción de las variables empleadas y sus estadísticos descriptivos (cont.)

CODE	VARIABLE	DESCRIPCIÓN	MEDIA±Desv. Tip.
<i>FIXBI</i>	Suscriptores de banda ancha de internet	El número de suscriptores de banda ancha con una línea digital, módem de cable u otra tecnología de alta velocidad, por 1.000 habitantes.	175,152±106,165
<i>INTERN</i>	Usuarios de internet	Los usuarios de internet son personas con acceso a la red mundial de internet, por 1000 habitantes.	542,012±233,420
Pilar 4: Sistema de innovación			
<i>TPATRE</i>	Solicitudes de patentes	Patentes concedidas por país de origen y oficina de patentes, por 1000 habitantes.	0,782±1,083
<i>SPUBLI</i>	Publicaciones científicas	Publicaciones científicas por país, por 1000 habitantes.	1,108±0,878

4.2. MÉTRICAS DE EVALUACIÓN

Las dos métricas seleccionadas para la evaluación de los clasificadores cuantifican la exactitud de las n etiquetas predichas para un conjunto de patrones dado $\{y_1^*, y_2^*, \dots, y_n^*\}$ respecto a los objetivos reales $\{y_1, y_2, \dots, y_n\}$. Estas medidas evalúan dos facetas diferentes que tienen que ser consideradas en los problemas de clasificación ordinal (Chu and Keerti, 2007): si los patrones son correctamente clasificados (*CCR*) y si los clasificadores predicen una clase tan cerca a la clase real como sea posible (*Mean Absolute Error, MAE*). La medida *CCR* se detalló junto a otras métricas en la sección 2.2 del capítulo VII para el caso de los clasificadores nominales. En este caso, las métricas seleccionadas se consideran más adecuadas para el caso de los clasificadores ordinales.

La primera métrica, la tasa de clasificación correcta o *Correct Classification Rate (CCR)*, es la fracción de las predicciones correctas sobre las muestras individuales.

$$C = \frac{1}{N} \sum_{i=1}^N I(y_i^* = y_i) \quad (9.3)$$

Donde $I(\cdot)$ es la función de pérdida cero-uno y N es el número de patrones en el conjunto de evaluación (conjunto de generalización o año 2009, en nuestro caso). Un buen clasificador trata de alcanzar el *CCR* más alto posible en un problema dado. La segunda métrica es la Media Absoluta del Error (*MAE*) que toma en cuenta la magnitud en clasificaciones erróneas. Es la desviación media en valor absoluto de la predicción del objetivo real,

$$MAE = \frac{1}{N} \sum_{i=1}^N |\mathcal{O}(y_i^*) - \mathcal{O}(y_i)| \quad (9.4)$$

donde $|\mathcal{O}(y_i^*) - \mathcal{O}(y_i)|$ es la distancia entre el valor real de la etiqueta y el valor de predicción, tomando dicha distancia como número de etiquetas en la escala ordinal.

4.3. RESULTADOS

El conjunto de experimentos considerado en este diseño experimental incluye todos los métodos citados en el apartado 2. La tabla 9.3 compara la exactitud o precisión de los seis métodos (tres nominales y tres ordinales) a la hora de predecir el estado o el progreso hacia una EC de un país dado según su clasificación en el grupo correspondiente. Un primer

objetivo fue comprobar si los clasificadores ordinales realmente tienen mejores resultados que los nominales (como es de esperar por la naturaleza de este problema económico) y, consecuentemente, escoger el mejor modelo para la predicción de la clase de pertenencia.

El conjunto de datos fue dividido en dos periodos de tiempo consecutivos. El primer periodo fue utilizado para entrenar el modelo y el segundo para evaluarlo. Consecuentemente, el conjunto de entrenamiento consistió en un conjunto de 108 observaciones país-año en el periodo 2007-2008 (datos anuales). El conjunto de generalización estaba formado por 54 observaciones país-año en el año 2009.

Tabla 9.3. CCR_G y MAE_G para el conjunto de generalización de los métodos evaluados^a

		Métricas	
Tipo de clasificador	Clasificador	CCR_G	MAE_G
NOMINAL	Logistic	83,33%	0,167
	1-NN	81,48%	0,185
	Slogistic	83,33%	0,185
ORDINAL	POM	85,19%	0,148
	SVOR-EXC	79,63%	0,204
	SVOR-IMC	87,04%	0,130

^a El mejor resultado aparece en negrita y el segundo mejor en cursiva.

Los resultados para las dos medidas de evaluación (CCR_G y MAE_G ; ver Eqs. (9.3) y (9.4)) se incluyen en la tabla 9.3. Basándonos en estos resultados, la primera conclusión es que valores de precisión altos (CCR_G) se obtienen mediante los clasificadores ordinales, lo que muestra que considerar el problema como una tarea de naturaleza ordinal puede proveernos de información más precisa del estado de transición hacia una economía del conocimiento que considerarlo como de clasificación nominal. Concretamente, el modelo SVORIM es el que proporciona mejores resultados ($CCR_G = 87,04\%$). Los valores de MAE_G también son bajos siendo el más bajo es el que se puede encontrar para el modelo SVORIM. Un valor de MAE_G

de 0,130 significa que las predicciones del clasificador son, de media, 0,130 categorías más bajas o más altas (en la escala ordinal) que las categorías objetivo, con siete patrones país-año mal clasificados en 2009 o año del conjunto de generalización.

Consecuentemente, el clasificador ordinal SVORIM es el seleccionado por su mayor precisión. Por otro lado, aunque se pudiera pensar que las diferencias no son excesivamente grandes con respecto a los clasificadores nominales, sería interesante tener en cuenta algunas consideraciones adicionales.

Primero, en situaciones reales como la presente, los grupos de interés y los decisores políticos encuentran de mayor utilidad las clases ordenadas tanto para monitorizar políticas específicas y estrategias como para llevar a cabo prácticas de *benchmarking*. Así, la identificación de grupos ordenados podría ayudar a los decisores a identificar perfiles o patrones de comportamiento en los países útiles para monitorizar, comparar y ofrecer una nueva perspectiva en un mundo donde más y más rankings están siendo manejados para este tipo de cuestiones (De Smet et al., 2012).

En segundo lugar, realizando una clasificación ordinal, se proyectan las observaciones/patrones en una dimensión, mientras el análisis discriminante emplea múltiples proyecciones. Para interpretar el modelo y estudiar las proyecciones, resulta más natural usar una sola proyección que ordene los patrones.

Finalmente, la tasa de clasificación correcta (*CCR*) mide la proporción de los objetos clasificados correctamente en una muestra de datos. Para los problemas ordinales esto no tiene por qué ser siempre el criterio más adecuado de decisión de los mejores resultados. El *CCR* tácitamente asume que el coste de las clasificaciones erróneas es el mismo para cada una de los objetos clasificados de forma incorrecta. La naturaleza ordinal del problema, sin embargo, implica que un error de clasificación en un grupo adyacente no es igual en términos de coste que un error en dos grupos adyacentes (por ejemplo).

En nuestro caso, el *CCR* para el grupo 1 es del 92% (23 patrones clasificados correctamente de un total de 25; ver tabla 9.4), siendo el grupo que incluye el mayor número de patrones (46,29% de los patrones en el conjunto de generalización), y sólo dos patrones no son clasificados correctamente y, además, en el grupo adyacente número 2. En ningún caso, los errores se producen en más de una clase adyacente.

BLOQUE 3- ANÁLISIS EMPÍRICO

El modelo seleccionado SVORIM clasificó correctamente la mayoría de los patrones (87,04%). Analizando los errores concretos cometidos en el conjunto de generalización, se comprueba que corresponden a los siguientes siete países en el año 2009: Bielorrusia, Bosnia, Chile, Letonia, Polonia, Rusia y Eslovenia. Letonia, Eslovenia y Polonia han sido mal clasificados tanto en los métodos nominales como los ordinales. Bielorrusia, Chile, Letonia y Eslovenia son clasificados en una categoría más baja que la que les corresponde según el agrupamiento jerárquico, mientras que Bosnia, Rusia y Polonia son clasificados en una categoría superior.

Tabla 9.4. Matriz de confusión para el mejor clasificador ordinal (SVORIM)

		Predicción (SVORIM) ^a			
#patrones	CLASE REAL	GRUPO 1	GRUPO 2	GRUPO 3	GRUPO 4
25	GRUPO 1	23	2	0	0
10	GRUPO 2	0	10	0	0
7	GRUPO 3	0	1	4	2
12	GRUPO 4	0	0	2	10

^a Las siete observaciones mal clasificadas son: Bielorrusia-09, Bosnia-09, Chile-09, Letonia-09, Polonia-09, Rusia-09 and Eslovenia-09.

Estas diferencias pueden deberse al hecho de que los datos cuantitativos empleados en nuestro modelo arrojan información sobre los datos históricos de los pilares seleccionados y sobre las características estructurales fundamentales. Son básicamente retrospectivos, mientras que en las ciencias sociales y a nivel macro pueden producirse situaciones inesperadas; un modelo no puede incorporar todas las posibilidades y situaciones que se pueden presentar en la realidad. El examen de sucesos futuros por parte de los expertos del área tiene que ser una herramienta complementaria para testear la vulnerabilidad de una país respecto a una variedad de shocks generados, tanto interna como externamente. Así, los aspectos cualitativos y de juicio del análisis son inevitables, incluso en la interpretación de indicadores cuantitativos.

Continuando esta línea de argumentación, las clasificaciones erróneas pueden ser explicadas con la ayuda de la literatura y las opiniones de los expertos. La primera razón se encuentra en el hecho de que cinco de los siete países mal clasificados son países de régimen post-socialista y tres de ellos son, además, Nuevos Estados Miembros (NEM). En la mayoría de los NEM se da poca importancia a las regiones y a su política. Después de su entrada en la UE, parece más difícil para estos países construir economías basadas en el conocimiento; la brecha en las actividades relacionadas con el conocimiento y la innovación se agrava, en parte, debido a su débil capacidad tecnológica (Kołodko, 2001; Orłowski, 2001). Polonia es un caso especial donde su transición al sistema capitalista se facilitó por el hecho de que un reducido número de empresas privadas había sido tolerado por el régimen comunista (Boettke et al., 2008).

En el caso de Letonia, cambia su clasificación del grupo 2 al grupo 1 en el año 2009 pero, los modelos de clasificación propuestos, continúan considerando (predicen) que pertenece al grupo 2. El análisis de las variables empleadas muestra el impresionante incremento en los resultados del pilar de las TICs junto a la elevada tasa bruta de ingreso en la educación terciaria, lo que llevó a su clasificación al grupo 1. A pesar de ello, un significativo descenso en el número de estudiantes se observó en los años previos debido principalmente a problemas demográficos. Esta característica no se corresponde con un país con un alto nivel educacional y con un fuerte incremento en la penetración de las TICs en 2009. Todo ello perfila una situación especial que nuestro modelo considera atípica y no lo suficientemente decisiva para el cambio de grupo.

Lo mismo se observa en el caso de Eslovenia: cambia del grupo 2 al grupo 1 en el año 2009 pero el mejor clasificador sigue considerando que pertenece al segundo grupo. Muestra un buen nivel de solicitudes de patentes y de publicaciones científicas, así como un incremento global en la tasa de matriculación en todos los niveles de educación contemplados. Esta es la razón para el cambio del grupo 2 al grupo 1 en los resultados del agrupamiento jerárquico. Sin embargo, este incremento no ha sido lo suficientemente relevante para considerarlo como miembro del grupo 1 para nuestro mejor clasificador SVORIM.

Rusia pertenece al grupo 4 en todos los años bajo estudio, pero los tres clasificadores ordinales seleccionados han realizado la predicción de que Rusia pertenecería al grupo 3 en 2009, es decir, no cambia de grupo pero los tres clasificadores sí que predicen que mejora su situación respecto a una EC. Esto podría deberse a su especial transición de una economía

comunista a una capitalista y a su marco regulatorio, que dificulta el modelar su comportamiento respecto a los demás países de la muestra.

5. DISCUSIÓN Y CONCLUSIONES DEL ANÁLISIS DESDE EL ENFOQUE DE UNA ECONOMÍA BASADA EN EL CONOCIMIENTO

1. En este tercer análisis empírico, se examinaron tres métodos de clasificación nominales y tres ordinales procedentes del campo del Aprendizaje Automático con un doble objetivo: i) obtener un modelo capaz de predecir la clasificación de los países de acuerdo a su estado de transición hacia una EC como una herramienta de monitorización y *benchmarking* para diseñadores de políticas públicas y grupos de interés. Previamente, se ejecutó un algoritmo de agrupamiento jerárquico para obtener los grupos en los que los países podrían ser agrupados de acuerdo a las variables usualmente empleadas para caracterizar una EC; y ii) para comprobar nuestra hipótesis sobre la naturaleza ordinal de este problema económico.

2. Para alcanzar este objetivo, se emplearon 162 patrones formados por 54 países (los países de Europa, países no europeos miembros de la OECD más China, Hong Kong y Túnez) en el periodo 2007-2009 y, mediante un algoritmo de agrupación jerárquico, se obtuvieron cuatro grupos: economías del conocimiento avanzadas, seguidoras, moderadas y en estado inicial o incipiente.

3. Una vez se obtuvieron los grupos de clasificación, se comparó la exactitud de los métodos de Aprendizaje Automático nominales y ordinales. Los resultados muestran que los clasificadores ordinales obtienen mejores resultados que los nominales y que el mejor método para predecir la clasificación de los países fue SVORIM, por su gran precisión (*CCR*) y su bajo valor de *MAE*. De esta forma, este trabajo ha confirmado cómo el orden de los grupos puede mejorar los resultados de los clasificadores nominales, con predicciones más exactas y apoyando la hipótesis inicial de la posible naturaleza ordinal de los grupos. Razones complementarias de tipo económico y metodológico se expusieron para apoyar esta afirmación.

4. De esta forma, la metodología presentada también puede ser útil, junto a las anteriores, como i) una herramienta de gestión para el apoyo a la toma de decisiones, dado que permite predecir el estado de transición de un país hacia una EC; ii) como una herramienta de *benchmarking* que compare los resultados con otros países de referencia y poder aprender de las

mejores prácticas. El clasificador podría ayudar a los expertos y a los principales grupos de interés a predecir la clasificación de los países de acuerdo a su estado de transición hacia una EC en años futuros o para otros países distintos de los seleccionados, de una forma similar a los rankings, índices compuestos, calificaciones y otras metodologías con datos de países disponibles.

5. La principal conclusión es que los algoritmos de clasificación ordinal ejecutan mejor la clasificación que los nominales, por lo que se puede considerar la naturaleza de este problema como ordinal. El modelo obtenido podría apoyar el seguimiento de las estrategias nacionales y de algunos aspectos clave relacionados con la creación del conocimiento y la innovación, en términos generales, al permitir comprobar el avance o retroceso de los países en función de los cambios en la clasificación en uno u otro grupo.

BLOQUE 4

DISCUSIÓN,

CONCLUSIONES Y

FUTURAS LÍNEAS DE

INVESTIGACIÓN

CAPÍTULO X. DISCUSIÓN, CONCLUSIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN

1. DISCUSIÓN

En el Bloque introductorio, se planteaban como principales objetivos generales: i) la evaluación de la eficiencia de las políticas públicas de I+D+i de los países de la UE en el marco de un desarrollo sostenible (DS); ii) el diseño de un modelo conceptual (red Bayesiana causal), ante la ausencia de este tipo de modelos en la evaluación de políticas públicas de I+D+i a nivel macro, que apoye esta evaluación desde la perspectiva del DS y otras complementarias y, iii) la aplicación de metodologías alternativas del campo de la Inteligencia Artificial y el Aprendizaje Automático para el estudio, monitorización o seguimiento del uso de los recursos de I+D+i, desde los enfoques adicionales contemplados en el modelo (resultados de la I+D+i y economía del conocimiento o EC) y gracias a la clasificación de los países en grupos.

El objetivo referido al diseño de un modelo conceptual causal (segundo objetivo) que sirva como apoyo al estudio de la eficiencia de las políticas públicas de I+D+i (primer objetivo), contribuye a confirmar la integración de la I+D+i, la consecución de una EC y un DS como un todo, posibilitando con su diseño estudiar el fenómeno bajo tres primas de una misma realidad (referida al crecimiento económico para el incremento del bienestar social) y plantear metodologías alternativas a las ya existentes para el estudio de los recursos de I+D+i (tercer objetivo). Cabe, según lo expuesto, concebir estos objetivos como un sistema integrado.

A continuación, se presentan y discuten las principales contribuciones aportadas por la investigación en función de los objetivos generales fijados:

1. Para la consecución del segundo objetivo general²⁷ de obtener un modelo conceptual causal que apoye a los grupos de interés en la evaluación de la eficiencia de las políticas públicas de I+D+i, desde una perspectiva que se centra en la consecución de un DS, se siguió la metodología Expert-based Cooperative Analysis (EbCA) propuesta por Gibert et al. (2010). Las principales características y ventajas de esta metodología son:

²⁷ Se abordó en primer lugar el segundo objetivo dado que el modelo servirá de apoyo a la consecución del primer objetivo general de la investigación: la evaluación de la eficiencia de los recursos públicos de I+D+i.

- a. Elicitar conocimiento implícito de forma iterativa.
 - b. Formalizar la incorporación del conocimiento experto (explícito e implícito) a la metodología empleada para la evaluación de la eficiencia.
 - c. Enriquecer los resultados obtenidos.
- En base a esta metodología, el primer paso consistió en construir una base de conocimiento, a través de la consulta a expertos y de una amplia revisión bibliográfica. Una vez construida la base de conocimiento, era necesario seleccionar un modelo que representase dicho conocimiento (dicha revisión bibliográfica y conceptual queda plasmada en los capítulos I, II y III del Bloque 1).
 - La representación del modelo que se seleccionó, siguiendo la Teoría de la Causalidad de Pearl (2009), fue el de una red Bayesiana. La base de conocimiento explícito permitió seleccionar las variables (aleatorias y continuas) y las relaciones de causalidad entre ellas. El modelo así diseñado representa la versión inicial de cómo relacionar el bienestar de una sociedad (en este caso, europea) con el DS y la EC, asumiendo las teorías del crecimiento económico endógeno. Estas teorías consideran al cambio tecnológico y a las innovaciones, junto a los recursos de I+D+i que los promueven, como las fuerzas motoras de dicho crecimiento económico.
2. Una vez diseñado el modelo de red Bayesiana causal, la siguiente tarea a realizar para alcanzar el primer objetivo fue evaluar la eficiencia de las políticas públicas de I+D+i. (apoyándonos en las variables, relaciones de causalidad y conocimiento que incluye la red Bayesiana). Con este fin, el método concreto seleccionado fue el Análisis Envolvente de Datos (DEA). Siguiendo los pasos propuestos por el método EbCA, la base de datos original es modificada aplicando una base de reglas de conocimiento, de forma que se incorpora este conocimiento experto al modelo. El diseño del modelo también permitió plantear el uso de un DEA en dos pasos enlazados, en función a la secuencia de fases del proceso que sigue la I+D+i, denominándolo 2LS-DEA (two-linked steps Data Envelopment Anaylisis).
 - La evaluación de la eficiencia de las políticas públicas de I+D+i de la UE, a través del 2LS-DEA aplicado a la base de datos transformada mediante la aplicación de la base de reglas, se realizó para cada uno de los escenarios socioeconómicos de DS en dos pasos enlazados. El realizar la evaluación de la eficiencia en dos pasos, constituye una aportación adicional de la investigación, surgida del planteamiento de utilizar las variables y las relaciones de causalidad del modelo diseñado.

- Se obtuvieron las puntuaciones de eficiencia de cada una de las políticas de I+D+i de los 27 EEMM de la UE. Estas eficiencias, así como las variables más relevantes para que la política correspondiente a cada país se tornara eficiente, se comentarán en el apartado de conclusiones.
 - Del análisis de los resultados de las puntuaciones de eficiencia obtenidas mediante el 2LS- DEA, siguiendo la metodología EbCA, se elicó nuevo conocimiento (implícito) que mejoró de forma iterativa la red Bayesiana causal inicial diseñada hasta la obtención del modelo final presentado, constituyendo así mismo un resultado del primer análisis del Bloque 3 empírico.
3. Respecto al tercer objetivo, se analizaron las dos facetas restantes: los resultados de I+D+i de forma global y la EC. En ambos casos, primero se emplearon técnicas de agrupamiento (*clustering*) para la clasificación de los países en grupos homogéneos, empleando el algoritmo de agrupamiento *k*-medias y el jerárquico. Posteriormente, se utilizaron técnicas de clasificación supervisada nominal y ordinal pertenecientes al campo de la Inteligencia Artificial y del Aprendizaje Automático para la construcción de modelos capaces de predecir la clasificación de los países en cada una de las clases o grupos obtenidos previamente gracias a las técnicas de agrupamiento citadas.

En el caso del estudio de los resultados de I+D+i, para la construcción del clasificador, se desarrollaron modelos de redes neuronales artificiales (RNAs) de Unidad Sigmoide (SUNN) y de Unidad Producto (PUNN) entrenadas con Algoritmos de Programación Evolutiva (APEs). Estas Redes Neuronales Evolutivas (ESUNN y EPUNN, respectivamente) se compararon con 9 métodos bien conocidos en la literatura (MLP, SLDA, C4.5, Adaboost100, LMT, NaiveBayes, Slogistic, MLogistic y SVMs). De la comparación se concluyó que el mejor modelo era el modelo ESUNN, debido a su gran precisión y a su menor número de conexiones.

Para el caso del estudio de la transición de los países hacia una EC, los clasificadores seleccionados fueron tres clasificadores ordinales (POM, SVOREX y SVORIM) y tres clasificadores nominales (SLogistic, MLogistic, SVM), a efectos de comparar y descubrir la naturaleza nominal u ordinal del problema. Se utilizó un agrupamiento jerárquico para definir con mayor flexibilidad los distintos grupos de países de acuerdo a su estado de transición hacia la consecución de una economía del conocimiento. De esta forma, los grupos obtenidos guardaron una relación de orden, al menos en la etiqueta que se le asignó, y nuestro objetivo fue el corroborar que ese orden estaba presente en las variables

de entrada. El modelo que presentó mejores resultados en las métricas de evaluación de los clasificadores fue el SVORIM, el cuál es un modelo ordinal. La obtención de un clasificador ordinal como mejor modelo, junto a varias consideraciones, concluyeron con la consideración de este tipo de problema como un problema de naturaleza ordinal.

Con ello, se dio cumplimiento al tercer objetivo perseguido de ofrecer un abanico de metodologías alternativas a las existentes en los estudios relacionados.

Los principales avances que suponen la consecución de estos tres objetivos quedan resumidos en los siguientes párrafos:

- i) La aportación de un modelo conceptual causal a este campo de estudio que reúne de forma global un conjunto de variables, relaciones de causalidad y escenarios, cuyo diseño ha reunido una importante y consistente base de conocimiento procedente de la revisión de la bibliografía especializada y de la consulta a expertos.
- ii) La posibilidad de emplear este modelo como herramienta en el diseño de políticas públicas de I+D+i, que afectan a los objetivos de los gobiernos europeos de alcanzar un desarrollo sostenible a través de una economía basada en el conocimiento. Una de las utilidades de esta herramienta es su poder visual, permitiendo concentrar las relaciones entre las variables y factores micro y macroeconómicos más relevantes en el crecimiento económico, ante la gran cantidad de indicadores, índices y cuadros de mando propuestos a nivel internacional (datos cuantitativos, al fin y al cabo), que no aportan una perspectiva global del problema de una forma tan representativa.
- iii) El modelo ha permitido la evaluación del uso de los recursos de I+D+i en distintos escenarios socioeconómicos, superando el enfoque economicista de la denominada “política de la ciencia”, detectando las políticas más y menos eficientes y evidenciando las áreas de trabajo a abordar a nivel de estrategia nacional y/o europea.
- iv) Gracias a la importante labor realizada construyendo una base de conocimiento experto, se ha podido plantear el diseño de análisis empíricos alternativos, tanto en el enfoque adoptado como en la metodología empleada, aplicando el DEA para la evaluación de la eficiencia técnica relativa, así como algoritmos de agrupamiento y técnicas clasificación procedente de los campos de la Inteligencia Artificial y el Aprendizaje Automático que, hasta la fecha, son escasamente empleadas en el campo objeto de estudio.

v) Cada uno de los objetivos propuestos, una vez alcanzados y lejos de cerrar las posibilidades de estudio, suponen un punto de partida para futuros trabajos en términos de las amplias posibilidades combinatorias a realizar entre las variables de los escenarios recogidos en la red Bayesiana, las técnicas y metodologías aplicables, los periodos de tiempo y los países seleccionados (dando lugar a otros tantos análisis empíricos alternativos a los aquí expuestos).

2. CONCLUSIONES

En cuanto a la comprobación de las hipótesis planteadas al inicio de la investigación y que están vinculadas a cada uno de los objetivos, pueden contrastarse simultáneamente a la consecución del objetivo correspondiente y, de esta forma, se obtienen las siguientes conclusiones:

- ✓ Hipótesis 1 (identificación de las variables más relevantes y sus relaciones para evaluar la eficiencia del uso de los recursos de I+D+i en base a la literatura existente y la opinión de los expertos que posibilite la construcción de un modelo conceptual de red Bayesiana causal).

Esta hipótesis queda verificada con la consecución del segundo objetivo, como se puede desprender de los párrafos anteriores. Efectivamente, la revisión bibliográfica especializada (tanto académica como de los organismos internacionales dedicados a la normalización y recopilación de datos) permitió justificar la selección de cada una de las variables de acuerdo al ámbito de los países de la UE o de un análogo nivel de desarrollo económico y, de forma relevante, las relaciones de causalidad entre ellas. Adicionalmente, aplicando la metodología EbCA, se ha elicitado conocimiento implícito que se ha incorporado al diseño del modelo de forma iterativa, mejorándolo y constatando posteriormente que se trata de una red Bayesiana, siguiendo los postulados de la Teoría de la Causalidad de Pearl (2009).

- ✓ Hipótesis 2 (se pueden detectar las variables más relevantes en el estudio del uso de los recursos de I+D+i en función de las metodologías propuestas).

Mediante la consecución del segundo y el tercer objetivos, esta hipótesis queda verificada de forma parcial. Así, del análisis 2LS-DEA se detectó que las variables más relevantes en la mejora de la eficiencia de las políticas públicas de I+D+i, en términos de porcentaje de ganancia del output para alcanzar la eficiencia en cada uno de los escenarios, fueron las

patentes, (escenario Tradicional) y el número de Doctores (escenario de Educación). Los daños causados por el dióxido de carbono (escenario Medio Ambiente) y el índice de dependencia de las personas mayores (escenario de Salud) son las variables a incrementar por países concretos, pero no destaca en estos dos escenarios especialmente una variable para incrementar su valor en todos los países. En el escenario de Sociedad de la Información, *INTUSE*₂₀₀₈, *BROSUS*₂₀₀₈ y *MOBILE*₂₀₀₈, (número de usuarios de internet, número de suscriptores de banda ancha y número de usuarios de móvil, respectivamente) son las variables que deben modificarse para Estonia, Letonia e Irlanda, sin destacar ninguna en general como en el escenario anterior.

En cuanto a las técnicas de clasificación procedentes del campo de la Inteligencia Artificial, el mejor modelo ESUNN seleccionado en el caso de la clasificación nominal y en el estudio de los resultados generales de las actividades de I+D+i, arrojan como resultado que las variables que más influyen en la probabilidad de pertenecer a la clase 2 (países orientados a la innovación) son el aprendizaje a lo largo de la vida, el total de solicitudes de patentes y el gasto bruto total en I+D+i. El número de estudiantes matriculados en el nivel terciario, también influye en la clasificación de un país en la clase 4 (países de alta innovación). Estas consideraciones, junto a la importancia del número de doctores en el caso del análisis 2LS-DEA destacan el papel fundamental de la educación. Además, resulta reveladora la importancia del número de patentes en estos dos análisis empíricos.

Sin embargo, respecto al clasificador ordinal, SVORIM, que fue seleccionado como el mejor en el caso de la economía del conocimiento, no es posible realizar una fácil interpretación de los coeficientes del modelo, por lo que se plantea para futuras investigaciones el análisis detallado de este tipo de modelos. Sí se puede apreciar que la mejor clase respecto al estado hacia la EC (grupo 1, EC avanzadas) cuenta con un valor elevado del número de patentes y muy alejado de la siguiente clase, por lo que se vuelve a constatar la relevancia de la actividad patentadora.

- ✓ Hipótesis 3 (el problema puede ser abordado desde distintas perspectivas dada la interrelación entre los conceptos de I+D+i, EC y DS en el bienestar de la sociedad).

Basándonos en el planteamiento teórico abordado en el Capítulo 3, esta hipótesis queda verificada con la consecución de los tres objetivos planteados, dado que no se han

encontrado incongruencias que impidieran diseñar y analizar los resultados de los tres análisis empíricos planteados desde cada una de dichas perspectivas.

- ✓ Hipótesis 4 (si un país presenta una política de I+D+i eficiente, también debería ser eficiente en el resto de escenarios socioeconómicos analizados).

Esta hipótesis, que se encuentra relacionada con el primer objetivo, queda verificada sólo parcialmente. Así, sólo se cumple en el caso de la política de I+D+i de Suecia (eficiente en todos los escenarios) y queda próxima a verificarse en el caso de Bélgica (eficiente en todos los escenarios excepto en el de Sociedad de la Información) y en el de Finlandia (eficiente en todos los escenarios menos en el de Salud y el de Medio Ambiente).

En el caso de Bulgaria, por el contrario, alcanza la eficiencia más baja en el escenario tradicional, siendo la tercera política menos eficiente de las 27 políticas de I+D+i (no es eficiente en ningún escenario), mientras que casi alcanza la eficiencia técnica relativa en los escenarios Económico, de Salud y de Medio Ambiente.

Resulta llamativo el caso de Alemania ya que sólo es eficiente en el escenario de Innovación y casi alcanza la eficiencia en el escenario Económico (escenarios fundamentales en el contexto que se está analizando). En el escenario de Medio Ambiente, también se encuentra muy cercano a la eficiencia, pero, en el que aparece como menos eficiente (84,04), es precisamente en el escenario Tradicional. Esto podría conducir a una importante implicación si se analizase con mayor detenimiento: aunque no emplea de forma eficiente los recursos públicos dedicados a la I+D+i que conducen a resultados como son las aplicaciones de patentes y las publicaciones científicas, sin embargo, una vez seleccionadas éstas como los inputs del escenario de Innovación y Económico, son empleados de forma eficiente en términos de outputs como el número de empresas innovadoras, beneficios procedentes de la innovación, crecimiento del PIB, de la productividad real del trabajo, de las exportaciones de alta tecnología y del nivel de desempleo.

Los países del Este son eficientes en el escenario Económico cuando no son eficientes en el resto de escenarios y presentan bajas eficiencias en el escenario Tradicional. Como se apuntó en el agrupamiento de países-año realizado en el capítulo VIII, estos buenos resultados económicos pueden deberse a la demanda creciente por el cambio de régimen político y económico y a la abundancia de recursos naturales que presentan estos países.

Por lo tanto, la presente hipótesis sólo se cumple de forma parcial.

- ✓ Hipótesis 5 (los países que invierten mayores recursos públicos en I+D+i, son los más eficientes en el escenario Tradicional).

Los países que más invierten en I+D+i por parte del sector público son Suecia, Finlandia y Francia, por ese orden. Por tanto, si los países más eficientes en el escenario Tradicional (que emplea esta variables como input) son Suecia, Finlandia, Bélgica y Lituania, se cumple sólo en la mitad de los casos (Suecia y Finlandia).

Respecto al personal del sector público dedicado a las actividades de I+D+i, los países que presentan mayor porcentaje respecto al total de empleados son Finlandia, Lituania y Reino Unido. La hipótesis vuelve a confirmarse sólo parcialmente (Finlandia y Lituania). Finlandia es el único país que cumple la hipótesis completamente. No es absolutamente necesario ser el país que más invierte en recursos públicos de I+D+i para ser eficiente en el uso de los mismos (Bélgica es eficiente con el nivel de recursos que emplea, no siendo uno de los que más invierte fondos públicos). Si se atiende al total de fondos dedicados a I+D+i (tanto públicos como privados) como porcentaje del PIB, el análisis arroja idénticos resultados.

- ✓ Hipótesis 6 (los países eficientes en el escenario de Educación también son los países eficientes en el escenario Tradicional).

Bélgica, Finlandia y Suecia son los países que son eficientes tanto en el escenario Tradicional como en el de Educación. Por lo tanto, tres de cuatro países cumplen la hipótesis planteada. Irlanda y Eslovaquia son eficientes en el escenario de Educación y son, respectivamente, casi eficiente e ineficiente, en el escenario Tradicional. Excepto por Eslovaquia, se podría concluir que ambos escenarios se encuentran muy relacionados y que vuelve a relucir el destacado papel de la Educación.

- ✓ Hipótesis 7 (los países del Norte de Europa son más eficientes en los distintos escenarios).

Efectivamente, del análisis de los resultados del enfoque 2LS-DEA se obtiene que Suecia, Finlandia y Dinamarca (los tres países nórdicos que pertenecen a la UE) se encuentran entre los más eficientes. Lo que no se puede verificar es las razones de esta eficiencia,

aunque se pueda suponer que las características del clima y su cultura puedan influir en gran medida en su sistema educativo y en el tradicional esfuerzo inversor en I+D+i²⁸.

- ✓ Hipótesis 8 (los países que muestran mejores resultados se repiten en los diversos casos de estudio propuestos dada la interrelación entre los conceptos).

Si se atiende a los grupos en los que se han clasificado los países en cada uno de los análisis, los países que obtienen mejores resultados respecto a las tres perspectivas y que se repiten, por tanto, en los tres análisis son Suecia, Finlandia, Dinamarca y Países Bajos. En el primer análisis (DS) y en el tercero (EC), se encuentran de forma común en los mejores grupos de clasificación Austria, Alemania y Bélgica.

Respecto al análisis 2LS-DEA, se puede comprobar qué países han sido utilizados un mayor número de ocasiones como referencia o *peers* en los distintos escenarios analizados. Éstos son Bélgica, Finlandia y Suecia. En escenarios concretos, los países que se utilizan como *peers* en un mayor número de ocasiones son España (escenario de Innovación), Finlandia (Educación), Italia (Salud) y Grecia (Sociedad de la Información).

Por tanto, la hipótesis de que algunos países podrían repetirse en el mejor de los grupos en los tres análisis se cumple en el caso de Suecia, Finlandia, Dinamarca y Países Bajos.

- ✓ Hipótesis 9 (se pueden detectar patrones homogéneos de comportamiento de los países seleccionados a la hora de clasificarlos en grupos o clases y se pueden construir modelos que pueden predecir dicha clasificación con un alto grado de precisión).

En el caso del estudio de los resultados de I+D+i, en su sentido más amplio y teniendo en cuenta al sector público y al sector privado, se empleó el algoritmo de agrupación *k*-medias obteniendo cuatro grupos de países: de desempeño moderado, orientados a la innovación, de baja innovación y seguidores. En el estudio realizado desde el punto de vista de la economía del conocimiento, se empleó la técnica de agrupamiento jerárquica. Los cuatro grupos fijados fueron economías del conocimiento avanzadas, seguidoras, moderadas y en fase incipiente (ver objetivo tercero).

En este último caso, se incluyeron países adicionales a los países de la UE a efectos comparativos: el resto de países europeos no miembros de la UE, países de la OECD, China, Hong Kong y Túnez). Se encuentra que 9 de los 27 países adicionales

²⁸ Un artículo muy interesante al respecto puede encontrarse en la revista *The Economist* titulado "If in doubt, innovate"; disponible en <http://www.economist.com/news/special-report/21570834-nordic-region-becoming-hothouse-entrepreneurship-if-doubt-innovate>.

seleccionados, (Australia, Canadá, Hong Kong, Islandia, República de Corea, Nueva Zelanda, Noruega, Suiza y Estados Unidos), pertenecen a la primera clase y que 14 de ellos pertenecen a la cuarta clase, es decir, se encuentran muy polarizados respecto a su estado de transición hacia una economía del conocimiento. Por lo que, se confirma la posibilidad de detectar patrones de comportamiento entre los países, incluso de aquellos que no pertenecen a la UE.

La posibilidad de construir modelos que clasifiquen con un alto grado de precisión la clase a la que pertenece cada país, también queda verificada al haberse llevado a cabo el tercer objetivo, empleando una gran variedad de ellos, tanto de tipo nominal como ordinal. Se obtuvieron porcentajes de buena clasificación para los conjuntos de datos de generalización cercanos al 100% en el primer caso y cercanos al 90% en el segundo. Los mejores modelos, como se ha comentado, fueron el ESUNN (para el problema nominal) y el SVORIM (para el problema ordinal).

- ✓ Hipótesis 10 (el concepto de evaluación de la eficiencia puede ser empleado, para el propósito de evaluar la eficiencia de las políticas públicas de I+D+i, desde su sentido más amplio).

Queda comprobada esta hipótesis gracias a los tres análisis empíricos realizados, en los que ha sido posible medir su eficiencia técnica relativa mediante la técnica no paramétrica DEA, como ser clasificados en grupos y predecir su clasificación a través de clasificadores procedentes del campo de la Inteligencia Artificial y el Aprendizaje Automático. En el primer caso, fue un análisis de eficiencia en sentido estricto. En el caso de los clasificadores, el hecho de analizar la pertenencia a un grupo y, mediante el correspondiente clasificador, predecir su pertenencia a una u otra clase, es una forma de evaluar la situación de un país respecto a las perspectivas planteadas, y de realizar un seguimiento. De esta forma, si cambian los valores de las variables seleccionadas, cambiaría la clasificación del grupo en cuestión. Dependiendo de la nueva clase asignada por el clasificador, se podría decir si ha mejorado o no respecto a su situación (clasificación) anterior y se podrían detectar las variables más relevantes en dicho cambio.

- ✓ Hipótesis 11 (posibilidad de abordar el estudio mediante técnicas de clasificación nominales y ordinales, comprobando la naturaleza de los problemas analizados).

En el caso del análisis de los resultados de I+D+i de los países de la UE, del análisis de las características de cada uno de los grupos y del valor de las variables empleadas, se

concluyó la naturaleza nominal de los mismos. Entre la clase 1 (desempeño moderado en innovación) y la clase tres (baja innovación), no queda clara la relación de orden para los datos analizados. Así, en la clase 1, tienen un buen nivel de estudios terciarios, pero no destacan en el aprendizaje a lo largo de la vida y es el grupo que presenta menor número de doctores. No son grandes inversores en I+D+i ni el Gobierno ni el sector privado, y cuentan con menos de la media de recursos humanos en actividades de ciencia y tecnología así como en personal investigador y de desarrollo. Sin embargo, son muy productivos a la hora de crear marcas comerciales (no patentes) y en los beneficios procedentes de las mismas. Rentabilizan su reducida producción de patentes, el estar en la media de exportaciones de productos de media y alta tecnología. Presentan un valor menor a la media en producción científica (lo cual puede ser lógico al tener la menor proporción de doctores). Económicamente, es el grupo en el que menos crece la productividad, cuenta un con bajo crecimiento del PIB y un bajo nivel de empleo.

La clase 3, a pesar de que cuenta con niveles muy bajos en casi todas las variables analizadas (pocos doctores, la menor población de estudios de nivel superior y menos aprendizaje a lo largo de la vida, bajos niveles de recursos humanos y de gasto del Gobierno y las empresas en I+D+i y son los que peores resultados muestran en cuanto a marcas comerciales, exportaciones de alta y media tecnología, número de patentes, ingresos por patentes y número de publicaciones por 1000 habitantes), sin embargo, económicamente, son altamente productivos y presentan el porcentaje de crecimiento del PIB más elevado.

Si el fin último de la inversión en I+D+i es el crecimiento económico que redunde en el bienestar general de la sociedad, por éste último hecho, no se puede afirmar categóricamente que la clase 1 preceda en orden a la clase 3, al menos con los datos del periodo empleado.

Además de los resultados expuestos, que demuestran la consecución de los objetivos y la verificación (o no) de las hipótesis planteadas, se pueden realizar otras consideraciones adicionales y resaltar algunas implicaciones y utilidades relacionadas con las contribuciones de esta investigación, como son las siguientes:

1. Tanto el modelo conceptual propuesto como la clasificación de los países en función de los tres conceptos expuestos, pretende ser de utilidad e interés en: i) la gestión de los recursos

públicos/privados en Investigación, Desarrollo e Innovación (I+D+i); ii) el diseño de políticas macroeconómicas de incentivación, iii) la monitorización o seguimiento de los avances realizados; iv) la realización de tareas de *benchmarking* entre países, con la adopción de las mejores prácticas detectadas entre ellos.

2. Los diversos escenarios socioeconómicos de eficiencia presentados pueden contribuir a la discusión sobre la orientación de las políticas y estrategias europeas. La iniciativa Europa 2020, El Espacio Europeo de Investigación y la inclusión de los indicadores relacionados de Eurostat para monitorizar su evolución, revelan la importancia de alcanzar una sociedad basada en el conocimiento y un desarrollo sostenible en la UE. Si no se consiguen estos objetivos, podría plantearse el giro hacia otro tipo de estrategias alternativas.

3. Los países de la UE deberían continuar apoyando los recursos dedicados a la educación, además de los fondos dedicados a la I+D+i, dado que estas habilidades son esenciales en la aplicación de nuevas tecnologías industriales y de los servicios basados en el conocimiento que conducen al crecimiento económico, gracias a la mayor productividad lograda a través de la actividad innovadora.

4. Las políticas de I+D+i europeas tienen aún un gran camino que recorrer en el uso eficiente de los recursos seleccionados, al margen de los tres países más eficientes detectados que, por otra parte, tienen una amplia tradición en la inversión en I+D+i y cuentan con unos sistemas educativos de gran calidad. Estos países, precisamente, se encuentran entre los primeros puestos del índice de competitividad que elabora el WEF (2012). Suecia, por ejemplo, se encuentra en tercer lugar en el informe correspondiente a 2011/2012 tras haber ocupado el segundo puesto en el informe a los años 2010/2011.

5. Si el surgimiento de una infraestructura específica para un sistema de política de innovación europea pudiera ser trazado, la monitorización es también una tarea relevante que arrojaría luz en el debate de lo que sería mejor para Europa: una política de innovación centralizada o una competición abierta de iniciativas y sistemas de innovación, regionales o internacionales, dependiendo de la evolución de los países de Europa hacia una mejor clasificación en su desempeño de I+D en el clúster correspondiente, una vez nuevos datos estén disponibles o nuevos Estados pasen a formar parte de la Unión Europea.

6. La Economía del Conocimiento (EC) se ha convertido en la principal tendencia en la sociedad internacional y en este globalmente competitivo siglo XXI. La exploración de las características de la EC y establecer un paradigma económico apropiado para acelerar la

innovación tecnológica que favorezca el crecimiento económico, son tareas prioritarias para los gobiernos. En este contexto, los principales grupos de interés necesitan herramientas útiles para el apoyo a la toma de decisiones, como representan el variado catálogo de indicadores compuestos facilitados por instituciones como la OECD, el WB o la UE. Por ello, las técnicas propuestas en este trabajo pretenden aunarse a dichos intentos de medición y seguimiento.

7. La monitorización es una tarea relevante para arrojar luz sobre el progreso de los países hacia las metas propuestas. La clasificación de estos países de acuerdo a sus avances hacia la EC (basada en los cuatro pilares utilizados), su eficiencia en función de los escenarios de DS y de sus resultados de I+D+i, puede ser también útil para evaluar su posición comparándose con otros países con respecto a alcanzar los objetivos de competitividad y sostenibilidad en el largo plazo. Los clústeres obtenidos y los modelos de clasificación nominales y ordinales reflejan una imagen global del estado de los países, lo que podría enriquecer y complementar el juicio de los grupos de interés, más que aportando un único valor cuantitativo de un indicador o de índices compuestos.

8. Por último, conviene insistir en la necesidad de diseñar y ejecutar actuaciones que permitan contribuir a un importante cambio cultural en la sociedad, no sólo desde el punto de vista de la eficiencia del uso de los recursos, sino también con el objetivo de que pueda mejorar el reconocimiento social tanto de los investigadores que generan nuevo conocimiento científico y tecnológico, como de los emprendedores. La conjunción de ambos hacen posible la puesta en marcha de nuevos proyectos empresariales, aplicando al sistema productivo el nuevo conocimiento (lo que hace necesario potenciar, así mismo, los sistemas de transferencia del conocimiento) y contribuyendo con ello al crecimiento económico y a la creación de empleo. El modelo elaborado, que contiene una variable representativa de la cultura empresarial del país, pretende potenciar el diseño de políticas que contribuyan a lo expuesto.

3. LÍNEAS FUTURAS DE INVESTIGACIÓN

Al culminar la presente investigación se cierra un trabajo, pero lo planteado permite abrir muchos más. En realidad, el objetivo de investigar no consiste solo en los resultados alcanzados en sí mismos, sino en las líneas de investigación que se abren para el futuro. El abanico de posibilidades que nos aporta la presente tesis nos permite formular una serie de cuestiones que se espera abordar en el futuro.

Entre las futuras investigaciones que se comentan, se podrían destacar de forma especial las siguientes:

1. Puesto que el modelo diseñado cumple la condición necesaria y suficiente para ser considerado una red Bayesiana, se podrían llevar a cabo simulaciones Monte Carlo para evaluar los efectos de inversiones alternativas (modificando los valores de los indicadores clave) y estudiar los efectos, en términos de probabilidades, sobre el crecimiento económico.
2. También podrían resolverse escenarios o partes *d-separadas* del modelo de red Bayesiana construyendo las distribuciones de probabilidad condicionada correspondientes, a través de los datos o de consultas a expertos.
3. A medida que los datos necesarios se encuentren disponibles, las implicaciones de nuestro enfoque se tornarían útiles para futuras investigaciones comparativas de los resultados entre países. Adicionalmente, sería de interés ampliar la investigación con datos desde el año 2009 ya que el periodo analizado en el presente estudio podría no recoger completamente el impacto de la crisis económica y financiera, tanto por el retraso en que los impactos económicos se reflejen en las magnitudes analizadas como por la disponibilidad de los datos (sobre todo, en lo relativo a las variables de educación).
4. Dado que en este trabajo se ha consultado a expertos en I+D+i, no se ha llegado a profundizar en la interacción con potenciales usuarios finales del modelo. Por ello, una línea de investigación podría consistir en definir distintos tipos de usuarios, dependiendo del conocimiento sobre la materia, y obtener un *feed-back* de las utilidades que les reporta el modelo o de sus deficiencias.
5. Simplificación de la red. A pesar de que nuestro grafo pretende ser útil para generar visiones intuitivas de las relaciones entre las variables implicadas, en el caso en que la red recoge un gran número de variables, esto puede restar claridad a su interpretación. Por añadidura, el tiempo necesario para ejecutar los algoritmos de resolución de la red crece exponencialmente con el tamaño de la red, máxime si las variables aleatorias seleccionadas son continuas.
6. Ante la dificultad de diseñar redes Bayesianas manualmente, con la ayuda de expertos y la consulta de la bibliografía correspondiente, sería de gran utilidad contar con una interfaz de apoyo para dichos expertos, donde pudieran combinarse datos y opiniones.

7. Selección de la muestra de países. Podría estudiarse la validez del modelo diseñado para que sea aplicado a otros países con distintos niveles de desarrollo económico o de distinta localización geográfica. El estudio podría distinguir entre modificar las relaciones de causalidad y los nodos de la red o escoger indicadores de medición alternativos.
8. La naturaleza nominal u ordinal del problema podría también estudiarse en el caso de los resultados por países de la I+D+i y de un desarrollo sostenible. En este último caso, se podrían aplicar además, como alternativa al método DEA, las técnicas de clasificación propuestas en los otros dos análisis empíricos.
9. El modelo SVORIM fue el que mejor se comportó para el caso del estudio de los países respecto a su grado de consecución de economía del conocimiento. Por tanto, se podría realizar un estudio de los países seleccionados como más importantes por dicho modelo (es decir, aquellos seleccionados como vectores soporte).
10. La elección de las variables seleccionadas podría sesgar los resultados en el agrupamiento de países realizado y en los métodos de clasificación, dado que no dejan de ser medidas indirectas de lo que realmente se quiere medir. Se propone la selección de variables alternativas, adicionales y/o combinación de ellas para cada pilar para más investigaciones.

Para concluir, los tópicos analizados en esta tesis doctoral se han mostrado relevantes, ya que existen muchos grupos de interés implicados e interesados en ellos, como las empresas, los académicos y los decisores políticos, con una complicada coyuntura económica como telón de fondo. Por este motivo, sería interesante continuar investigando en esta línea, también a nivel microeconómico, con la perspectiva planteada en el modelo conceptual causal pero con datos de empresas de varios países. Si se continúa investigando en estos temas, podría contarse con mayor evidencia empírica, lo cual permitirá a los diferentes agentes conocer y entender mejor la naturaleza del problema y tomar mejores decisiones, y para lo que esta tesis pretende ser el primer paso de un arduo camino por recorrer.

REFERENCIAS

- Abramovitz, M., 1986. Catching up, forging ahead, and falling behind. *Journal of Economic History* 46, 385–406.
- Abramowitz, M., 1989. *Thinking about Growth*. Cambridge University Press, Cambridge.
- Acemoglu, D., 2009. Introduction to modern economic growth. *Introduction to modern economic growth*, forthcoming. Princeton University Press, Princeton (NJ).
- Afuah, A., 1999. *La dinámica de la innovación organizacional*. Oxford University Press, México.
- Agazzi, E., 1998. From technique to Technology: The rol of modern science. *Society for Philosophy and Tecnology* 4, (disponible en <http://scholar.lib.vt.edu/ejournals/SPT/v4n2/AGAZZI.html>).
- Agénor, P. R., 2000. *The Economics of Adjustment and Growth*. Academic Press, New York.
- Aghion, P., 2012. Growth policy and the State. Working paper 3, *The Growth Dialogue*, Washington, D.C.
- Aghion, P., Howitt, P., 1997. *Endogenous Growth Theory*. MIT Press, Cambridge, MA.
- Aguado, I., Echebarria, C., Barrutia, J.M., 2009. El desarrollo sostenible a lo largo de la historia del pensamiento económico. *Revista de economía mundial* 21, 87–110.
- Aguilera, P.A., Fernández, A., Fernández, R., Rumí, R., Salmerón, A., 2011. Bayesian networks in environmental modelling. *Environmental Modelling & Software* 26, 1376–1388.
- Al Shami, A., Lotfi, A., Coleman, S., 2012. Unified Knowledge Based Economy Neural Forecasting Map. WCCI 2012 IEEE World Congress on Computational Intelligence (IJCNN), Brisbane, Australia.
- Anderson, R.D., Vastag, G., 2004. Causal modeling alternatives in operations research: Overview and application. *European Journal of Operational Research* 156, 92–109.
- Archibald, R. B., Pereira, A. M., 2003. Effects of public and private R&D on private-sector performance in the United States. *Public Finance Review* 31, 429–451.
- Archibugi, D., A. Coco, A., 2005. Measuring technological capabilities at the country level: a survey and a menu for choice. *Research Policy* 34, 175–194.
- Arrow, K. J., 1962. The Economic Implications of Learning By Doing. *Review of Economics Studies* XXIX, 155–173.
- Arrow, K., 1979. El bienestar económico y la asignación de recursos para la invención, en: Rosenberg, N. (Eds), *Economía del Cambio Tecnológico*. Fondo de Cultura Económica, México.
- Auluck, R., 2002. Benchmarking: A tool for facilitating organizational learning? *Public Administration and Development* 22, 109–122.
- Balconi, M., Brusoni, S., Orsenigo, L., 2010. In defence of the linear model: An essay. *Research Policy* 39, 1–13.
- Banker, R.D., Charnes, A., Cooper, W.W., 1984. Some Models for Estimating Technical and Scale Inefficiencies in Data Envelopment Analysis. *Management Science* 30, 1078–1092.
- Baskanran, A., Muchie, M., 2006. *Bridging the Digital Divide: Innovation Systems for ICT in Brazil, China, Thailand and Southern Africa*. Adonis and Abbey, London.

- Baumol, W. J., 2003. Innovations and growth: Two common misapprehensions. *Journal of Policy Modeling* 25, 435–444.
- Becker, J., Prischmann, M., 1993. Supporting the design process with neural networks. A complex application of cooperating neural networks and its implementation. *Journal of Information Science Technology* 3, 79–95.
- Bishop, C. M., 2006. *Pattern Recognition and Machine Learning*. Springer, New York.
- Bishop, M., 1995. *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford.
- Block, J.H., Thurik, R., Zhou, H., 2012. What turns knowledge into innovative products? The role of entrepreneurship and knowledge spillovers. *Journal of Evolutionary Economics*, DOI: 10.1007/s00191-012-0265-5.
- Bode, J., 1998. Decision support with neural networks in the management of research and development: Concepts and application to cost estimation. *Information Management* 34, 33–40.
- Boettke, Peter J., Coyne, Christopher J., Leeson, Peter T., 2008. Institutional Stickiness and the New Development Economics. *American Journal of Economics and Sociology* 67, 1536–7150.
- Bonaccorsi, A., Daraio, C., Simar, L., 2006. Advanced indicators of productivity of universities. An application of robust nonparametric methods to Italian data. *Scientometrics* 66, 389–410.
- Bontis, N., Dragonetti, N.C., Jacobsen, K., Roos, G., 1999. The knowledge toolbox: a review of the tools available to measure and manage intangible resources. *European Management Journal* 17, 391–402.
- Boose, J.H., 1989. A survey of knowledge acquisition techniques and tools. *Knowledge Acquisition* 1, 3–37.
- Booyen, F., 2002. An Overview and Evaluation of Composite Indices of Development. *Social Indicators Research* 59, 115–151.
- Borsuk M.E., Stow C.A., Reckhow K.H., 2004. A Bayesian network of eutrophication models for synthesis, prediction, and uncertainty analysis. *Ecological Modelling* 173, 219–239.
- Boulanger, P.-M., Brochet, T., 2005. Models for policy-making in sustainable development: The state of the art and perspectives for research. *Ecological Economics* 55, 337–350.
- Bound, J., Cumins, C., Griliches, Z., Hall, H.H., Jaffe, A., 1984. *R&D, Patent, and Productivity*. University of Chicago Press, Chicago.
- Braun, E., 1986. *Tecnología rebelde*, Tecnos/Fundesco, Madrid.
- Bresnahan, T.F., Trajtenberg, M., 1995. General purpose technologies ‘Engines of growth?’, *Journal of Econometrics* 65, 83–108.
- Brundtland, G.H., 1987. *Informe de la Comisión Mundial sobre el Medio Ambiente y el Desarrollo (Comisión Brundtland): Nuestro Futuro Común*. Comisión Brundtland, New York.
- Buck, S.F., 1960. A method of estimation of missing values in multivariate data suitable for use with an electronic computer, *Journal of Royal Statistics Society. Series B (Methodology)* 22, 302–306.
- Busch, E. D., Trexler, J.C., 2003. The importance of monitoring in regional ecosystem initiatives, en: Busch, E. D., Trexler, J. C. (Eds), *Monitoring ecosystems: Interdisciplinary approaches for evaluating ecoregional initiatives*. Island Press, Washington, D.C.

- Carayannis, E.G., Campbell, D.F.J., 2010. Triple Helix, Quadruple Helix and Quintuple Helix and how do knowledge, innovation and the environment relate to each other? A proposed framework for a trans-disciplinary analysis of sustainable development and social ecology. *International Journal of Social Ecology and Sustainable Development* 1, 41–69.
- Cardoso, J.S., Da Costa, J.F., 2007. Learning to Classify Ordinal Data: The Data Replication Method. *Journal of Machine Learning Research* 8, 1393–1429.
- Cerulli, G., A. Filippetti, A., 2012. The complementary nature of technological capabilities: Measurement and robustness issues. *Technological Forecasting Social Change* 79, 875–887.
- Charalambous, C., Charitou, A., Kaourou, F., 2000. Comparative Analysis of Artificial Neural Network Models: Application in Bankruptcy Prediction, *Annals of Operation Research* 99 403–425.
- Charniak, E., 1991. Bayesian networks without tears. *Airsoft International Magazine* 12 , 50–63.
- Chelimsky, E., 1998. The role of Experience in Formulating Theories of Evaluation Practice. *American Journal of Evaluation* 19, 35–55.
- Chen Derek, H.C., Dahlman C. J., 2006. The Knowledge Economy, The Kam Methodology and World Bank Operations. The International Bank for Reconstruction and Development/The World Bank, Washington.
- Chiesa, V., Frattini, F., Lazzarotti, V., Manzini, R., 2009. Performance measurement in R&D: exploring the interplay between measurement objectives, dimensions of performance and contextual factors. *R&D Management* 39, 487–519.
- Chu, W., Keerthi, S. S., 2007. Support vector ordinal regression. *Neural Computation* 19, 792–815.
- Chu, W., Keerthi, S.S., 2005. New Approaches to Support Vector Ordinal Regression, en: *Proceedings of the 22nd International Conference of Machine Learning (ICML'05)*. Association for Computer Machinery, New York, pp. 145–152.
- Cohen, J., 1960. A coefficient of agreement for nominal scales, *Educational and Psychological Measurement* 20, 37–46.
- Comisión Europea, 2007. COM(2007) 161 final. Bruselas 4.4.2007, Libro Verde, El Espacio Europeo de Investigación: nuevas perspectivas.
- Comisión Europea, 2010. Europe 2020 – a strategy for smart, sustainable and inclusive growth. COM (2010) 2020. Brussels.
- Comisión Europea, PRO INNO Europe, 2011. <http://www.proinno-europe.eu/>, sitio web oficial (último acceso, Febrero 2012).
- Conte, A., Schweizer, P., Dierx A. and Ilzkovitz, F., 2009. An Analysis of the Efficiency of Public Spending and National Policies in the Area of R&D. *European Economy. Occasional Papers*, 54. European Commission, Brussels.
- Cooper, W.W., Seiford, L.M., Tone, K., 2007. *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA–Solver Software* (2^a ed.). Springer, New York.
- Cooper, W.W., Seiford, L.M., Zhu, J., 2004. Data envelopment analysis: history, models and interpretations, en: Cooper, W.W., Seiford, L.M., Zhu, J.(Eds), *Handbook on Data Envelopment Analysis*. Kluwer Academic Publishers, Boston, 1–45.
- Cortes, C., Vapnik, V., 1995. Support-vector networks. *Machine Learning* 20, 273–297.

- Cover T.M., Hart P.E., 1967. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory* 13, 21–27.
- Cowan, R., David, P. and Foray, D., 2000. The explicit economics of codification and tacitness, *Industrial and Corporate Change* 9, 211–53.
- Cozzens, S.E., Kaplinsky, R., 2009. Innovation, poverty and inequality: cause, coincidence, or co-evolution?, en: Lundvall, B.-A., Joseph, J.K., Chaminade, C., Vang, J. (Eds), *Handbook of Innovation Systems and Developing Countries. Building Domestic Capabilities in a Global Context*. Edward Elgar, Cheltenham.
- Darwiche, A., 2010. What are Bayesian networks and why are their applications growing across all fields? *Communications of the Association for Computing Machinery* 53, 80–90.
- David, P., Foray, D., 1995. Accessing and Expanding the Science and Technology Knowledge Base. *STI Review* 16, 16-38.
- David, P.A., Hall, B. H., 2000. Heart of Darkness: Modeling Public–Private Funding Interactions Inside the R&D Black Box. *Research Policy* 29, 1165–1183.
- De Smet Y., Nemery, P., Selvaraj, R., 2012. An exact algorithm for the multicriteria ordered clustering problem. *Omega* 40, 861–869
- De Smet, Y., 2004. Towards multicriteria clustering: An extension of the k–means algorithm, *Eur. J. Oper. Res.* 158(2) (2004) 390–398.
- Degtyarev, A., 2004. Policy Analysis as Applied Discipline. *Political Studies* 1, 56–68.
- Denison, E., 1985. Trends in American economic growth, 1929–82. Brookings Institution, Washington.
- Díez, F.J., 1999. From causal graphs to Bayesian networks. En *Proceedings of the Workshop on Conditional Independence Structures and Graphical Models*, Toronto, Canadá.
- Diialiwal, J.S., Benbasat, I., 1990. A framework for the comparative evaluation of knowledge acquisition tools and techniques. *Knowledge Acquisition* 2, 145–166.
- Domar, E., 1946. Capital Expansion, Rate of Growth and Employment. *Econometrica* 14, 13-147.
- Doumpos, M., Kosmidou, K., Baourakis, G., Zopounidis, C., 2002. Credit risk assessment using a multicriteria hierarchical discrimination approach: A comparative analysis. *European Journal of Operational Research* 138, 392–412.
- Drucker, P. F., 1994. *La Sociedad Postcapitalista*. Editorial Norma, Bogotá.
- Drucker, P., 1969. *The Age of Discontinuity*. Harper and Row, New York.
- Druzdzel, M., 1996. Qualitative verbal explanations in Bayesian belief networks. *Artificial Intelligence and simulation of behaviour Quarterly*, special issue on Bayesian networks 94, 43-54.
- Duc, D., Umemoto, K., 2009. Modeling the development toward the knowledge economy: a national capability approach. *Journal of Knowledge Management* 13, 359–372.
- Dunn W.N., 1994. *Public policy analysis. An introduction.*, Prentice- Hall. Englewood Cliff, NJ.
- Dunn, W.N., 1981. *Public Policy Analysis: An introduction*. Prentice–Hall, Englewood Cliffs, NJ.

- Durbin, R., Rumelhart, D., 1989. Products Units: A computationally powerful and biologically plausible extension to backpropagation networks. *Neural Computation* 1, 133–142.
- Eden, C., Ackermann, F., Cropper, S., 1992. The analysis of cause maps. *Journal of Management Studies* 29, 309–324.
- Ehrlich, I., 1990. The Problem of Development: Introduction. *Journal of Political Economy* 98, S1–S10.
- Erbes, A., Robert, V., Yoguel, G., Borello, J., Lebedinsky, V., 2006. Regímenes tecnológico, de conocimiento y competencia en diferentes formas organizacionales: la dinámica entre difusión y apropiación. *Revista Desarrollo Económico* 46, 33-61.
- 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, 109–23.
- Fagerberg, J., Landström, H., Martind, B.R., 2012. Exploring the emerging knowledge base of ‘the knowledge society’. *Research Policy* 41, 1121–1131.
- Fagerberg, J., Verspagen, B., 2009. Innovation studies - the emerging structure of a new scientific field. *Research Policy* 38, 218–233.
- Fahrenkrog, G., Polt, W., Rojo, J., Tübke, A., Zinöcker, K., 2002. RTD Evaluation toolbox, assessing the socioeconomic impact of RTD policies. Joint Research Centre, Institute for Prospective Technological Studies, Seville.
- Farrell, M.J., 1957. The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society. Series A (General)* 120, 253–290.
- Filippettia, A., Archibugia, D., 2011. Innovation in times of crisis: National Systems of Innovation, structure, and demand. *Research Policy* 40, 179–192.
- Fogel, L.J., Owens, A.J., Walsh, M.J., 1966. *Artificial Intelligence through Simulated Evolution*. John Wiley, New York.
- Fouad, S., Tino, P., 2012. Adaptive Metric Learning Vector Quantization for Ordinal Classification. *Neural Computation* 24, 2825–2851.
- Frank, E., Hall, M., 2001. A simple approach to ordinal classification, en: De Raedt, L. Flach P. (Eds), *Proceedings of the Twelfth European Conference on Machine Learning*. Springer–Verlag, Berlin Heidelberg, pp. 145–156.
- Freeman, C., 1987. *Technology and Economic Performance: Lessons from Japan*. Pinter Publishers, London.
- Freeman, C., Soete, L., 2009. Developing science, technology and innovation indicators: What we can learn from the past. *Research Policy* 38, 583–589.
- Friedman, J., Hastie, T., Tibshirani, R., 2000. Additive logistic regression: a statistical view of boosting. *Annals of Statistics* 38, 337–374.
- Fritz–Vannahme, J., García Schmidt, A., Hierlemann, D., Vehrkamp, R., 2010. Lisboa, la segunda. *Spotlight Europe 2010/02*, Fundación Bertelsmann Stiftun.
- Furman, J.L., Porter, M.E., Stern, S., 2002. The determinants of national innovative capacity. *Research Policy* 31, 899–933.

- Gadrey, J., Jany-Catrice, F., 2006. *The New Indicators of Well-Being and Development*, Palgrave, Houndmills.
- Galindo, M.A., 2011. Crecimiento Económico, *Revista de Economía ICE* 858, 39–55.
- García-Alonso, C.R., Torres-Jiménez, M., Hervás-Martínez, C., 2009. Expert-Oriented Monte-Carlo DEA: Technical Efficiency Assessment of the Agrarian Sector, en: *Proceedings of the 23 European Conference on Operational Research*, Bonn, 242–242.
- García-Alonso, C.R., Campoy-Muñoz, P., Salazar Ordoñez, M., 2013. A Multi-Objective Evolutionary Algorithm for enhancing Bayesian Networks hybrid-based modeling. *Computers & Mathematics with Applications* (disponible en línea desde el 9 de febrero de 2013; DOI: 10.1016/j.camwa.2013.01.029).
- Gemmell, N., 1995. Endogenous growth, the Solow model and human capital, *Economics of Planning* 28 (2-3), 169–183.
- Gibert, K., García-Alonso, C., Salvador-Carulla, L., 2010. Integrating clinicians, knowledge and data: expertised cooperative analysis in health care decision support. *Health Research Policy and Systems* 8, 28–44.
- Gibert, K., Sonicki, Z., 1999. Clustering based on rules and medical research. *Applied Stochastic Models in Business and Industry, formerly JASMDA* 15, 319–324.
- Godin, B., 2002. Technological gaps: an important episode in the construction of S&T statistics. *Technology in Society* 24, 387–413.
- Godin, B., 2007. Science, accounting and statistics: The input-output framework. *Research Policy* 36, 1388–1403.
- Goldberg, D.E., 2000. The design of innovation: lessons from genetic algorithms, lessons for the real World. *Technological Forecasting and Social Change* 64, 7–12.
- Goldin, I., Winters, L.A., 1995. Economics policies for sustainable development, en: Goldin I., Winters, L.A. (Eds), *The Economics of Sustainable Development*. University Press, Cambridge, pp. 1-15.
- Greif, S., 1985. Relation between R&D expenditure and patent applications. *World Patent Information* 7, 190–195.
- Griliches, Z., 1990. Patent statistics as economic indicators: a survey. *Journal of Economic Literature* 28, 1661–1707.
- Griliches, Z., 1992. The search for R&D spillovers. *Scandinavian Journal of Economics* 94, 29–47.
- Griliches, Z., 1998. *R&D and Productivity: The Econometric Evidence*. The University of Chicago Press, Chicago Press.
- Gross, J.E. 2003. Developing conceptual models for monitoring programs. Available at http://science.nature.nps.gov/im/monitor/docs/Conceptual_Modelling.pdf.
- Gruca, T.S. , Klemz, B.R., 1998. Using Neural Networks to Identify Competitive Market Structures from Aggregate Market Response Data. *Omega, International Journal of Management Science* 26, 49–62.
- Grupp, H., Schubert, T., 2010. Review and new evidence on composite innovation indicators for evaluating national performance. *Research Policy* 39, 67–78.

- Guan, J. C., Wang, J. X., 2004. Evaluation and interpretation of knowledge production efficiency. *Scientometrics* 59, 131-155.
- Gupta, S., Kim, H.W., 2008. Linking structural equation modeling to Bayesian networks: Decision Support for customer retention in virtual communities. *European Journal of Operational Research* 190, 818–833.
- Gutiérrez, P.A., Segovia–Vargas, M. J., Salcedo–Sanz, S., Hervás–Martínez, C., Sanchís, A., Portilla–Figueras, J. A., Fernández–Navarro, F., 2010. Hybridizing logistic regression with product unit and RBF networks for accurate detection and prediction of banking crises. *Omega, International Journal of Management* 38, 333–344.
- Hájková, V., Hájek, P., 2010. Analysis of Regional Innovation Systems by Neural Networks and Cluster Analysis, *Proc. Communication and Management in Technological Innovation and Academic Globalization (COMATIA '10)*, 46–51.
- Hall, B.H., R.H. Ziedonis, 2001. The Patent Paradox Revisited: an Empirical Study of Patenting in the U.S. Semiconductor Industry. *RAND Journal of Economics* 32, 101–128.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I., 2009. The WEKA Data Mining Software: An Update, *ACM SIGKDD Explorations News* 1(11), 10–18.
- Hanley, S., Malafsky, G., 2002. A Guide for Measuring the Value of Km Investments, en: Holsapple, C.W. (Eds), *Handbook on Knowledge Management 2*. Springer–Verlag, Berlin, Heidelberg, pp. 369–394.
- Hardi, P., Zdan, T., 1997. *Assessing Sustainable Development. Principles in Practice*. The International Institute for Sustainable Development, Winnipeg
- Harrod, R. F., 1939. *La Teoría Dinámica*, en Sen, Amartya. (1970, 1979). *Economía del Crecimiento*. FCE, pp. 43–62
- Hartigan, J. A., Wong, M. A., Algorithm AS 136: a k–means clustering algorithm, *Journal of Royal Statistics Society C–APP* 28 (1979) 100–108.
- Hartmann, D., Pyka, A., 2013. Innovation, economic diversification and human development. Discussion Paper 65. Center for Research on Innovation and Services, University of Hohenheim.
- Hastie, T., Tibshirani, R. y Friedman, J. H.: 2001, *The Elements of Statistical Learning*. Springer, New York.
- Hausman, A., Stock, J.R., 2003. Adoption and implementation of technological innovations within long–term relationships, *Journal of Business Research* 56, 681–686.
- Haykin, S., 1994. *Neural Networks: A comprehensive Foundation*, MacMillan, New York.
- Heckerman, D., 1996. Bayesian networks for data mining. *Data Mining and Knowledge Discovery* 1, 79–119.
- Helpman, E., 1998. *General Purpose Technologies and Economic Growth*. Massachusetts Institute of Technology, USA.
- Herbrich, R., Graepel, T., Obermayer, K., 1999. Support vector learning for ordinal regression, en: Ninth international conference on artificial neural networks, ICANN 1, 97–102.
- Hervás–Martínez, C., Martínez Estudillo, F.J., 2007. Logistic regression using covariates obtained by product–unit neural network models. *Pattern Recognition* 40, 52–64.

- Hirsch–Kreinsen, H., Jacobson, D., Laestadius, S., Smith, K., 2003. Low–tech industries and the knowledge economy: state of the art and research challenges. Working Paper 1, University of Dortmund.
- Holland, H., 1975. *Adaptation in natural and artificial systems*. The University of Michigan Press, Ann Arbor.
- Holland, J., 1975. *Adaptation in Natural and Artificial Systems*, The University of Michigan Press, Ann Arbor, MI:
- Hsu, C.W., Lin, C.J., 2002. A comparison of methods for multi-class support vector machines. *IEEE Transaction on Neural Networks* 13, 415–425.
- Hsu, G.J.Y., Lin, Y.–H., Wei, Z.–Y., 2008. Competition policy for technological innovation in an era of knowledge–based economy. *Knowledge Based Systems* 21, 826–832.
- Jensen, F.V., 2001. *Bayesian Networks and Decision Graphs*. Springer–Verlag, New York.
- Jensen, M.B., Johnson, B., Lorenz, E., Lundvall, B.A., 2007. Forms of knowledge and modes of innovation. *Research Policy* 36, 680–693.
- Johannessen, J.A., Olsen, B., 2010. The future of value creation and innovations: Aspects of a theory of value creation and innovation in a global knowledge economy. *International Journal of Information Management* 30, 502–511.
- Johnson, B., Lundvall, B.–A., Edquist., C., 2003. *Economic Development and the National System of Innovation Approach*. First Globelics Conference, Rio de Janeiro.
- Joint Committee on Standards for Educational Evaluation, 1994. *The Program Evaluation Standards (2nd ed.)*. How to Assess Evaluations of Educational Programs. Thousand Oaks, Sage.
- Jones, C.I., Williams, J.C., 1998. Measuring the social return to R&D. *The Quarterly Journal of Economics* 113, 1119–1135.
- Jørgensen, S.E., 1986. Structural dynamic model. *Ecological Modelling* 13, 1–9.
- Kerssens–van Drongelen, I.C., A. Cook, A., 1997. Design Principles for the Development of Measurement Systems for Research and Development Processes, *R&D Management* 27, 345–357.
- Keynes, J.M., 1936. *The General Theory of Employment, Interest and Money*, Macmillan, London.
- Kim, H., Park, Y., 2008. The impact of R&D collaboration on innovative performance in Korea: A Bayesian network approach. *Scientometrics* 75, 535–554.
- Kline, S.J., Rosenberg, N., 1986. An overview of innovation, en: Landau, R., Rosenberg, N. (Eds), *The Positive Sum Strategy: Harnessing Technology for Economic Growth*. National Academy Press, Washington, D.C., pp. 275–305.
- Kolodko, G.W., 2001. The 'New Economy' and the Old Problems. *Prospects for Fast Growth in Postsocialist Countries*. *Zagreb International Review of Economics & Business* 4, 71–90.
- Korhonen, P., Tainio, R., Wallenius, J., 2001. Value efficiency analysis of academic research. *European Journal of Operational Research* 130, 121–132.
- Krugman, P., 1993. Toward a Counter–Counterrevolution, en: *Development Theory*. Proceedings of the World Bank Annual Conference on Development Economics. The World Bank, Washington, D.C., 15–38.

- Kurtz, J. C., Jackson, L. E., Fisher, W. S., 2001. Strategies for evaluating indicators based on guidelines from the Environmental Protection Agency's Office of Research and Development. *Ecological Indicators* 1, 49–60.
- Lamirel, J. C., Al Shehabi, S., Francois, C., Polanco, X., 2004. Using a compound approach based on elaborated neural network for Webometrics: An example issued from the EICSTES Project. *Scientometrics* 61, 427–441.
- Landwehr, N., Hall, M., Frank, E., 2005. Logistic model trees. *Machine Learning* 59, 161–205.
- Laskey, K.B., Mahoney, S.M., 1997. Network fragments: representing knowledge for constructing probabilistic models, en: D. Geiger, P.P. Shenoy (Eds), *Uncertainty in Artificial Intelligence: Proceedings of the Thirteenth Conference*. Morgan Kaufmann, San Mateo, CA.
- Lazhano, E., Sierra, B., Astigarraga, A., Martínez–Otzeta, J.M., 2007. On the use of Bayesian Networks to develop behaviours for mobile robots. *Robotics and Autonomous Systems* 55, 253–265.
- Le Cessie, S. and van Houwelingen, J., 1992. Ridge estimators in logistic regression. *Applied Statistics* 41, 191–201.
- Leydesdorff, L., 2006. *The Knowledge–based Economy: Modeled, Measured, Simulated*. Universal Publishers, Boca Raton, FL.
- Leydesdorff, L., Wagner, C., 2009. Macro–level indicators of the relations between research funding and research output, *Journal of Econometrics* 3, 353–362.
- Li, L., Lin, H., 2007. Ordinal regression by extended binary classification. *Advances in neural information processing systems* 19, 865–872.
- Lin, H.–Y., 2013. Feature selection based on cluster and variability analyses for ordinal multi–class classification problems. *Knowledge Based System* 37, 94–104.
- Lippmann, R. P., 1996, *Neural Networks, Bayesian a posteriori probabilities, and pattern classification*, en Cherkassky, V., Friedman, J. H., Wechsler, H., (Eds), *From Statistics to Neural Networks: Theory and Pattern Recognition Applications*. Springer-Verlag, Berlin, pp. 83–104.
- Loikkanen, T., Ahlqvist, T., Pellinen, P., 2009. The role of the technology barometer in assessing the performance of the national innovation system. *Technological Forecasting and Social Change* 76, 1177–1186.
- Lucas, R.E., 1986. Adaptive Behavior and Economic Theory. *The Journal of Business* 59, S401-S426
- Lucas, R.E., 1988. On the mechanics of economic development. *Journal of Monetary Economics* 22, 3–42.
- Luján, J.L., Moreno, L., 1996. El cambio tecnológico en las ciencias sociales: el estado de la cuestión. *Reis* 74/96, 127–161.
- Lundvall, B.A., 1992. *National Innovation Systems: Towards a Theory of Innovation and Interactive Learning*. Pinter, London.
- Lundvall, B.A., Johnson, B., 1994. The learning economy. *Industry and Innovation* 1, 23–42.
- Machlup, F., 1962. *The Production and Distribution of Knowledge in the United States*, Princeton University Press, Princeton, NJ.
- Malthus, T. R., 1820. *Principios de economía política*, Fondo de Cultura Económica, Méjico.

- Margoluis, R., Stem, C., Salafsky, N., Brown, M., 2009. Using conceptual models as a planning and evaluation tool in conservation. *Evaluation and Program Planning* 32, 138–147.
- Martínez–Estudillo, A.C., Martínez–Estudillo, F.J., Hervás–Martínez, C., García–Pedrajas, N., 2006. Evolutionary product unit based neural networks for regression. *Neural Networks* 19, 477–486.
- Mathieson, M. J., 1995. Ordinal models for neural networks, en: Refenes, A., Abu–Mostafa, Y., Moody, J. (Eds), *Neural networks in Financial Engineering*. World Scientific, Singapore, pp. 523–536).
- May, C., 2002. *The Information Society: A Sceptical View*. Polity Press, Cambridge.
- Meng, W., Hu, Z., & Liu, W., 2006. Efficiency evaluation of basic research in China. *Scientometrics*, 69, 85–101.
- McCullagh, P., 1980. Regression models for ordinal data. *Journal of Royal Statistics Society. Series B (Methodological)* 42, 109–142.
- McKim, V.R., Turner, S.P., 1997. *Causality in Crisis?* University of Notre Dame Press, Notre Dame.
- Meadows, D. H., Meadows, D. L., Randers, J., Behrens, W. W., 1972. *Los límites del crecimiento*. Fondo de Cultura Económica, México.
- Meyer, M., 2000. Does science push technology? Patents citing scientific literature. *Research Policy* 3, 409–434.
- Mokyr, J. 2005, The Intellectual Origins of Modern Economic Growth. *Journal of Economic History* 65, 285–351.
- Molero, J. y Buesa, M., 1997. Innovación y cambio tecnológico, en: García Delgado, J.L., (Eds), *Lecciones de Economía Española*. Civitas, Madrid, pp. 61-69.
- Monnier, E., 1995. *Evaluación de la acción de los poderes públicos*. Instituto de Estudios Fiscales, Ministerio de Economía y Hacienda, Madrid.
- Mukerji, S., 1989. *A Fragile Power. Scientists and the State*. Princeton University Press, Princeton (NJ).
- Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., 2005. *Handbook on Constructing Composite Indicators: Methodology and User Guide*. OECD Statistics Working Papers, OECD Publishing, Paris.
- Nelson R.R., 1996. *The sources of economic growth*. Harvard University Press, Cambridge, MA.
- Nelson, R. R., Sidney G. W., 2002. Evolutionary Theorizing in Economics. *Journal of Economic Perspectives* 16, Spring, 23–46.
- Nelson, R. y Winter, S., 1982. *An evolutionary theory of economic change*. The belknap press of Harvard University Press, Cambridge, USA.
- Nelson, R.R., 1993. *National Innovation Systems: A Comparative Analysis*. Oxford University Press, Oxford.
- Nelson, R.R., 1998. The agenda for growth Theory: a different point of view. *Cambridge Journal of Economics* 22, 497–520.
- Newton, A.C., 2010. Use of a Bayesian network for Red Listing under uncertainty. *Environmental Modelling & Software* 25, 15–23.
- Norusis, M., 2007. *SPSS 15.0 Advanced Statistical Procedures Companion*. Prentice Hall Press, Upper Saddle River.

- OECD, 1996. *The Knowledge Based Economy*. OECD/GD (96) 102. OECD, Paris.
- OECD, 1996a. *Employment and Growth in the Knowledge-based Economy*. OECD, Paris.
- OECD, 1996b. *Technology, Productivity and Job Creation*. OECD, Paris.
- OECD, 1996c. *Transitions to Learning Economies and Societies*. OECD, Paris.
- OECD, 2003. *Proposed Standard Practice for Surveys of Research and Experimental Development—Frascati Manual (5^a ed.)*. OECD, Paris.
- OECD, 2009. *Patents Statistics Manual*. OECD, Paris.
- OECD, 2010. *Main Science and Technology Indicators (MSTI) (2nd Ed)*. OECD, Paris.
- OECD, Eurostat, 2005. *Oslo Manual guidelines for collecting and interpreting innovation data (3^a ed.)*. OECD and Eurostat, Paris.
- OECD, Eurostat, 1995. *The measurement of Scientific and Technological activities. Manual on the measurement of human resources devoted to S&T “Canberra manual”, OECD and Eurostat (European Commission), Paris.*
- Ojanen, V., Vuola, O., 2003. *Categorising the Measures and Evaluation Methods of R&D Performance - A State-of-the-art Review on R&D Performance Analysis*, □ Telecom Business Research Centre Working Paper 16, Lappeenranta University of Technology, Finland.
- Olson, J. R., Reuter, H. H., 1987. *Extracting Expertise from Experts: Methods for Knowledge Acquisition*. Technical report 13, Cognitive Science and Machine Intelligence Laboratory, University of Michigan.
- Orlowski, W., 2000. *Knowledge Economy and Knowledge-Based Growth: Some Issues in a Transition Economy*, en: Kuklinski, A. (Eds.), *The Knowledge-Based Economy: The European Challenges of the 21st Century*. State Committee for Scientific Research of the Republic of Poland, Warsaw, pp. 89–96.
- Oszlak, O., O'Donnell, G., 1995. *Estado y Políticas estatales en América Latina: Hacia una estrategia de investigación*, en: Kliksberg B., Sulbrandt J. (Eds), *Para investigar la Administración Pública*. Instituto Nacional de Administración Pública, Madrid, pp. 89-138.
- Padmore, T., Schuetze, H., Gibson, H., 1998. *Modeling systems of innovation: An enterprise-centered view*. *Research Policy* 26, 605-624.
- Pampillón, R., 1991. *El déficit tecnológico español*. Colección Estudios, Instituto de Estudios Económicos, Madrid.
- Parayil, G., 1991. *Technological knowledge and technological change*. *Technology in Society* 13, 289–304.
- Parker, S.C., 2008. *The economics of formal business networks*. *Journal of Business Venturing* 23, 627–640.
- Parlamento Europeo, Consejo y Comisión de la Unión Europea. *Carta de los Derechos Fundamentales de la Unión Europea*, DOUE 2000/C 364/01 de 18/12/2000.
- Parsons, W., 1995. *Public Policy. An Introduction to the Theory and Practice of Policy Analysis*. Cheltenham, Edward Elgar, UK.
- Patton, M. Q., 1997. *Utilization-Focused Evaluation*. *The New Century Text (3rd. ed)*. Thousand Oaks, Sage.

- Pavitt, K., 1998. The inevitable limits of EU R&D funding. *Research Policy* 27, 559–568.
- Pearl, J., 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo, CA.
- Pearl, J., 1995. Causal diagrams for empirical research. *Biometrika* 82, 669–688.
- Pearl, J., 1999. Reasoning with cause and effect. Technical Report, Department of Science. University of California, Los Angeles, California (descargado en noviembre de 2012 de <http://bayes.cs.ucla.edu/jphome.html>).
- Pearl, J., 2009. *Causality: Models, Reasoning, and Inference* (2ª ed.), Cambridge University Press, Cambridge.
- Pedraja, F., Salinas, J., Suárez, J., 2001. La medición de la eficiencia en el sector público, en: Álvarez Pinila, A., (Eds), *La medición de la eficiencia y la productividad*. Pirámide, Madrid. pp. 243–268.
- Pennington, N., Hastie, R., 1988. Explanation based decision making: Effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory and Cognition* 14, 521–533.
- Peffers, A., Tenenbaum, J.B., Griffiths, T.L., Xu, F., 2011. A tutorial introduction to Bayesian models of cognitive development. *Cognition* 120, 302–321.
- Pezzey, J., 1992. Sustainable Development Concept. An Economic Analysis. The International Bank for Reconstruction and Development/The World Bank, Washington, D.C.
- Powell, W.W., Snellman, K., 2004. The Knowledge Economy. *Annual Review of Sociology* 30, 199–220.
- Pyon, Y.S., Li, J., 2009. Identifying gene signatures from cancer progression data using ordinal analysis, en: *BIBM '09 Proceedings of the 2009 IEEE International Conference on Bioinformatics and Biomedicine*. IEEE Computer Society, Washington, DC, pp. 136–141.
- Questier, F., Walczaka, B., Massarta, D.L., Bouconb, C., de Jong, S., 2002. Feature selection for hierarchical clustering. *Analytica Chimica Acta* 466, 311–324.
- Rechenberg, I., 1973. *Evolutions strategie: Optimierung technischer Systemen nach Prinzipien der Biologischen Evolution*. Franmann–Holzboog Verlag, Stuttgart.
- Ricardo, David, 1821. *On the Principles of Political Economy and Taxation*. Library of Economics and Liberty (descargado en Febrero de 2013 de <http://www.econlib.org/library/Ricardo/ricP.html>).
- Romer, P. M., 1986. Increasing returns and long–run growth. *Journal of Political Economy* 94, 1002–1037.
- Romer, P., 1990. Endogenous Technological Change. *Journal of Political Economy* 98, S71–S102.
- Ros, J., 2001. *Development Theory & The Economics of Growth*. The University of Michigan Press, USA.
- Rosenberg, N., 1982. *Inside the Black Box*. Cambridge University Press, Cambridge.
- Rosenblatt, M., 1956. Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics* 27, 832–837.
- Rosenstein–Rodan, P.N., 1943. Problems of Industrialization of Eastern and South–Eastern Europe. *Economic Journal* 53, 202–211.

- Rousseau, S., Rousseau, R. 1997. Data analysis as a tool for constructing scientometric indicators. *Scientometrics* 40, 45-46.
- Ruiz, C. F., Bonilla, R., Chavarro, D., Orozco, L. A., Zarama, R., Polanco, X., 2009. Efficiency measurement of research groups using Data Envelopment Analysis and Bayesian networks. *Scientometrics* 83, 711–721.
- Saen, R.F., 2010. Restricting weights in supplier selection decisions in the presence of dual–role factors. *Applied Mathematical Modelling* 34, 2820–2830.
- Salvador-Carulla, L., Garcia-Alonso C., González-Caballero J.L., Garrido–Cumbrera M., 2010. Use of an operational model of community care to assess technical efficiency and benchmarking of small mental health areas in Spain. *Journal of Mental Health Policy and Economics* 10, 63–76.
- Schmookler, J., 1966. *Invention and Economic Growth*. Harvard University Press, Cambridge.
- Schumpeter, J. A., 1912. *Teoría del desarrollo económico*, Fondo de Cultura Económica, México.
- Schumpeter, J.A., 1961. *Business Cycles. A Theoretical, Historical, and Statistical Analysis of the Capitalist Process*. McGraw-Hill Book Company, New York.
- Sen, A., 1999. *Development as Freedom*. Oxford University Press, Oxford.
- Sharif, N., 2006. Emergence and development of the national innovation systems concept. *Research Policy* 35, 745–766.
- Shashua, A., Levin, A., 2003. Ranking with Large Margin Principle: Two Approaches. *Advances in Neural Information Processing Systems* 15, 937–944.
- Sheskin, D.J., 2011. *Handbook of parametric and nonparametric statistical procedures*. Chapman and Hall/CRC, Boca Raton.
- Singha, R.J., Murtyb, H.R., Guptac, S.K., Dikshitc, A.K., 2012. An overview of sustainability assessment methodologies. *Ecological Indicators* 15, 281–299.
- Smith, A. (1937). *The Wealth of Nations (1776)*. Modern Library, New York.
- Smith, K., Dietrichs, E., Nas, S., 1995. *The Norwegian National Innovation System: A Pilot Study of Knowledge Creation, Distribution and Use*. OECD Workshop on National Innovation Systems, Vienna.
- Smith, K.A., Gupta, J.N.D., 2000. Neural networks in business: techniques and applications for the operations researcher. *Computational Operational Research* 27, 1023–1044.
- Smith, P. C., Street, A., 2000. Measuring the efficiency of public services: the limits of analysis. *Journal of the Royal Statistical Society. Series A (Statistics in Society)* 168, 401–417.
- Solow, R. M., 1956. A Contribution to the Theory of Economic Growth. *The Quarterly Journal of Economic* 70, 65–94.
- Solow, R. M., 1957. Technical Change and the Aggregate Production Function. *Review of Economics and Statistics* XXXIX, 312–320.
- Solow, R. M., 1987. *Growth Theory and After*. Nobel Prize in Economics documents 1987-1, Nobel Prize Committee, Stockholm.
- Solow, Robert M., 1962. Technical Progress, Capital Formation, and Economic Growth. *The American Economic Review* LII, 76–86.

- Soudien, C., 2002. Education in the network age: globalisation, development and the World Bank. *International Journal of Educational Development* 22, 439–450.
- Spirtes, P., Glymour, C., Schienens, R., 2000. *Causation, Prediction, and Search* (2^a ed.). MIT Press, Cambridge, MA.
- Starfield, A.M., Farm, B.P., Taylor, R.H., 1989. A rule-based ecological model for the management of an estuarine lake. *Ecological Modelling* 46, 107–119.
- Stoneman, P., 1987. *The Economic Analysis of Technology Policy*. Clarendon Press, Oxford.
- Thomas, V.J., Sharma, S., Jain, S.K., 2011. Using patents and publications to assess R&D efficiency in the states of the USA. *World Patent Information* 33, 4–10.
- Thurow, Lester C., 1996. *The Future of Capitalism*. William Morrow and Company, Inc New York.
- Toffler, A., 1982. *La tercera Ola*. Plaza & Janés, Madrid.
- Trippi, R., Turban, E., 1994. *Neural Networks in Finance and Investing*. Probus Publishing Co., Cambridge.
- United Nations Development Programme, 2011. *Human Development Report 2011. A Better Future for All*. United Nations, New York.
- Vapnik, V.N., 1999. *The Nature of Statistical Learning Theory*. Springer, Berlin.
- Vedung, E., 1998. Policy Instruments: Typologies and Theories, in: Bemelmans-Videc, M. L., Rist, R.C., Vedung, E. (Eds), *Carrots, Sticks & Sermons*. Transaction Publishers, New Brunswick and London, pp. 21-58.
- Vegara, J. M., 1989. *Ensayos económicos sobre innovación tecnológica*. Alianza Editorial, Madrid.
- Von Hippel, E., 1988. *The sources of innovation*. Oxford University Press, New York.
- Von Krogh, G., Ichijo, K., Nonaka, I., 2000. *Enabling knowledge creation: How to unlock the mystery of tacit knowledge and release the power of innovation*. Oxford University Press, New York.
- Wang, E.C., Huang, W., 2007. Relative efficiency of R&D activities: A cross-country study accounting for environmental factors in the DEA approach. *Research Policy* 36, 260–273.
- Watts, R.J., Porter, A.L., 1997. Innovation forecasting. *Technological Forecasting and Social Change* 56, 25–47.
- Weiss, C. H., 1998. *Evaluation. Methods for Studying Programs and Policies* (2^a ed.). Prentice Hall, Upper Saddle River (NJ).
- Whittaker, E. T., 1922. On a new method of graduation. *Proceedings of the Edinburgh Mathematical Society* 41, 63–75.
- Woolgar, S., 1991. *Ciencia: Abriendo la caja negra*. Anthropos, Barcelona.
- World Bank, 1999. *Knowledge for Development, World Development Report 1998–1999*. World Bank, Washington (DC).
- World Bank, 2007. *Building Knowledge Economies: Advanced Strategies for Development*. World Bank, International Bank for Reconstruction and Development, Knowledge for Development Program, (descargado en noviembre de 2012 de <http://go.worldbank.org/851HK6EUH0>).

- World Bank, 2012. Knowledge for Development (K4D), Knowledge Assessment Methodology 2012, (descargado en enero de 2013 de www.worldbank.org/kam).
- World Economic Forum, 2010. The Global Competitiveness Report 2010–2011. World Economic Forum, Geneva, Switzerland.
- Wright, P.M., Dunford, B.B., Snell, S.A., 2001. Human resources and the resource based view of the firm. *Journal of Management* 27, 701–21.
- Wu, J., Xiong, H., Chen, J., 2009. Towards understanding hierarchical clustering: A data distribution perspective. *Neurocomputing* 72, 2319–2330.
- Wu, J.J., David, L., 2002. A Spatially Explicit Hierarchical Approach To Modeling Complex Ecological Systems: Theory And Applications. *Ecological Modelling* 153:7–26.
- Xavier Ballat, X., 1992. ¿Cómo evaluar programas y servicios públicos? Aproximación sistemática y casos de estudio. Ministerio para las Administraciones Públicas (1ª ed.). Secretaría General Técnica, Instituto Nacional de Administración Pública, Madrid.
- Xu, X., Zhou, C. & Wang, Z. (2009). Credit scoring algorithm based on link analysis ranking with support vector machine. *Expert Systems with Applications*, 36, 2625–2632.
- Yao, X., Liu, Y., 1997. A new evolutionary system for evolving artificial neural networks. *IEEE Transactions on Neural Networks* 8, 694–713.
- Yu, S., Yu, K., Tresp, V., Kriegel, H., 2006. Collaborative ordinal regression, en: Proceedings of the 23rd international conference on Machine learning, 1096. ACM.
- Yu, Z., 2005. Trade, Market Size, and Industrial Structure: Revisiting the Home-Market Effect. *The Canadian Journal of Economics* 38, 255-272.
- Zanak, S. H., Becerra–Fernandez, I., 2005. Competitiveness of nations: A knowledge discovery examination, *European Journal of Operational Research* 166, 185–211.
- Zopounidis, C., Doumpos, M., 2002. Multicriteria classification and sorting methods: A literature review, *European Journal of Operational Research* 138, 229–246.

1. TRABAJOS PUBLICADOS

Derivadas de la presente tesis, las publicaciones aceptadas y publicadas son las siguientes y se adjuntan a continuación:

De la Paz Marín, M., Campoy Muñoz, P., Hervás Martínez, C., 2012. Non-linear multiclassifier model based on Artificial Intelligence to predict Research and Development Performance in European Countries, *Technological Forecasting and Social Change* 79, 1731-1745. (INDICIO DE CALIDAD PRESENTADO).

- Índice de Impacto (JCR 2012): 1,709.
- Área de conocimiento: Planning and Development. Ranking 9/54. Primer Cuartil.

Capítulo de Libro:

De la Paz-Marín, M., Campoy Muñoz, P., Hervás Martínez, C., 2012. Evolutionary Neural Network Classifiers for monitoring Research, Development and Innovation performance in European Union Member States, in: *Methods for Decision Making in an uncertain Environment*. Proceedings of the XVII SIGEF Congress, 309-333.

Por otro lado, fruto del empleo de las técnicas de clasificación ordinal en temas socioeconómicos también se derivó la siguiente publicación, que aunque no se encuentra relacionada con la temática de la tesis, sí que resultaron de utilidad dichas técnicas para otro tema económico y de actualidad como es el de la calificación de la deuda soberana por las Agencias de Rating:

Fernández-Navarro, F., Campoy-Muñoz, P., De la Paz-Marín, M., Hervás-Martínez, C., Yao, X., 2013. Addressing the EU Sovereign Ratings Using an Ordinal Regression Approach, *IEEE Systems, Man, And Cybernetics, PART C: Applications And Reviews*, D.O.I.: 10.1109/TSMCC.2013.2247595.

- Índice de Impacto (2012): 2,009.



Non-linear multiclassifier model based on Artificial Intelligence to predict research and development performance in European countries

Mónica de la Paz-Marín ^{a,*}, Pilar Campoy-Muñoz ^a, César Hervás-Martínez ^b

^a Department of Management and Quantitative Methods, ETEA, Business Administration Faculty, Escritor Castilla Aguayo 4, 14004 Córdoba, Spain

^b Department of Computing and Numerical Analysis, University of Córdoba, Albert Einstein Building, 14071 Córdoba, Spain

ARTICLE INFO

Article history:

Received 23 July 2011

Received in revised form 14 April 2012

Accepted 3 June 2012

Available online 24 June 2012

Keywords:

R&D performance

Neural networks

Evolutionary algorithms

k-means clustering

Multiclassification

European Union

ABSTRACT

This paper deals with one of the most important keys for economic growth: scientific knowledge and innovation, following the linear Research and Development (R&D) model. Patents, scientific publications and expenditure in R&D as well as the personnel involved in these activities are taken into account as proxy indicators, together with variables related to education and economy in order to classify R&D performance in 25 European Union (EU) Member States. This study classifies these countries using a set of variables which characterize them from 2005 to 2008 and analyses the most relevant ones for this classification. The Multilayer Perceptron Model (MLP) and the Product-Unit Neural Network (EPUNN) models, both trained by evolutionary algorithms (EA), were used to classify yearly country observations in clusters previously defined by employing unsupervised algorithm k-means clustering, obtaining four different classes of national R&D performance: low, moderate, high and innovation driven economies. Finally, our methodology is compared to other classification methods normally used in machine learning. The results show that while various methods of classification exist, our methodology obtains models with a significantly lower number of coefficients without decreasing their accuracy in predicting the classification of other European countries or in these countries in the following years.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Research, Development (R&D) and innovation, especially at this time of economic downturn and budget constraint, have become both the key for economic and smart growth in a knowledge-based society as well as a driving force for national competitive advantage [1]. Europe's competitiveness, our future standard of living, depends on its ability to incorporate innovation into products, services, business and social processes and models. The dynamics of modern economies have increased through globalization, greater competition and rapid technological change; consequently, R&D and innovation were first placed at the heart of the European Lisbon Strategy (2000), then re-launched in the Barcelona Council (2002) and today are found in the Europe 2020 Strategy and the European Research Area.

In general, the R&D effort is a very complex structure with multiple factors to resource allocation strategies and to convert them into innovations. In response to international competitive pressure, firms' survival and competitive advantage rely upon R&D ability, and hence, upon innovation.

However, although the design and estimation of models for efficiency (performance) assessment have made an important progress [2], the problem of defining a model to measure the R&D efficiency at the country level does not seem to have been resolved yet in spite of the various attempts in the field. Scholars and experts started investigating the topic from different, albeit complementary, perspectives. Prior studies generally test and evaluate the efficiency/performance of R&D investments mainly using statistical indicators and non-parametric techniques like Data Envelopment Analysis (DEA) and Stochastic Frontier Analysis, at macro

* Corresponding author. Tel.: +34 957 218404; fax: +34 957 218030.

E-mail addresses: mpaz@uco.es (M. de la Paz-Marín), mpcampoy@etea.com (P. Campoy-Muñoz), chervas@uco.com (C. Hervás-Martínez).

[3,4] and micro level (R&D programs, projects and firms [5]¹). It is evident that the purposes and the subjects of measurement efforts, and the context which these concepts are suited for, can be quite different [6].

The context of statistical indicators has been a field of application for composite indicators to measure national Science, Technology and Innovation (STI) performance and comparisons among countries have been a field of application for such composite indicators. Academics, international organizations (with their well-known efforts of normalization of STI indicators and concepts thanks to the OECD, the National Science Foundation or NSF and the EU) and policymakers largely adopt this approach for recommendations in this field [7]. These indicators share two similar characteristics: i) they are “composite” indicators² made up of various other sub-indicators, and (ii) they provide a synthesis of sub-indicators by aggregating them through using arithmetic mean (weighted or un-weighted) [8].

At European country level, a leading example is the study of R&D and innovation performance or, more widely, of the competitiveness and the state of the knowledge-based economy contained in the Innovation Union Scoreboard with respect to the countries under study. The Scoreboard examines differences in efficiency in R&D policies by assuming that efficiency is the ratio of outputs over inputs. Here, it is measured by comparing the ratio between a composite indicator score for one or more input dimensions and one or more output dimensions. The first edition of the Innovation Union Scoreboard in 2010 (based on the previous European Innovation Scoreboard) and the second edition in 2011 include innovation indicators and trend analyses for all 27 EU Member States and others [9]. Thus, the methodology captures more dimensions and variables of a country's innovation performance than in this study.

Based on their Summary Innovation Index (SII) score and the growth rate of the SII, the countries included in the analysis are also divided in four groups or clusters: average performance is measured using a composite indicator building on data for 24 indicators going from a lowest possible performance of 0 to a maximum possible performance of 1. This average in the 2010 report reflects performance in 2008/2009 and the 2011 report refers to the years 2009/2010 due to a lag in data availability. The methodology used for calculating this composite innovation indicator goes through seven steps ranging from identifying and replacing outliers to calculating the composite innovation indexes.

Finally, a wide range of evaluation methodologies and indicators for assessing the socio-economic impacts of R&D have also been presented as a toolbox document by the European Commission [2]. The report presents evaluation from a user perspective and highlights the expectations of different participants describing aspects of evaluation in four broad policy areas, i.e. R&D financing, the provision of R&D infrastructures, technology transfer and its legal framework. Eleven evaluation methodologies are reviewed with their descriptions, requirements, applications and good practice examples and outline the synergies among these evaluation methods.

With respect to the methodology employed in this paper, in the last two decades Artificial Neural Networks (ANNs) have been successfully used in economics for detecting and predicting banking crises [10], modeling financial time series [11], bankruptcy prediction [12] and identifying market structure [13]. ANNs have many advantages over conventional methods. They have the ability to analyze complex patterns with a high degree of accuracy and make no assumptions about the nature of the distribution of the data. Compared with an econometric model, it could be said that it is easier to use ANNs where a forecast needs to be obtained in a shorter period of time. One of the applications of neural networks is pattern recognition³. Like the DEA, a ANN is a non-parametric technique that make weaker assumptions and, from the point of view of application in R&D performance analysis, both are able to detect non-linear functional relationships that are hidden in case data and are able to apply these relationships to new data; finally, the problem can be described by a few independent variables, resulting in low dimensionality or less complexity [14].

In spite of this, to date hardly any works related to R&D performance have been based on Artificial Neural Networks (ANNs) and, basically, these are focused on R&D management activities [15]. We find that Data Mining and Cluster Analysis techniques have been used in a similar way dealing with Regional Innovation Systems [16] and studying the competitiveness of nations [17]. In [16], i.e., the authors use Regional Innovations Systems instead of countries. Thus, a wide range of 265 EU regions at NUTS 2 level are grouped into 9 clusters by employing 11 variables related to economy, education and R&D.

Regarding the neural network models available, the most popular could be the Multilayer Perceptron (MLP) due to its simple architecture yet powerful problem-solving ability, where neurons are grouped in layers and only forward connections exist [18]. However, alternatives to MLP have arisen in the last few years: Product Unit Neural Network (PUNN) models are an alternative to MLPs and are based on multiplicative neurons instead of additive ones. They correspond to a special class of feed-forward neural networks introduced by Durbin and Rumelhart [19].

In many cases, neural networks that use sigmoid unit basis functions⁴ and product unit basis functions (PUNNs) are trained by using evolutionary algorithms (EAs), obtaining with this method significant advantages with respect to traditional training approaches [15]. One of the most important aspects of employing evolutionary programming (EP) algorithms as a modeling methodology is their ability to make the complexity of the network more flexible during training through the use of structural mutation operators that add and delete nodes and add and delete connections.

The aim of this work focuses on predicting the classification of R&D and innovation performance throughout 25 European countries thanks to their assignment into clusters, which will help to monitor European strategies for R&D and innovation and some key features related to the EU innovation policy, in general. First, *k*-means clustering (an unsupervised algorithm) is applied to detect behavioral patterns in 25 EU Member States between 2005 and 2008. As a result of clustering, a number of classes will be set to define the characteristics of each one. Then a multiclass classifier is built to assign each country–year observation to its corresponding cluster

¹ For a survey, see V. Ojanen and O. Vuola (2003) [20].

² Reviews and contribution about the methodology of the composite indicators can be found on the European Commission-Joint Research Centre, website: <http://composite-indicators.jrc.ec.europa.eu/>, last access March 6th, 2011.

³ For a book of neural networks see S. Hayken (1994) [21] and R.P. Lippman (1989) [22].

⁴ Also called MLP or SUNN in this paper.

according to a set of specific features that characterize that country. To build the classifier, we develop neural network models using Sigmoid Unit Neural Networks (SUNN) and Product Unit Neural Networks (PUNN) trained with an EA. These EANNs use structural operation mutations that allow us to build models with a reduced number of basis functions and connections, which determine a lower number of coefficients in the model.

The study has the following structure: after drawing the relevance of R&D performance evaluation, the methodology section is devoted to the clustering methodology and class descriptions, neural networks, multiclass classification and evolutionary algorithm methodologies; then, the experimental study is carried out with the description of the database, the results and the analysis of the best model. Finally, and after a discussion section, conclusions are presented about experimental results.

2. Research, development and innovation performance

The relevance of R&D performance evaluation is highlighted even further due to the current worldwide economic crisis and, especially, with the adjustment processes that establish budget cuts and/or interrupt on-going programs. European governments have increased their countries' levels of R&D expenditures and established expenditure intensity targets relying on their impact on economic performance and, ultimately, to achieve the goal of knowledge-intensive economies where research, its commercial exploitation, education, and other intellectual work play a growing role.

R&D activities "comprise creative work undertaken on a systematic basis in order to increase the stock of knowledge, including knowledge of man, culture and society, and the use of this stock of knowledge to devise new applications" [23], encompassing a multitude of phases, from basic research to application and development, and leading to innovations that could create competitive advantages.

In R&D-performance assessment literature, there are many differences in perspectives and definitions of performance vary widely [24]. The R&D performance measurement is reflected in terms such as effectiveness, efficiency, productivity, quality, capability, success, control, monitoring, assessment, benchmarking evaluation, even technological capabilities, which are very often employed as synonyms [25]. This study uses the concept of performance measurement as part of the broader concept of evaluating the efficiency of the R&D activities carried out by public and private sectors in the European Union that deal with the relationships between inputs and outputs of such activities (the quantity of output which can be efficiently produced for a given input level).

The underlying argument for R&D public [26] and private expenditure is that the more resources invested by a country in innovation, the higher its expected productivity and its economic growth [27–30], and its competitiveness [31], leading to a more inclusive, knowledge-based and sustainable economy.

Thus, two important underlying assumptions must be considered and stressed in this paper when selecting the variables employed to measure a country's R&D performance. Considering that one of the main determinants of productivity and economic growth in the long run is the technical change and that modern endogenous growth theories [32] postulate that innovation should be regarded as an output, our first assumption is the role of technical change as the major source of productivity growth and R&D as a major source of such technical change; it was during the 1990's when the so-called "Schumpeterian-growth models" in Endogenous Growth Theory emerged with the publication of four articles and their rapid development followed the general evolutionary process of creative destruction [28,33–35].

The essential feature of these models is the incorporation of technological progress, which is generated by the endogenous introduction of product and/or process innovations, putting technological change at the center of an endogenous explanation of economic growth and showing that in order to accelerate growth it is not enough to increase R&D expenditures; it is also necessary to increase the total quantity of inputs related to the R&D process [36]. Although the Endogenous Growth Theory formalizes only a subset of Schumpeter's ideas, it is much closer to the concept of creative destruction than other existing economic growth theories and provides a true formalization of Schumpeter's insights into endogenous technological change as the core of economic development [37].

Regarding the second assumption, after the use of the linear model in the 1950s, one of the key elements of the innovation process was that research, development and innovation are linked in a non-linear and interactive manner, denominated *chain linked model* [38,39]. Since then, even more complex models like the *cycle model* [40] have emerged. Nevertheless, we make the assumption of a *linear-model* because, as Godin (2006) [41] remarked, this model functions as a social fact and official databases are collected following this linear model where everything originates with a discovery and then developmental activities follow in a sequential process until achieving the successful introduction of the innovation in the market happens.

Even in times of economic downturn, some authors claim that it is fundamental to continue supporting R&D and innovation activities [42]. For all these reasons, assessments of a country's R&D performance play a major role in current discussions of R&D and innovation performance, and result in an increasing request for and interest in for finding overall efficiency and productivity measures for R&D investment and innovation.

3. Methodology

3.1. *k*-means clustering and class descriptions

In many situations, decision makers have to group the object of their analysis into homogeneous classes. The two mostly employed techniques for grouping objects with similar properties are classification and clustering. The clustering technique adopted is based on the *k*-means algorithm, which is one of the most widely used unsupervised technique. This method allows grouping the alternatives

into categories so as to keep the shortest possible distances between them within the same category, while maintaining the largest distances between the centers of different categories [43].

Our initial task is to employ a cluster technique to detect patterns among European countries in relation to their R&D performance. One of the advantages of the cluster technique is to reveal patterns hidden within the data simultaneously over a multidimensional set of variables.

Before the application of the model, the number of classes desired must be specified. As this information is not necessarily obvious, we executed experiments for a number of classes ranging from 1 to 9. Partitions in three or four classes appear to be good compromises: the sum of intra-class distances divided by the total sum of distances is relatively low while the number of classes is also small. Considering more classes will not lead to a significant increase in the quality of the partition [40]. On the other hand, considering a partition in four classes, instead of three classes, permits the quality of clustering to increase.

The simple k -means clustering algorithm [44] was used due to its easy implementation and low computational cost in WEKA software [45]. As commented previously, given that there was a significant decrease in the cluster sum of the squared errors when we increase k , a value of $k=4$ was chosen. Then, all variables in each cluster were compared to their centroid to see which are nearest or farthest away from it according to the Euclidian distance. Using an economic analysis of these patterns and cluster centroids, the profile of each group was drawn and labeled⁵ into classes as follows (for descriptions of variables, clusters centroids and standard deviations, see Table 1):

Moderate innovation countries (cluster 1): this cluster comprises the countries in which R&D plays a less than central role given that R&D intensity is not well developed either in the public and private sectors or, on average, in the number of researchers and human resources in Science & Technology (S&T). So, these countries have lagged behind in performance with very low patenting intensity and number of scientific publications. The amount of investment in R&D can only be described as a complementary driver of growth. These countries are not economically strong (in terms of employment rates, Gross Domestic Product (GDP) and real labor productivity growth) and give little emphasis to Ph.D. graduates and life-long learning, although they present good performance in community trademarks and license and patent revenues.

Innovation-driven countries (cluster 2): this class involves the most developed countries. They are economically prosperous (best employment rates, although GDP and real labor productivity growth per hour, RLPGH, are slower due to their former situation), and they present the largest proportion of investment in R&D, researchers, human resources in S&T and doctoral graduates. We also find emphasis on tertiary and life-long education, which is all representative of their education systems. This has provided the workforce with the skills needed for a changing environment and has laid the groundwork for high levels of technological adoption and innovation. They present their best performances in scientific publications and patent revenues as well.

Low innovation countries (cluster 3): these are mainly represented by the transition economies of the New Member States (except Greece and Portugal). As new economies, they involve a high growth in GDP and in RLPGH due to their abundance of natural resources, the low population and a high domestic demand following a period of socialism, among other reasons, but characterized also by a low rate of employment. Nonetheless, their economies do not seem to be based either on education or on R&D investment and this could be an obstacle for future growth and for their future socioeconomic development.

High innovation countries or followers (cluster 4): countries in this cluster fit well in the model of technological competitiveness with high average values in most variables and are close to cluster 2 in the number of Ph.D. graduates, the employment rate, number of patents by residents, and medium and high-technology exports (these two latter cases hold the highest position). However, they present lower average values in terms of the profitability of their trademarks and license and patent revenues as well as in the number of persons enrolled in tertiary education. Life-long learning is not as good as one might suppose in this group of countries. Thus, their economic and education performances, regardless of the above-mentioned variables, are worse than in the innovation-driven countries cluster, although it falls close behind.

Mapping these clusters (see Figs. 1–3) it can be observed that only two country–year patterns change their classification throughout the time span: Malta and Estonia. Both change from a low to a moderate innovation cluster, mainly due to the increase of world wide patent applications of residents filed through the Patent Cooperation Treaty procedure or with a national patent office (TPATRE) and the license and patent revenues from abroad (PATREV). But while the former moves in 2006, the latter moves in 2008.

3.2. Neural networks

Artificial Neural Networks (ANNs) are calculable structures consisting of a number of the same type elements, which execute relatively simple functions. Processes in ANNs are sometimes associated with those that take place in the nervous system of living organisms. One of the most useful and successful applications of neural networks in data analysis is the Multilayer Perceptron

⁵ In a similar way, and based on their average innovation performance, the Innovation Union Scoreboard (IUS) classifies all Member States and another European countries into four groups: innovation leaders, innovation followers, moderate innovators and modest innovators. In the case of cluster 2 (innovation-driven countries), for example, we must say that it is a commonly used term that can be found in a number of studies and reports related to the competitiveness of countries.

Table 1

Descriptions of variables, means of cluster centroids and standard deviations.

Code	Variable ^e	Unit	#Pattern/total	Cluster 1 ^a	Cluster 2 ^b	Cluster 3 ^c	Cluster 4 ^d
			100	24	16	40	20
<i>R&D enablers</i>							
HUMLF	Human Resources in Science and Technology	% Of labor force	31.39 ± 6.08	30.09 ± 6.03	40.52 ± 6.08	27.70 ± 6.08	33.04 ± 6.11
RDPER	R&D personnel (include all persons employed directly in R&D, plus persons supplying direct services)	% Of labor force	0.99 ± 0.48	1.00 ± 0.48	1.59 ± 0.49	0.66 ± 0.48	1.17 ± 0.47
GERDBU	Gross Domestic Expenditure on R&D: business sector	% Of GDP	0.94 ± 0.72	0.78 ± 0.69	2.00 ± 0.73	0.40 ± 0.72	1.38 ± 0.71
GERDGO	Gross Domestic Expenditure in R&D: government and higher education sectors	% Of GDP	0.56 ± 0.21	0.41 ± 0.21	0.89 ± 0.21	0.45 ± 0.21	0.70 ± 0.21
<i>R&D results</i>							
SPUBLI	Number of scientific publications	Per 1000 population	1.23 ± 0.65	1.09 ± 0.63	2.29 ± 0.65	0.74 ± 0.65	1.53 ± 0.64
TPATRE	Worldwide patent applications filed through the Patent Cooperation Treaty procedure or with a national patent office (residents)	Per 1000 population	0.14 ± 0.14	0.07 ± 0.14	0.26 ± 0.14	0.06 ± 0.14	0.31 ± 0.14
<i>Education</i>							
PHD06	Number of Ph.D. graduates	Per 1000 population	0.17 ± 0.09	0.10 ± 0.09	0.25 ± 0.09	0.14 ± 0.09	0.24 ± 0.09
TERTIT	Population having completed tertiary education	% Of population aged 30–34	31.09 ± 10.29	38.20 ± 10.39	41.58 ± 10.40	22.79 ± 10.29	30.76 ± 10.56
LLEARN	Life-long learning	% Of persons aged 18 to 64	15.90 ± 6.71	12.76 ± 6.54	27.26 ± 6.80	11.71 ± 6.71	18.94 ± 6.92
<i>Economy</i>							
GDPGRO	Growth rate of GDP volume	% Of change on previous year	3.01 ± 3.17	1.63 ± 2.94	2.17 ± 2.99	4.50 ± 3.17	2.35 ± 3.04
RLPGH	Real labor productivity growth per hour worked	Real output per unit of labor input (measured by the total number of hours worked).	1.89 ± 2.36	0.80 ± 2.21	1.08 ± 2.29	2.84 ± 2.36	1.93 ± 2.30
EMPLO	Total employment rate	% Of persons aged 20 to 64 in employment	70.74 ± 5.59	69.45 ± 5.43	77.47 ± 5.63	67.90 ± 5.59	72.59 ± 5.43
PATREV	License and patent revenues from abroad	% Of GDP	0.39 ± 0.51	0.55 ± 0.53	1.00 ± 0.52	0.12 ± 0.51	0.27 ± 0.52
TRADEM	Community trademarks	Per billion GDP (in PPSE)	4.59 ± 3.22	7.65 ± 3.34	5.38 ± 3.28	2.39 ± 3.22	4.69 ± 3.24
MHTEXP	Medium and high-technology exports	% Of total manufactured exports	49.16 ± 12.00	48.65 ± 12.08	47.80 ± 11.94	45.98 ± 12.00	57.21 ± 11.92

^a Belgium05–08, Cyprus05–08, Ireland05–08, Luxembourg05–08, Malta06–08, Spain05–08, Estonia08.^b Finland05–08, Sweden05–08, Denmark05–08, Netherlands05–08.^c Czech Republic05–08, Estonia05–07, Greece05–08, Hungary05–08, Italy05–08, Latvia05–08, Lithuania05–08, Malta05, Poland05–08, Portugal05–08, Slovakia05–08.^d Austria05–08, France05–08, Slovenia05–08, United Kingdom05–08, Germany05–08.^e Data source: Eurostat for all variables except SPUBLI (SCImago JCR), TPATRE (World Bank Database), TRADEM and MHTEXP (Innometrics/Eurostat).

Model (MLP). MLP models are non-linear neural networks models that can be used to approximate almost any function with a high degree of accuracy [21] and they are especially suitable for classification [22].

The key difference between neural networks and other problem-solving methods is that neural networks learn by example rather than by being programmed with specific and pre-established rules; they can modify themselves to fit the data presented. Furthermore, neural models are also able to learn from very noisy, distorted, or incomplete sample data [45].

An ANN consists of an input layer, the hidden layers and an output layer. The input layer takes in information (e.g. independent variables) and passes it on to the downstream neurons via connections. These links are assigned weights in the artificial neural network and thus control the flow of information.

The base unit of any neural network is the neuron (processor). Each unit or neuron has a number of inputs that are combined to produce a single output. Each neuron is able to sum up many inputs, whether these inputs are from a database or from other neurons, with each input modified by an adjustable weight w_i . The sum of these weighted inputs s is added to an adjustable threshold for the neuron and then passed through a modifying (transfer) function f that determines the final output y .

Typical activation functions in neural networks are Logistic, Symmetric Logistic, Sigmoid, Hyperbolic Tangent, Gaussian, and Sinusoidal [46].

The hidden layer is the series of relationships calculated in the network's training process. There is no theoretical limit to the number of hidden layers but typically there will be one or two. ANNs have become an important tool for classification since recent

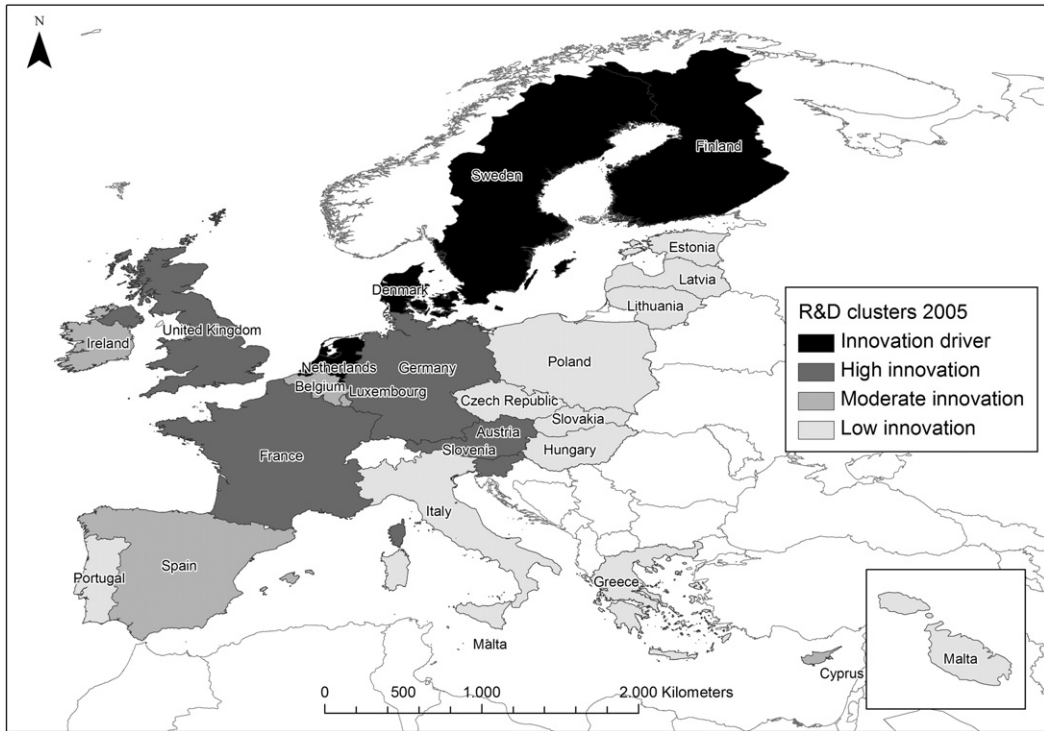


Fig. 1. Clusters with country-year patterns 2005.

research activities have identified them as a promising alternative to conventional classification methods, such as linear discriminant analysis or decision trees. Different types of ANNs are now being used for classification purposes: Multilayer Perceptron Model (MLP), where the transfer functions are Sigmoidal-Unit basis functions, SU; Radial basis functions (RBF), kernel functions where the transfer

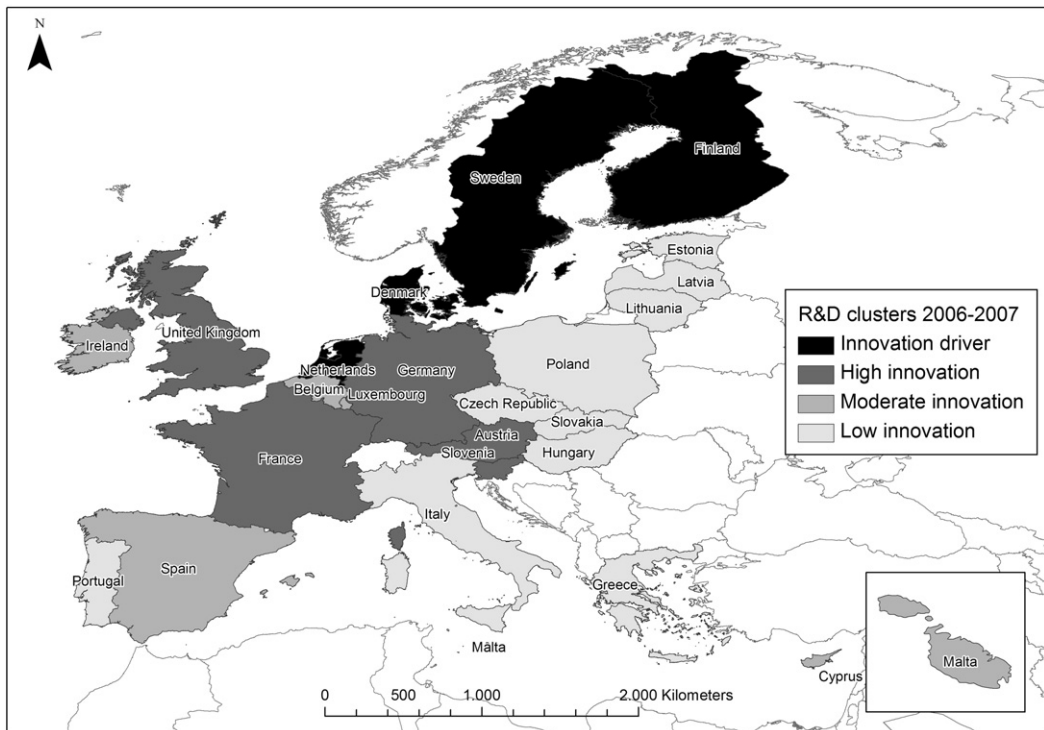


Fig. 2. Clusters with country-year patterns 2006–2007.

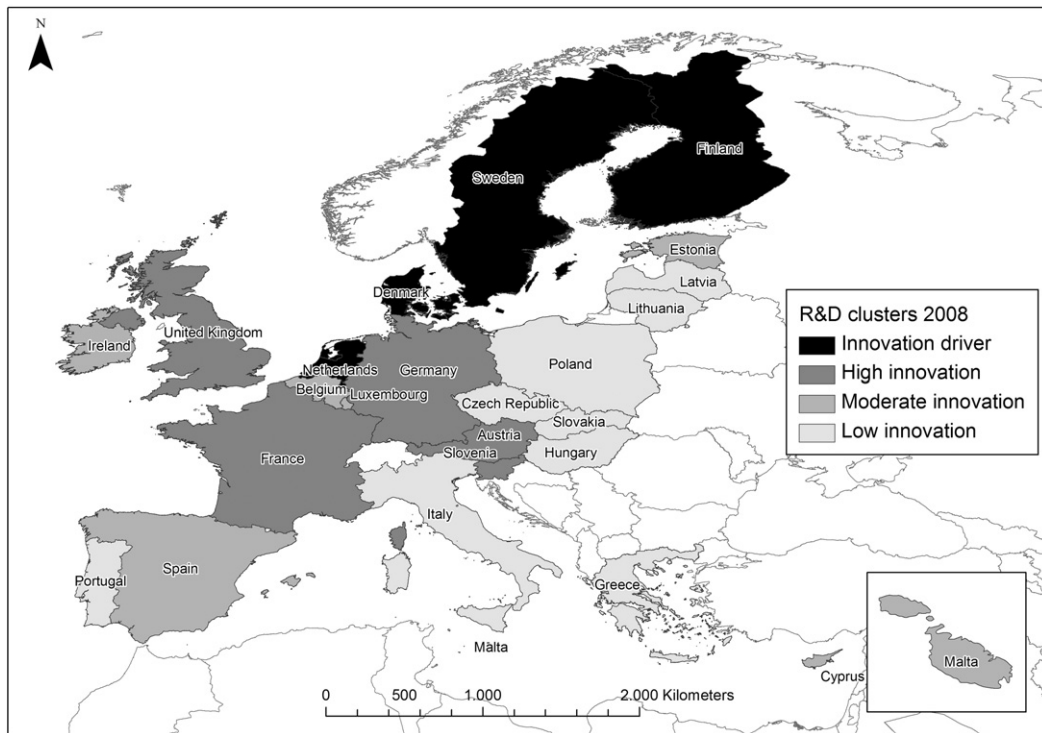


Fig. 3. Clusters with country–year patterns 2008.

functions are usually Gaussian; the General Regression Neural Networks (GRNN); a class of multiplicative ANNs which comprises such types as sigma-pi networks; and Product-Unit basis functions, where a multiplicative node is given as $y_i = \prod_{j=1}^k x_j^{w_{ij}}$ where k is the number of inputs.

In order to solve the problem of selecting the proper size for the hidden layer of the net, a new category of algorithms has been introduced to automatically determine the structure of the network. These include the orthogonal least-squares algorithm; individual training of each hidden unit based on functional analysis; constructive methods, where the structure of the network is built incrementally; pruning methods that start with an initial selection of a large number of hidden units which is reduced as the algorithm proceeds; and the simultaneous selection of network structure and parameters by employing optimization methods based on genetic programming algorithms, EP (see Section 2.3.). This training methodology will be applied in this research study on Evolutionary Sigmoidal-Unit Neural Networks (ESUNN) and Evolutionary Product-Unit Neural Networks (EPUNN) models.

3.3. Multi-class classification

Multi-class pattern recognition is a problem of building a system that accurately maps an input feature space to an output space of more than two pattern classes. While the two-class classification problem is well understood, there has been relatively less research on multi-class classification. Many pattern classification systems have been developed for two-class classification problems and theoretical studies of learning have focused almost entirely on learning binary functions including the support vector machines (SVM), ANNs and the error backpropagation (BP) algorithms.

A neural network system F can be a system of neural networks or a single neural network whose weights are determined by a neural learning algorithm (evolutionary algorithm in our case).

Let a training sample be $D = \{(\mathbf{x}_n, \mathbf{y}_n); n = 1, 2, \dots, N\}$ where $\mathbf{x}_n = (x_{1n}, \dots, x_{kn})$ is the vector of measurements taking values in $\Omega \subset \mathbb{R}^k$ and y_n is the class level of the n -th individual. We adopt the common technique of representing the class levels using a “1-of- Q ” encoding vector $\mathbf{y} = (y^{(1)}, y^{(2)}, \dots, y^{(Q)})$ such as $y^{(l)} = 1$ if \mathbf{x} corresponds to an example belonging to class l and $y^{(l)} = 0$ otherwise. Based on the training sample we wish to find a classifier $ANN(\hat{\theta})$ that can result in a good classification of the patterns in the generalization set. This would be the same as obtaining a decision function $C: (\Omega, \theta) \rightarrow \{1, 2, \dots, Q\}$ to classify the individuals (the neural networks). A misclassification occurs when C assigns an individual (based on a measurement vector) to class j when it is actually coming from a class $l \neq j$. See the Appendix A for further technical details.

3.4. Evolutionary algorithm

Evolutionary algorithms (EAs) are a class of stochastic search algorithms and optimization heuristics derived from the classic evolution theory, that mimic the metaphor of natural biological evolution operating on a population of potential solutions and applying the principle of survival of the fittest. These population dynamics follow the basic rule of the Darwinistic evolution theory: the basic idea is that if only those individuals which meet a certain selection criteria in a population reproduce, and the other individuals in the population die, population will converge to those individuals that best meet the selection criteria, just as in natural adaptation [47].

EAs are applicable to a wide range of problems in learning and optimization. They have been applied to numerous problems in combinatorial optimization, function optimization, artificial neural network learning, fuzzy logic system learning, etc. This paper first introduces EAs and their basic operators. Three major branches of EAs exist and they are genetic algorithms by John Holland, one of the early pioneers of evolutionary computation [48], evolution strategies [49] and, the one we follow in this study, evolutionary programming [50].

As mentioned in the introductory section, one of the most important aspects of employing evolutionary programming (EP) algorithms as a modeling methodology is their ability to make the complexity of the network more flexible during training through the use of five structural mutation operators that add and delete nodes and add and delete connections and node fusion. Thus, to estimate the structure and learn the weights of artificial neural network models, an Evolutionary Artificial Neural Network (EANN) was applied due to the simplicity of the basis functions in the best net models within a basic framework that is detailed in [Appendix B](#).

4. Experimental study: predicting country R&D performance

4.1. Dataset description

The insights gained from the use of the EANN were applied to a dataset of 25 EU Members for R&D performance classification. Due to data availability (until 2008 in most official databases), 100 items were selected considering each country and year from 2005 to 2008 as a single composed item since no changes between clusters of country–year items are expected. However, the incorporations of Bulgaria and Romania are dated in 2007, thus scientific and technological European policy and strategy were not directly applicable and so their inclusion has not been considered appropriate for this classification. The variables were selected based on the type of consensus which international OECD Manuals and reports [23,51], PRO INNO Europe studies (an initiative of the Directorate of General Enterprise and Industry of the European Commission) [52] and expert literature have reached about the type of indicators necessary to measure R&D performance in a country. Besides data availability, other reasons for selecting and isolating the period 2005–2008 were to avoid the effects of economic cycles and that the data pertaining to the beginning of the worldwide economic crisis could lead to ambiguous results. Thus, it could have been quite a while after the onset of the crisis before its effects were observable in data. To avoid this, we selected our data until 2008, just after the sub-prime episode date in August 2007.

Both the difficulty in finding data about R&D and innovation as well as the missing data for a number of countries force us to select proxy variables in most cases. Thus, fifteen main variables were selected, as shown in [Table 1](#), for characterizing R&D performance at country level following the input/output orientation in their interpretation. In the side of the inputs, we ascribe to the idea that the production of knowledge requires specific investments in R&D and that the role of human capital is of great importance, thus the amount of funds and the number of researchers were selected as a proxy for both the capital input and labor input of R&D intensity, respectively. The R&D expenditure was broken down into public and business sectors.

Data were obtained from various sources (Eurostat's official website, World Bank Database and SCImago Journal & Country Rank in the case of scientific publications per country), adjusted when necessary to 1000 per population. Despite this, an imputation has been carried out for missing data when necessary (it occurred in up to 1.3% of the cases); this imputation was made by linear regression [53]. Six models were fitted for the tasks at hand, one for each variable with missing data, (PHD06, GERDGO, GERDBU, TPATRE, PATREV and RDPERS). The dependent variable is the year for which there is no data. The covariates are the remaining years without missing data. All models were fit to a significant level of $\alpha = 0.05$ and a determination coefficient $R^2 > 0.90$.

Among the resulting set of variables of R&D activity [54], scientific publications have been considered the major output of research and they are widely used to evaluate the performance of researchers. Patents were also selected since they are often used as the direct output of the innovation process because technical advance and technological innovation are difficult to measure [55].

R&D performance includes the perspective of education with two variables, Ph.D. graduates and tertiary education students, who are among the main objectives of government R&D policies because human resources, with their managerial and organizational skills, play an ever increasing role in the performance of R&D activities.

Finally, economic variables like Gross Domestic Product growth (GDPGRO), per-hour-growth of real labor productivity and medium and high-technology exports are taken into account to identify patterns among countries.

4.2. Results

We have compared Evolutionary Sigmoidal Unit Neural Networks (ESUNNs) and Evolutionary Product Unit Neural Networks (EPUNNs) with nine state-of-the-art methods that are well known in the literature. Eight of them have been configured and run in WEKA [45] and the Simple Linear Discriminant Analysis (SLDA) method is available in SPSS [56]. The methods used for comparison are: MLP: a neural network classifier that uses back-propagation to calculate weights; SLDA: a multivariate statistical procedure that

derives equations to classify instances [57]; C4.5: a classifier tree for generating a pruned or unpruned decision tree; AdaBoost100: a classifier tree based on the AdaBoost.M1 algorithm with a maximum of 100 iterations; LMT: a classifier for building logistic model trees, which are classification trees with logistic regression functions at the leaves [58]; NaiveBayes: a classifier whose numeric estimator precision values are chosen based on the analysis of the training data; SLogistic: a classifier for building linear logistic regression models; MLogistic: a classifier for building a multinomial logistic regression model with a ridge estimator; SVM: a classifier for building a linear model (maximum-margin-hyperplane) over a kernel space. A Gaussian kernel is used for the selection of the SVM hyper-parameters (regularization parameter C , and width of the Gaussian functions γ); a grid search algorithm was applied with a 10-fold cross validation, using the following ranges: $C \in \{2^{-5}, 2^{-3}, \dots, 2^{-15}\}$ and $\gamma \in \{2^{-15}, 2^{-33}, \dots, 2^3\}$.

Table 2 shows a comparison of the accuracy of eleven methods (three stochastic, from which the best model is chosen, and eight deterministic) in predicting the R&D performance of each country. From a descriptive point of view, the Best ESUNN, the Best EPUNN, the Best MLP, LMT, SLogistic, MLogistic and SVM models obtain equal results in CCR_G (100), MS_G (100) and K_G (1) for the generalization set. With respect to the number of connections for each model (see last column in Table 2), the Best ESUNN model presents the lowest number of coefficients (20); thus this is the selected model, which will be explained further on.

With respect to the results obtained in the 30 runs for the three stochastic algorithms (see Table 3), it is clear that the best results are given by the MLP neural network, although the ESUNN methodology provides the second best results with a high mean (98% for CCR_G) and a low standard deviation, SD, (2.73% for CCR_G). The principal difference between these methodologies is the number of connections or model coefficients, because for MLP the mean is 44 and for ESUNN the mean is 26.57 with a SD = 3.47.

Thus, considering accuracy and the lower number of connections, we recommend the ESUNN methodology for country–year classification.

4.3. Best model

Once the algorithm has been run 30 times to satisfy central limit theorem, 30 best models are obtained, both for ESUNN and EPUNN. Then, the following ESSUN model is chosen: the model consists in nine variables in the input layer, two hidden nodes with sigmoid transfer function units and three nodes in the output layer which return the probability a pattern has of belonging to each class. From them, the probability of membership in the fourth class is calculated. The selected input variables were TERTIT, LLEARN, GERDGO, TRADEM, RLPGH, GPGROW, TPATRE, PATREV and HUMLF.

Interest in the proposed best ESUNN model is due to the fact that it improves the results of other approaches (see Table 2) and because it provides an (non-linear) interpretable model. Table 4 shows the expression of the probability of this best ESUNN model as well as its performance.

Table 4 also shows the construction of sigmoid unit SU_1 as a function of variables TERTIT, HUMLF (which represent labor force skills), PATREV, TRADEM (they embody the profitability of innovation through patents and trademarks) and RLPGH (real labor productivity growth per hour worked). The same table shows the construction of SU_2 as a function of variables TERTIT, HUMLF (as well as SU_1) and LLEARN (which denotes the importance of education for labor force skills), GERDGO (public policy support for R&D investment that stimulates innovation processes in the private sector), TPATRE (main proxy indicator for measuring a country's innovation capability) and GDPGRO (the most common indicator of an economy's well-being).

On the other hand, if one looks at the coefficients of the sums of the exponents that belong to the sigmoidal units, along with the importance of variables according to the probability each one has of belonging to class (g_i) in Table 5, the variables which have the strongest effects for measuring R&D performance at country level are those that present (++) and (--). Variables with (+) or (–) have a lower effect and those with (=) have no significant effects. This means that an increase in variable value implies mainly an increase (+), a decrease (–) or non-significant changes (=) in the probability of belonging to the corresponding cluster. The analysis was performed normalizing data values to [0.1–0.9] and giving the maximum value 0.9 to each variable for testing the variations on such probability of being in one of the four clusters. In this manner, we were able to identify the key variables for improving their value in order to increase the R&D performance of a concrete European country.

Table 2

Statistical results of the CCR_G , MS_G , K_G and number of connections (#conn.).

Method ^a	CCR_G (%)	MS_G (%)	K_G	#Conn.
Best ESUNN	100	100	1	20
Best EPUNN	100	100	1	26
Best MLP	100	100	1	44
SLDA	96	80	0.94	24
C4.5	84	77	0.78	13
AdaBoost100	52	0	0.31	24
LMT	100	100	1	27
NaiveBayes	92	80	0.89	124
SLogistic	100	100	1	27
MLogistic	100	100	1	48
SVM	100	100	1	96

^a The best quantitative result is represented in bold face.

Table 3
Mean and standard deviation (SD) for stochastic methods^b.

Method ^a	CCR _G (%)	MS _G (%)	K _G	# Conn.
	Mean ± SD	Mean ± SD	Mean ± SD	Mean ± SD
ESUNN	98.00 ± 2.73	92.33 ± 9.60	0.97 ± 0.05	26.57 ± 3.47
EPUNN	94.00 ± 4.79	84.55 ± 11.13	0.92 ± 0.07	29.93 ± 4.66
MLP	100.00 ± 0.00	100.00 ± 0.00	1.00 ± 0.00	44.00 ± 0.00

^a The best quantitative result is represented in bold face.

^b In the case of these stochastic algorithms, the experimental design was conducted using a holdout procedure with 30 runs. Approximately 75% of the patterns were randomly selected for the training set and the remaining 25% for the test set.

5. Discussion

As explained in Section 3.3, after nine out of fifteen variables were selected in our best model (TERTIT, LLEARN, GERDGO, TRADEM, RLPGH, GPGROW, TPATRE, PATREV and HUMLF), our analyses pinpointed (see Table 5) the variables that could exert a major influence on the probability of belonging to the cluster that represents better R&D performance (“innovation-driven countries” or cluster 2). These variables were life-long learning (LLEARN), total application patents of residents (TPATRE) and Gross Expenditure in R&D from the public sector that involves government and higher education sectors (GERDGO). They correspond to Education, R&D results and R&D enabler indicators respectively.

While these variables improve the R&D performance classification of a country, TRADEM and PATREV tend to reduce it. These variables represent economic aspects of R&D and it could be read as they cannot be isolated for interpretation but within the context of an innovation system as a whole. On the other hand, if we increase the values of GDPGRO and RLPGH to their maximum normalized value (0.9), it does not vary the probability of a country to be classified in one of the four groups.

Interestingly, the number of tertiary students (TERTIT) along with LLEARN tends to increase a country's classification on R&D performance since education is basic for improving the necessary high-skills in a knowledge-based economy and for achieving a sustainable development [59].

Thus, one contribution of this paper is to confirm the relevance of variables related to knowledge-generating activities, specifically human capital and education, as well as the need to increase resources devoted to R&D and innovation. As the latter activities, the accumulation of human capital has long been considered an important engine of economic growth in theoretical models [60]. The analysis of our model supports the proposition that countries that invest in human capital have stronger R&D and economic performance [61] and the academic analysis that include education resources to explain cross-country variations of innovation productivity [i.e. 62].

European countries could consider to continue supporting the resources devoted to education in addition to R&D funds since these skills are essential in the application of new industry and service technologies that finally leads to economic growth due to a higher productivity by means of innovation activity [63].

Finally, due to the present fragmentation in innovation policy between regional, national and EU political arenas and institutions, monitoring such policy using this model could contribute to the debate of the future governance model of European innovation policies that could go from a centralized one to a progressive decentralization to regional, national or international initiative scenarios [64], once new data are available for future years or new Member States.

6. Conclusions

This study examined nine machine learning and two neural network classifiers to predict R&D performance classification in 25 EU Member States through the identification of the best model and analyzing major factors out of 15 variables that could influence the probability of belonging to a R&D performance class, previously established by means of an unsupervised *k*-means clustering.

Table 4
Expression of probability and the performance of the best ESUNN model.

Best ESUNN model
$f_1(x, \theta_1) = 4.69 - 18.69 * SU_1$ $f_2(x, \theta_2) = 9.30 - 13.62 * SU_2$ $f_3(x, \theta_3) = 0.49 + 13.49 * SU_1 - 23.59 * SU_2$ $f_4(x, \theta_4) = 0$
$SU_1 = \frac{1}{1 + \text{EXP}(-(-9.32 + 8.07 * \text{TERTIT} + 9.31 * \text{TRADEM} - 4.08 * \text{RLPGH} + 5.03 * \text{PATREV} + 3.68 * \text{HUMLF}))}$
$SU_2 = \frac{1}{1 + \text{EXP}(-(-12.44 + 3.42 * \text{TERTIT} - 9.61 * \text{LLEARN} - 9.04 * \text{GERDGO} - 3.07 * \text{GDPGRO} + 7.32 * \text{TPATRE} + 4.47 * \text{HUMLF}))}$
CCR _G = 100%; MS _G = 100%; K _G = 1

Table 5
Influence of each variable on a higher or lower probability of appearing in a given cluster.

Variable	Probability cluster1 ^a	Probability cluster2	Probability cluster3	Probability cluster4
TERTIT	(--)	(--)	(+)	(++)
LLEARN	(--)	(++)	(+)	(--)
GERDGO	(-)	(++)	(-)	(--)
TRADEM	(--)	(--)	(++)	(++)
RLPGH	(+)	(=)	(=)	(--)
GDPGRO	(=)	(=)	(=)	(=)
TPATRE	(--)	(++)	(+)	(--)
PATREV	(--)	(--)	(++)	(++)
HUMLF	(-)	(=)	(-)	(=)

^a Probability of Cluster *l* is $g_l(x, \theta_l)$.

Thus we are able to find out the most relevant issues for stakeholders involved in decision making processes to provide better allocation of scarce resources. In this sense, the model proposed could be useful for evaluating the effectiveness and efficiency of EU Strategies devoted to R&D and innovation. If the emergence of a specific infrastructure of a European innovation policymaking system could be traced, monitoring is also a relevant task in relation to shed on light in the debate of what is a better choice for Europe: a centralized innovation policy or an open competition between national, regional or international innovation systems and initiatives, depending on the evolution of European countries to better R&D performance classification clusters.

The classification of these countries according to their R&D and innovation performance (based on the traditional R&D linear model and the postulates of the endogenous economic growth theories) can be also useful to assess their position as compared to other countries with respect to reaching competitiveness and long-term sustainability objectives.

Therefore, the model presented can be useful in terms of i) managerial tools supporting decision making in R&D, since it provides information about strengths and weaknesses in innovation activities and about their contribution to the country's competitiveness; ii) a benchmarking tool, allowing the EU to compare its own innovation activities, inputs and results with other countries of reference and with European Members national initiatives and iii) a monitoring and evaluation tool for institutional and governmental bodies, since it allows some critical innovation indicators to be monitored.

The proposed technique has to be considered a complementary tool to manage the extensive information provided by the number of European official reports of indicators and methodologies (European Innovation Scoreboard, EU2020 Monitoring reports, ERAWATCH, etc.). Our aim is to help experts and the main stakeholders to classify countries or predict their R&D performance in future years in a similar way to innovation scores and other methodologies with publicly available country-data using this model to identify the major contributing factors that could estimate their R&D and innovation performance with the fewest variables possible.

The results of our study confirm the findings presented in the above-mentioned previous works related to countries' scores based on R&D and innovation indicators. European Countries will also need to continue investing in education, in human capital, to increase necessary skills in knowledge-based economies since those skills are basic in the application of new industry and service technologies.

Although this research demonstrates that nine out of fifteen selected variables like those in our dataset could be used in a neural network model to predict the R&D performance of a country reasonably well and with fewer coefficients than other machine learning techniques, overlooked variables must be considered in future studies.

As the appropriate data are available, the implications of our model would appear to be worthwhile for future comparative-research in cross-country performance. Additionally, it would be of interest to amplify the research with data after the year 2008 because the period analyzed in the present study may not fully capture the impact of the economic and financial crisis on innovation performance due to the delay in the economic impacts and in data availability.

Acknowledgment

This work was partially subsidized by the Spanish Inter-Ministerial Commission of Science and Technology under Project TIN2011-22794, the European Regional Development fund, and the "Junta de Andalucía", Spain, under Project P08-TIC-374.

Appendix A

For a g ANN($\hat{\theta}$) classifier, a classification problem with Q classes and N training or testing patterns is considered, obtaining a $Q \times Q$ contingency or confusion matrix $M(g)$:

$$M(g) = \left\{ n_{ij} \sum_{i,j=1}^Q n_{ij} = N \right\} \quad (\text{A.1})$$

where n_{ij} represents the number of times the patterns are predicted by classifier g to be in class j when they really belong to class i .

Let us denote the number of patterns associated with class i by $f_i = \sum_{j=1}^Q n_{ij}$, $i = 1, \dots, Q$.

To evaluate the performance of the classifiers, three scalar measures are defined that take the elements of the confusion matrix into consideration from different points of view. Let $MS_i = n_{ii}/f_i$ be the number of patterns correctly predicted to be in class i with respect to the total number of patterns in i (sensitivity for class i). Therefore the sensitivity for class i estimates the probability of correctly predicting a class i example. From the above quantities the first comparison measure of a multiclass classifier define the sensitivity S of the classifier as the minimum value of sensitivities for each class $M = \min \{S_i; i = 1, \dots, Q\}$. The second measure is the correct classification rate or accuracy,

$$CCR = (1/N) \sum_{i=1}^Q n_{ii} \tag{A.2}$$

that is the rate of all the correct predictions. The third measure is Cohen's Kappa coefficient, as an association measure between the class that a priori belongs to a pattern and the assignation that a posteriori assigns the classifier to that class.

Thus, we consider the three-dimensional measure (CCR, MS, K) associated with classifier g . The measure tries to evaluate three features of a classifier throughout the generalization set: global performance in the whole dataset, the performance in each class and the degree of association between pattern distributions in classes before and after classifier application.

We consider standard sigmoidal ANN, also called Multilayer Perceptron (MLP) as the base classification model. It can overcome to a great extent the longer training time and the difficulty entailed in determining hidden layer units of backpropagation networks. For determining the best ANN model, we apply an evolutionary algorithm EA that is applied to find the basis functions:

$$B(\mathbf{x}, \mathbf{W}) = \{B_1(\mathbf{x}, \mathbf{w}_1), B_2(\mathbf{x}, \mathbf{w}_2), \dots, B_m(\mathbf{x}, \mathbf{w}_m)\} \tag{A.3}$$

corresponding to the nonlinear part of the discriminant functions, $f_i(x, \theta_i)$. We have to determine the number of basis functions m and the weight matrix W . To apply evolutionary neural network techniques, we consider an ANN with softmax outputs defined in (Eq. A4) and the standard structure: an input layer with a node for every input variable; a hidden layer with several sigmoidal nodes; and an output layer with $J-1$ nodes, where J is the number of classes.

The output layer of the ANN($\hat{\theta}$) classifier is interpreted from the point of view of probability, which considers the soft-max activation function:

$$g_l(x, \theta_l) = \frac{\exp f_l(x, \theta_l)}{\sum_{l=1}^Q \exp f_l(x, \theta_l)} \text{ for } l = 1, \dots, Q \tag{A.4}$$

where $g_l(x, \theta_l)$ is the probability a pattern x has of belonging to class l , $\theta_l = (\beta_1, w_1, \dots, w_M)$, $\beta_1 = (\beta_1^1, \dots, \beta_1^M)$, is the l -th vector of weights of the output node, M is the number of hidden nodes, $w_j = (w_0^j, \dots, w_k^j)$ for $j = 1, \dots, M$, is the vector of input weights of the hidden node j , and $f_l(x, \theta_l)$ is the output of the l -th output node for pattern x given by:

$$f_l(x, \theta_l) = \beta_0^l + \sum_{j=1}^M \beta_j^l \sigma_j \left(w_0^j + \sum_{i=1}^k w_i^j x_i \right) \text{ for } l = 1, \dots, Q - 1 \tag{A.5}$$

$$f_Q(x, \theta_Q) = 0$$

where $\sigma(\bullet)$ is the sigmoidal activation function of the nodes in the hidden layer.

The classification rule coincides with the optimal Bayes rule. Thus the classification rule makes an individual be assigned to that class which has the maximum probability, given vector measurement x

$$C(\mathbf{x}) = \hat{l}, \text{ where } \hat{l} = \arg \max_l g_l(\mathbf{x}, \hat{\theta}_l), \text{ for } l = 1, \dots, Q \tag{A.6}$$

Appendix B

The basic framework of the EA employed in this work is the following: the search begins with an initial population of ANNs and the population is updated in each iteration using a population-update algorithm, which evolves both the structure and the weights. The population is subject to the operations of replication and mutation. Crossover is not used due to its potential disadvantages in evolving artificial networks [64]. The algorithm evolves architectures and connection weights simultaneously for each individual that is a fully specified Evolutionary Sigmoidal Unit Neural Network, (ESUNN). ANNs are represented using an object-oriented approach and the algorithm deals directly with the ESUNN phenotype. Each connection is specified by a binary value indicating if the connection exists, and a real value representing its weight. As the crossover is not considered, this object-oriented representation does not assume a fixed order between different hidden nodes.

The general structure of the EA is the following:

- (1) Generate a random population of size N .
- (2) Repeat until the stopping criterion is fulfilled
 - (a) Calculate the fitness of every individual in the population.
 - (b) Rank the individuals with respect to their fitness.
 - (c) The best individual is copied into the new population.
 - (d) The best 10% of population individuals are replicated and substitute the worst 10% of individuals.

To that intermediate population we:

- Apply parametric mutation to the best (p_m)% of individuals.
- Apply structural mutation to the remaining $(1-p_m)$ % of individuals.

The weights are initialized the same way, regardless of the hidden node (product nodes, PU, or sigmoidal nodes, SU). Weights are assigned using uniform distribution defined throughout two intervals: $[-5,5]$ for connections between the input layer and hidden layer and for all kinds of neurons, $[-10,10]$ for connections between the hidden layer and the output layer.

Parametric mutation is accomplished for each weight of the ANN with Gaussian noise $w_{ji}(t+1)w_{ji}(t) + \xi(t)$, where $\xi(t)$ represents a one-dimensional normally distributed random variable. The variance of the normal distribution is updated throughout the evolution of the algorithm applying the simplest heuristic 1/5 success rule of Rechenberg [48]. On the other hand, structural mutation implies a modification in the ANN structure and allows explorations of different regions in the search space while helping to keep up the diversity of the population. There are five different structural mutations: node deletion, connection deletion, node addition, connection addition and node fusion. These five mutations are applied sequentially to each network. For more details about the general structure of the EA and the parametric and structural mutations, readers can see [65,66].

The parameters used in the evolutionary algorithm are $N = 1000$, the number of generations 350, while the maximum number of hidden nodes is 3. The number of nodes that can be added or deleted in a structural mutation is within the [1,2] interval. The number of connections that can be added or deleted in a structural mutation is within the interval [1,4]. Parametric mutation is applied to the best ($p_m = 10$)% of individuals. The stop criterion is reached when the following condition is fulfilled: for 20 generations there is no improvement either in the average performance of the best (p_m)% of the population or in the fitness of the best individual.

The algorithm will evolve as the fitness function $A(g) = \frac{1}{1 + E(g, \theta_1)}$, where $E(g, \theta_1)$ is the cross-entropy error function or Q -class multinomial deviance, given by:

$$E(g, \theta_1) = -\frac{1}{N} \sum_{n=1}^N \sum_{l=1}^Q y_n^{(l)} \log g_l(\mathbf{x}_n, \theta_1) \tag{B.1}$$

Table A1. Confusion matrix of a classifier.

Class	1	2	...	Q	Priors
1	n_{11}	n_{12}	...	n_{1Q}	f_1
2	n_{21}	n_{22}	...	n_{2Q}	f_2
...
Q	n_{Q1}	n_{Q2}	...	n_{QQ}	f_Q

References

[1] R.J. Watts, A.L. Porter, Innovation forecasting, Technol. Forecast. Soc. Chang. 56 (1) (1997) 25–47.
 [2] G. Fahrenkrog, W. Polt, J. Rojo, A. Tübke, K. Zinöcker, RTD Evaluation Toolbox, Assessing the Socioeconomic Impact of RTD Policies, Joint Research Centre, Institute for Prospective Technological Studies, Seville, 2002.
 [3] T. Loikkanen, T. Ahlqvist, P. Pellinen, The role of the technology barometer in assessing the performance of the national innovation system, Technol. Forecast. Soc. Chang. 76 (9) (2009) 1177–1186.
 [4] A. Conte, An analysis of the efficiency of public spending and national policies in the area of R&D, European Commission, Occasional Papers, 54, 2009.
 [5] J. Cheng Guan, C. Kam Mok, R.C.M. Yam, K.S. Chin, K. Fai Pun, Technology transfer and innovation performance: evidence from Chinese firms, Technol. Forecast. Soc. Chang. 73 (6) (2006) 666–678.
 [6] I.C. Kerssens-van Drongelen, A. Cook, Design principles for the development of measurement systems for research and development processes, R&D Manag. 27 (4) (1997) 345–357.
 [7] D. Archibugi, A. Coco, Measuring technological capabilities at the country level: a survey and a menu for choice, Res. Policy 34 (2005) 175–194.
 [8] G. Cerulli, A. Filippetti, The complementary nature of technological capabilities: Measurement and robustness issues, Technol. Forecast. Soc. Chang. 79 (2012) 875–887.
 [9] European Commission, European Innovation Scoreboard 2011, Comparative Analysis of Innovation Performance, European Commission, DG Enterprise, Brussels, 2012.
 [10] P.A. Gutiérrez, M.J. Segovia-Vargas, S. Salcedo-Sanz, C. Hervás-Martínez, A. Sanchis, J.A. Portilla-Figueras, F. Fernández-Navarro, Hybridizing logistic regression with product unit and RBF networks for accurate detection and prediction of banking crises, Omega, Int. J. Manag. Sci. 38 (2010) 333–344.
 [11] R. Trippi, E. Turban, Neural Networks in Finance and Investing, Probus Publishing Co., Cambridge, 1994.
 [12] C. Charalambous, A. Charitou, F. Kaourou, Comparative analysis of artificial neural network models: application in bankruptcy prediction, Ann. Oper. Res. 99 (2000) 403–425.

- [13] T.S. Gruca, B.R. Klemz, Using Neural Networks to Identify Competitive Market Structures from Aggregate Market Response Data, *Omega, Int. J. Manag. Sci.* 26 (1998) 49–62.
- [14] J. Bode, Decision support with neural networks in the management of research and development: concepts and application to cost estimation, *Inf. Manag.* 34 (1) (1998) 33–40.
- [15] J. Becker, M. Prischmann, Supporting the design process with neural networks. A complex application of cooperating neural networks and its implementation, *J. Inf. Sci. Technol.* 3 (1) (1993) 79–95.
- [16] V. Hájková, P. Hájek, Analysis of regional innovation systems by neural networks and cluster analysis, *Proc. Communication and Management in Technological Innovation and Academic Globalization (COMATIA'10)*, 2010, pp. 46–51.
- [17] S.H. Zanak, I. Becerra-Fernandez, Competitiveness of nations: a knowledge discovery examination, *Eur. J. Oper. Res.* 166 (2005) 185–211.
- [18] K.A. Smith, J.N.D. Gupta, Neural networks in business: techniques and applications for the operations researcher, *Comput. Oper. Res.* 27 (2000) 1023–1044.
- [19] R. Durbin, D. Rumelhart, Products units: a computationally powerful and biologically plausible extension to backpropagation networks, *Neural Comput.* 1 (1989) 133–142.
- [20] V. Ojanen, O. Vuola, Categorising the Measures and Evaluation Methods of R&D Performance - A State-of-the-art Review on R&D Performance Analysis, *Telecom Business Research Centre Working Paper 16*, Lappeenranta University of Technology, Finland, 2003.
- [21] S. Hayken, *A comprehensive foundation. Neural Networks*, Macmillan, Nueva York, 1994.
- [22] R.P. Lippman, Review of neural networks for speech recognition, *Neural Comput.* 1 (1) (1989) 1–38.
- [23] OECD, *Proposed Standard Practice for Surveys of Research and Experimental Development-Frascati Manual*, 5th ed., OECD, Paris, 1994.
- [24] F.J. O'Donnell, A.H.B. Duffy, Modeling design development performance, *Int. J. Oper. Prod. Man.* 22 (11) (2002) 1198–1222.
- [25] V. Chiesa, F. Frattini, V. Lazzarotti, R. Manzini, Performance measurement in R&D: exploring the interplay between measurement objectives, dimensions of performance and contextual factors, *R&D Manag.* 39 (5) (2009) 487–519.
- [26] M. Coccia, What is the optimal rate of R&D investment to maximize productivity growth? *Technol. Forecast. Soc. Chang.* 76 (3) (2009) 433–446.
- [27] Z. Griliches, *R&D and Productivity: The Econometric Evidence*, 1st ed. The University of Chicago Press, Chicago, 1998.
- [28] G.M. Grossman, E. Helpman, *Innovation and growth in the global economy*, MIT Press, Cambridge, 1991.
- [29] D. Guellec, B. van Pottelsberghe de la Potterie, R&D and productivity growth: panel data analysis of 16 OECD countries, *OECD Econ. Stud.* 33 (3) (2001) 103–126.
- [30] E.C. Wang, R&D efficiency and economic performance: a cross-country analysis using the stochastic frontier approach, *J. Policy Model.* 29 (2007) 345–360.
- [31] M.E. Porter, *The Competitive Advantage of Nations*, Free Press, New York, 1990.
- [32] R.L. Brinkman, J.E. Brinkman, The new growth theories: a cultural and social addendum, *Int. J. Soc. Econ.* 28 (5,6,7) (2001) 506–526.
- [33] P.M. Romer, Endogenous technological change, *J. Polit. Econ.* 98 (5) (1990) 71–101.
- [34] P.S. Segerstrom, T.C.A. Anant, E. Dinopoulos, A Schumpeterian product life cycle, *Am. Econ. Rev.* 80 (1990) 1077–1091.
- [35] P. Aghion, P. Howitt, A model of growth through creative destruction, *Econometrica* 60 (2) (1992) 323–351.
- [36] P.M. Romer, Should the government subsidize supply or demand in the market for scientists and engineers? in: A.B. Jaffe, J. Lerner, S. Stern (Eds.), *Innovation Policy and the Economy*, National Bureau of Economic Research, Inc., Massachusetts, MA, 2001.
- [37] A. Alcouffe, T. Kuhn, Schumpeterian endogenous growth theory and evolutionary economics, *J. Evol. Econ.* 14 (2) (2004) 223–236.
- [38] S.J. Kline, N. Rosenberg, An overview of innovation, in: R. Landau, N. Rosenberg (Eds.), *The Positive Sum Strategy: Harnessing Technology for Economic Growth*, National Academy Press, Washington, D.C., 1986.
- [39] OECD, *Technology and the economy: the key relationships*, OECD, Paris, 1992.
- [40] T. Padmore, H. Schuetze, H. Gibson, Modeling systems of innovation: an enterprise-centered view, *Res. Policy* 26 (6) (1998) 605–624.
- [41] B. Godin, The historical construction of an analytical framework, *Sci. Technol. Hum. Values* 31 (2006) 639–667.
- [42] A. Filippetta, D. Archibugia, Innovation in times of crisis: national systems of innovation, structure, and demand, *Res. Policy* 40 (2011) 179–192.
- [43] Y. De Smet, Towards multicriteria clustering: an extension of the k-means algorithm, *Eur. J. Oper. Res.* 158 (2) (2004) 390–398.
- [44] J.A. Hartigan, M.A. Wong, Algorithm AS 136: a k-means clustering algorithm, *J. R. Stat. Soc. Ser. C* 28 (1979) 100–108.
- [45] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, I. Witten, *The WEKA Data Mining Software: An update*, *ACM SIGKDD Explorations News*, 11(1), 2009, pp. 10–18.
- [46] C.M. Bishop, *Neural Networks for Pattern Recognition*, Clarendon Press, Oxford, 1995.
- [47] D.E. Goldberg, The design of innovation: lessons from genetic algorithms, lessons for the real world, *Technol. Forecast. Soc. Chang.* 64 (2000) 7–12.
- [48] J.H. Holland, *Adaptation in natural and artificial systems*, The University of Michigan Press, Ann Arbor, 1975.
- [49] I. Rechenberg, *Evolutionstrategie: Optimierung technischer Systeme nach Prinzipien der Biologischen Evolution*, Franmann-Holzboog Verlag, Stuttgart, 1973.
- [50] L.J. Fogel, A.J. Owens, M.J. Walsh, *Artificial Intelligence through Simulated Evolution*, John Wiley, New York, NY, 1966.
- [51] OECD, *Science, Technology and Innovation Indicators in a Changing World: Responding to Policy Needs*, OECD, Paris, 2007.
- [52] European Commission, *PRO INNO Europe*, 2011 <http://www.proinno-europe.eu/>, official web site. Last accessed: February 25, 2011.
- [53] S.F. Buck, A method of estimation of missing values in multivariate data suitable for use with an electronic computer, *J. R. Stat. Soc. B Met.* 22 (1960) 302–306.
- [54] OECD, *Main Science and Technology Indicators*, 2nd ed., OECD, Paris, 2010.
- [55] Z. Griliches, Patent Statistics as Economic Indicators: A Survey, *J. Econ. Lit.*, 28(4), American Economic Association, 1990, pp. 1661–1707.
- [56] M. Norusis, *SPSS 15.0 Advanced Statistical Procedures Companion*, Prentice Hall Press, Upper Saddle River, 2007.
- [57] N. Landwehr, M. Hall, E. Frank, Logistic model trees, *Mach. Learn.* 59 (2005) 161–205.
- [58] D.J. Sheskin, *Handbook of parametric and nonparametric statistical procedures*, 5th ed. Chapman and Hall/CRC, Boca Raton, 2011.
- [59] C. Soudien, Education in the network age: globalisation, development and the World Bank, *Int. J. Educ. Dev.* 22 (5) (2002) 439–450.
- [60] R.E. Lucas, On the mechanics of economic development, *J. Monet. Econ.* 22 (1) (1988) 3–42.
- [61] N. Gemmill, Endogenous growth, the Solow model and human capital, *Econ. Plan.* 28 (1995) 169–183.
- [62] J.L. Furman, M.E. Porter, S. Stern, The determinants of national innovative capacity, *Res. Policy* 31 (2002) 899–933.
- [63] N. Varsakelis, Education, political institutions and innovative activity: a cross-country empirical investigation, *Res. Policy* 35 (7) (2000) 1083–1090.
- [64] S. Kuhlmann, Scenarios of technology and innovation policies in Europe: investigating future governance, *Technol. Forecast. Soc. Chang.* 70 (7) (2003) 619–637.
- [65] X. Yao, Y. Liu, A new evolutionary system for evolving artificial neural networks, *IEEE Trans. Neural Netw.* 8 (1997) 694–713.
- [66] C. Hervás-Martínez, F.J. Martínez-Estudillo, Logistic regression using covariates obtained by product-unit neural network models, *Pattern Recogn.* 40 (2007) 52–64.
- [67] A.C. Martínez-Estudillo, F.J. Martínez-Estudillo, C. Hervás-Martínez, N. García-Pedrajas, Evolutionary product unit based neural networks for regression, *Neural Netw.* 19 (2006) 477–486.

Mónica de la Paz-Marín. She received a B.S. degree in Business Administration from the University of Córdoba (ETEA, Business Administration Faculty) in 1996, a Bachelor of Law Degree (B.L.) and she is currently a Ph.D candidate in European Union Doctoral Programme. She initiated and continues her professional activities at the University of Córdoba in the Quality Management Service and she has worked as advisor in the Regional Government of Andalusia (Regional Ministry of Economy, Innovation and Science, Provincial Delegation). Her research interests are focused on the assessment of Research and Development (R&D) Public Policies and innovation, modeling tools for public decision-making support.

Pilar Campoy-Muñoz. Pilar Campoy-Muñoz received the B.S. degree in Business Administration in 2005 and M.S degree in Research Methods on Economics and Business Sciences in 2011, both from the University of Córdoba, Spain. She is currently pursuing the Ph.D degree and working as researcher in the Department of Department of Management and Quantitative Methods, ETEA, Business Administration Faculty in the University of Córdoba, Spain. Her current research interests include economy and computational economics.

César Hervás-Martínez. He received the B.S. degree in statistics and operations research from the Universidad Complutense, Madrid, Spain, in 1978, and the Ph.D. degree in mathematics from the University of Seville, Seville, Spain, in 1986. He is currently a Professor of Computer Science and Artificial Intelligence in the Department of Computer Science and Numerical Analysis, University of Córdoba, Córdoba, Spain, and an Associate Professor in the Department of Quantitative Methods, School of Economics, University of Córdoba. His current research interests include neural networks, evolutionary computation, and the modeling of natural systems.

Addressing the EU sovereign ratings using an ordinal regression approach

Francisco Fernández-Navarro, Pilar Campoy-Muñoz, Mónica de la Paz-Marín, César Hervás-Martínez *Member, IEEE*, Xin Yao, *Fellow, IEEE*,

Abstract—The current European debt crisis has drawn considerable attention to credit rating agencies’ news about sovereign ratings. From a technical point of view, credit rating constitutes a typical ordinal regression problem because credit rating agencies generally present a scale of risk composed several categories. This fact motivated the use of an ordinal regression approach for addressing the problem of sovereign credit-rating in this paper. Therefore, the ranking of different classes will be taken into account for the design of the classifier. To do so, a novel model is introduced in order to replicate sovereign rating, based on the Negative Correlation Learning framework. The methodology is fully described in the paper, and applied to the classification of the 27 European countries’ sovereign rating during the 2007-2010 period based on Standard and Poor’s reports. The proposed technique seems to be competitive and robust enough to classify the sovereign ratings reported by this agency when compared to other existing well-known ordinal and nominal methods.

Index Terms—Country risk detection, Neural Networks, Ordinal Regression, Negative Correlation Learning.

I. INTRODUCTION

Nowadays, Credit Rating Agencies (CRAs) play a crucial role in global financial markets due to the production of relevant credit information and its distribution to market participants, diminishing asymmetric information between investors and issuers. A wide range of stakeholders, including investors, issuers and policy makers use the rating agencies’ information in their decision-making processes [1]. Moreover, Basel Capital Accord III has driven a renewed interest in credit ratings, allowing banks to employ them to determine the default probabilities of their debtors by calculating an adequate

This work was supported in part by the Spanish Inter-Ministerial Commission of Science and Technology under Project TIN2011-22794, the European Regional Development fund, and the “Junta de Andalucía” (Spain), under Project P2011-TIC-7508. Xin Yao was partially supported by European Commission through its FP7 grant (No. 270428). Francisco Fernández-Navarro’s research has been subsidized by the “Junta de Andalucía” Predoctoral Program, grant reference P08-TIC-3745.

F. Fernández-Navarro is with Advanced Concepts Team, European Space Research and Technology Centre (ESTEC), European Space Agency (ESA), Noordwijk, Netherlands, e-mail: francisco.fernandez.navarro@esa.int

F. Fernández-Navarro and C. Hervás-Martínez are with the Department of Computer Science and Numerical Analysis, University of Córdoba, Campus de Rabanales, C2 building, 14004 - Córdoba, Spain, e-mail: i22fenaf@uco.es, chervas@uco.es.

Pilar Campoy-Muñoz and Mónica-de la Paz are with the Department of Economics, ETEA, University of Córdoba, Escritor Castilla Aguayo 4, 14004 - Córdoba, Spain, e-mail: mpcampoy@gmail.com, ma2pamam@uco.es.

Xin Yao is with the Centre of Excellence for Research in Computational Intelligence and Applications (CERCIA), School of Computer Science, University of Birmingham, Birmingham B15 2TT, U.K., e-mail: X.Yao@cs.bham.ac.uk

amount of capital to cover their credit risks (see, inter alia, [2], [3]).

These CRAs have been in the spotlight during the ongoing European sovereign debt crisis. This crisis has been a theater of sovereign credit rating downgrades, a widening of sovereign bond and credit default swap (CDS) spreads and pressures on stock markets. Interestingly, financial markets throughout the Euro zone have been under pressure although credit rating actions concentrated on a few countries such as Greece, Iceland, Ireland, Portugal and Spain [4]. The current debate about CRAs echoes previous discussions during the 1997-98 Asian crisis [4], fuels the earliest criticisms about the correct assignment of ratings by these agencies [5], [6] and highlights the inherent conflicts of interest within their business model that is characterised by a lack of transparency and poor communication [7]. To overcome those limitations and promote competition among agencies, formal regulation of the credit rating industry was introduced by the European Union in April 2009 and May 2011 and a proposal for a third regulation was published by the European Commission in November 2011. However, currently the three largest CRAs, namely Moody’s, S&P and Fitch, continue dominating the sovereign ratings business, which is relatively new; S&P presented its first set of foreign currency sovereign ratings in January 1961 and Moody’s followed in January 1974 [8].

In this regard, sovereign credit ratings could be considered a condensed assessment of a government’s ability and willingness to repay its public debt both in principal and in interests in a timely fashion [9]. They are ordinal measures that can both reflect the current financial position of sovereign nations and provide an overview of their future financial positions. Thus, the relevance of sovereign credit ratings is reflected in several ways: (i) they improve the capability of countries to access international markets and to attract foreign investments [10]; (ii) they face the growing international diversification of investors’ portfolios and their needs for more accurate and greater amounts of information regarding country risk due to the impacts that a re-assessment of a country’s risk can have on its portfolios [11], [12]; (iii) they can represent a ceiling for the ratings assigned to non-sovereign entities within a country [13], since there remains a “sovereign ceiling lite” [14] even though the sovereign ceiling rule has recently been eliminated by main agencies; (iv) and, finally, the credit rating events have a notable effect on financial market developments, as negative news impact the country’s bond and stock market [15], [11] and cause spillovers to other countries’ equity and bond markets [16], while upgrades have limited or negligible effect

[17], [18]. Rating agencies' signals also triggered foreign exchange market reactions [19].

To date, a wide range of classification techniques have been proposed in the literature on sovereign credit-scoring. Recent empirical studies have emerged from two research lines. The first group focuses on the application of statistical techniques such as ordinary least squares [13], [20], [21], discriminant analysis [22] and ordered response models [23], [24], [25]. The second group makes use of artificial intelligence methods, especially machine learning techniques such as neural networks [26], [27] or support vector machines [28]. However, recent studies suggested that the combination of multiple classifiers, which is known as ensemble learning, may lead to better performance [29], [30].

In this paper, the sovereign rating problem has been addressed using an ordinal regression approach because of the ordinal nature of the dependent variable. Ordinal regression algorithms are intended to take advantage of this order information in order to improve the classification performance. An ordinal categorical variable may be inherently a continuous variable termed as a grouped continuous variable [31]. On the other hand, an ordinal categorical variable such as attitude or opinion can be considered a manifestation of an underlying continuous variable (see, e.g., [32]). In both cases, it is sensible to relate the observed ordinal categorical variable to an underlying continuous variable.

A more commonly used idea is to transform the ordinal regression problem into a multinomial classification one, or to add additional constraints to traditional classification formations. In this sense, Crammer and Singer [33] generalized the online perceptron algorithm with multiple thresholds in order to seek the direction and thresholds for ordinal regression. Shashua and Levin [34] proposed two large-margin principles, namely, a fixed-margin principle and a sum of margins principle, to handle direction and multiple thresholds. More recently, Chu together with Ghahramani presented a probabilistic kernel approach to ordinal regression based on Gaussian processes [35] and, together with Keerthi, two new support vector machine approaches for ordinal regression, which optimized multiple thresholds to define parallel discriminant hyperplanes for ordinal scales [36]. In Cardoso and Pinto da Costa [37], a novel framework for ordered classes, based on replicating the dataset, was introduced in the context of support vector machines. Finally, Sun et al. [38] have proposed a new support vector approach for ordinal regression which maximizes the sum of the margins for the computation of parallel discriminant hyperplanes because they argue that, although SVM-based methods have shown great promise in ordinal regression, they do have some drawbacks.

On the other hand, an ensemble of multiple classifiers is expected to reduce generalisation error by considering the opinions from multiple classifiers. Therefore, diversity becomes an important issue to take into account [39], [40]. If every classifier had the same opinion, then the construction of multiple classifiers would become meaningless. Diversity can be described as the degree to which classifiers come to different decisions about the same problem. This degree allows voted accuracy to be greater than the one achieved by a single

classifier.

Concerning ordinal regression problems, there are several ensemble related approaches in the literature. These approaches transform the ordinal classification problem into a nested binary classification one, and then combine the resulting classifier predictions to obtain the final decision. For instance, Frank and Hall [41] proposed a general algorithm that enables standard binary classifiers to make use of order information in attributes. Waegeman and Boullart [42] imposed explicit weights over the patterns of each binary classifier, in such a way that errors on training patterns are penalized proportionally to the absolute difference between their rank and its ranking. However, both approaches take a base binary classification algorithm rather than a base ordinal regression one.

Negative Correlation Learning (NCL) [43], [44] is an ensemble learning method of neural networks that uses different mechanisms to induce diversity. It has shown a number of empirical successes in various applications and competitive results. NCL try to induce diversity directly by changing its error function into a "diversity-encouraging" one. In this work, the NCL methodology was adapted to ordinal regression problems and validated using a EU sovereign-rating dataset. At the best of the authors' knowledge, this research work would be the first adaptation of the NCL to ordinal regression problems.

The aim of this paper is to ascertain whether the NCL methodology can improve the classification accuracy when compared to the more common methods employed in machine learning. To achieve this aim, this work carries out an empirical analysis of foreign-currency sovereign-debt ratings with the data on 27 EU member countries provided by S&P during the financial crisis between 2007 and 2010. The S&P foreign-currency long-term rating history is selected instead of other agencies' historical ratings due to data availability. Moreover, S&P seems more prone to produce more rating revisions and to lead the re-rating of other agencies [45], [18]. Foreign-currency rating announcements by S&P also seem to convey a greater stock market impact on the country itself and do not seem to be fully expected by the market [11]

The remainder of our paper proceeds as follows. The next section reviews the relevant literature on classification methodologies in rating assignments. Secondly, the methodology applied in this paper and its motivation are described; thirdly, the experimental study is carried out with the description of the variables and the database employed, and the presentation of empirical results, the discussion and, finally, the main conclusions about these experimental results.

II. LITERATURE REVIEW

Although a number of methodologies have been employed in the empirical modelling of credit ratings, the main focus was on corporate bonds rather than on sovereign risk [26]. In the latter case, the earliest approach applies statistical techniques such as linear regression methods [13], [20], [21], ordered response models [23], [24], [25] and discriminant analysis (DA), derived from previous works of Frank and Cline and

Sargen [46], [22] or Carleton and Lerner [47], which refer to sub-sovereign entities.

While logit, probit and DA models require such assumptions as the normality of variable distribution and independent predictors, non-parametric and non-linear models such as artificial neural networks (ANNs) [48], [49], [50] do not rely on these assumptions that are adopted to turn traditional statistical methods into more tractable ones. Comparing ANNs and statistical models for sovereign risk prediction, most related studies pointed out that ANNs outperformed statistical models. Bennell [26] demonstrates that ANNs represent improved technology for calibrating and predicting sovereign ratings compared to ordered probit modelling. In this regard, Yim and Mitchell [27] also prove that hybrid ANNs produce better results than ordered response models and discriminant analysis. However, some authors pointed out that ANNs have difficulty in determining the difference between adjacent rating classes [51]. In fact, Huang et al. [52] found out that the probabilities of misclassification within one class away from the real one were over 90%. Since ANNs do not require the prior specification of theoretical model, it is argued that their functional form is totally unrelated to the economic theory [53]. ANNs are constructed from the data and not from an economic theory. However, when economic theory presents difficulties to be successfully implemented, the use of ANNs models can be an acceptable option to have good predictions on the EU sovereign ratings (while the economic community researches new alternatives). ANNs also have other technical drawbacks including the risk of over-fitting, difficulty in determining the values of control parameters and the number of processing elements in the hidden nodes [28].

The Support Vector Machines (SVM) algorithm has recently become a solution often employed to solve prediction problems because of its robustness and high accuracy. An SVM solution is globally optimal because SVMs seek to minimize structural risk. Conversely, the solutions found by ANN models tend to fall into local optimum because they seek to minimize empirical risk. SVMs were originally designed for two-class tasks and therefore are not naturally geared for multi-class classifications, which apply to credit ratings: even so researchers have made attempts to extend the original SVM to multi-class classification (MSVM). The prior studies that applied MSVMs to credit rating issues were not designed to reflect the ordinal nature of this domain (e.g. [52]) although recently there have been developments which take ordinal characteristics into account to efficiently and effectively handle multiple ordinal classes such as [28]. In this study, the MSVM approach improves the performance of classification as compared to other typical multi-class classification techniques. However, Huang et al. [52] pointed out that there are two major aspects that should be considered cautiously when SVM methods are applied to solve rating problems. The first is how to select the optimal input feature and the second is how to set good kernel parameters.

As can be seen, both statistical and machine learning techniques have been explored for credit rating, although there are no consistent conclusions about which performs better. Accordingly, most recent studies suggest combining multiple

classifiers, i.e., ensemble learning, which have shown better performance than any individual methods. Ensemble models in general have been successfully used for classification and regression. They not only introduce more stable predictions through linear combination, but also provide sufficient power for approximating complicated target functions [54]. Ensembles are trending in financial topics such as the prediction of financial crisis or the credit scoring due to these techniques outperform single classification techniques, but to the best of our knowledge this technique has not been applied to sovereign credit rating problems [55].

The next section introduces the motivation for using ensembles in ordinal regression problems and then the proposed ensemble methodology is described.

III. ENSEMBLE LEARNING SUITABILITY FOR ORDINAL REGRESSION PROBLEMS

The aim of this paper is to validate ensemble models in the ordinal regression problem of classifying the sovereign ratings provided by the S&P international agency. Therefore, this section describes the motivation for the use of ensemble models in ordinal regression problems.

In a regression problem, we can define why and how differences between individual predictor outputs contribute to overall ensemble accuracy. However, in a classification problem, there is no such neat theory. There is a subtle point here, often overlooked. Difficulties in quantifying classification error diversity are not intrinsic to ensemble-tackling classification problems. It is possible to reformulate any classification problem as a regression one by approximating posterior class probabilities [56]. For the regression context, the question can be clearly phrased as “how could we quantify diversity when our predictors are output real-valued numbers and are combined by a convex combination?”. For the case of Tumer and Ghosh [56] study, the question is the same, except for the fact that the “real-valued” numbers are probabilities. A much harder question appears when we are restricted in that our predictors can only output discrete class labels. In this case, the outputs do not present any intrinsic order among them, and so the concept of “covariance” is not so simple.

A step toward understanding this question can be taken by considering where the bias-variance-covariance decomposition comes from: it falls neatly out of the bias-variance decomposition of the ensemble error. However, when our classification of a pattern is either correct or incorrect, we have a zero-one loss function.

Tumer and Ghosh [56], [57] provided a theoretical framework for analysing the simple averaging combination rule when our predictor outputs are estimates of the posterior probabilities of each class.

For an input variable of only one dimension, x , the solid curves show the true posterior probabilities of \mathcal{C}_1 and \mathcal{C}_2 , these are $P(\mathcal{C}_1)$ and $P(\mathcal{C}_2)$, respectively (Fig. 1). The dotted curves show estimates of the posterior probabilities, from one of the individuals of the ensemble, these are $\hat{P}(\mathcal{C}_1)$ and $\hat{P}(\mathcal{C}_2)$. The solid vertical line at θ indicates the optimal decision boundary. The dark shaded area, called the Bayes error, is an irreducible

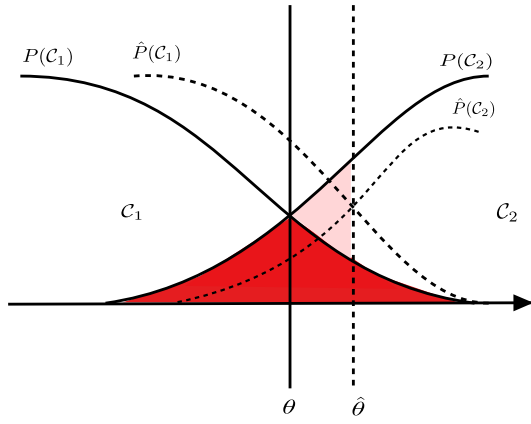


Fig. 1: Tumer and Ghosh's analysis of ordinal regression error

quantity. The dotted vertical line at $\hat{\theta}$ indicates the boundary placed by the individual, which is a certain distance from the optimal. The light shaded area indicates the added error that the predictor makes in addition to the Bayes error. The individual i approximates the posterior probability of \mathcal{C}_1 as:

$$\hat{P}_i(\mathcal{C}_1|x) = P_i(\mathcal{C}_1|x) + \varepsilon_i(\mathcal{C}_1|x), \quad (1)$$

where $P_i(\mathcal{C}_1|x)$ is the true posterior probability of \mathcal{C}_1 and $\varepsilon_i(\mathcal{C}_1|x)$ is the estimation error. Let us assume the estimation errors on different classes \mathcal{C}_1 and \mathcal{C}_2 are independent and identically distributed random variables with zero mean and variance σ_{ε_i} . The individual's expected added error of classifying classes \mathcal{C}_1 and \mathcal{C}_2 is defined as:

$$E_{\text{add},i} = \frac{2\sigma_{\varepsilon_i}^2}{P'_i(\mathcal{C}_1|x) - P'_i(\mathcal{C}_2|x)}, \quad (2)$$

where $P'_i(\mathcal{C}_1|x)$ and $P'_i(\mathcal{C}_2|x)$ are the derivatives of the true posterior probability of classes \mathcal{C}_1 and \mathcal{C}_2 , respectively. If the decision boundary was determined by an ensemble of individuals, the authors show that the expected added error of the ensemble is:

$$E_{\text{add}}^{\text{ens}} = E_{\text{add}} \left(\frac{1 + \delta(M-1)}{M} \right), \quad (3)$$

where M is the number of classifiers, E_{add} is the expected added error of the individual classifiers (they are assumed to have the same error), and δ is a correlation coefficient. If δ is zero (the classifiers in the ensemble are statistically independent because they are normally distributed), we have $E_{\text{add}}^{\text{ens}} = \frac{1}{M} E_{\text{add}}$, i.e. the error of the ensemble will be M times smaller than the error of the individuals. If δ is 1, i.e. perfect correlation, then the error of the ensemble will just be equal to the errors of the individuals ($E_{\text{add}}^{\text{ens}} = E_{\text{add}}$).

Taking the work of Tumer and Ghosh into consideration, it seems reasonable to think that in the case of ordinal regression, we should take the correlation between individuals into account. However, to our knowledge, there is no previous work where individuals were designed taking into account the correlation among them.

IV. DESCRIPTION OF THE NEGATIVE CORRELATION METHODOLOGY IN ORDINAL REGRESSION

Ordinal regression is similar to regression because the labels in both \mathcal{Y}_r and \mathbb{R} represent ordinal information. Nevertheless, unlike the real-valued regression labels in \mathbb{R} , the discrete ranks in \mathcal{Y}_r do not carry metric information. That is, ordinal ranking deals with qualitative ranks while regression focuses on quantitative, real-valued outcomes. To model ordinal ranking problems from a regression perspective, it is often assumed that some underlying real-valued outcomes exist, although they are unobservable [31]. The hidden local scales "around" different ranks can be quite different, but the actual scale (metric) information is not encoded in the ranks.

Under the above assumption, each rank represents a contiguous interval on the real line. Then, ordinal ranking can be approached by the algorithm described in Fig. 2.

Threshold Regression Algorithm:

- 1: Estimate a potential function $f(\mathbf{x})$ that predicts (a monotonic transform of) the real-valued outcomes.
- 2: Determine a threshold vector $\theta \in \mathbb{R}^{J-1}$ to represent the intervals in the range of $f(\mathbf{x})$, where $\theta_1 \leq \theta_2 \leq \dots \leq \theta_{J-1}$

Fig. 2: Threshold Regression Algorithm

In the threshold regression algorithm, the potential function tries to uncover the nature of the assumed underlying outcome, and the threshold vector estimates the possibly different scales around different ranks. The two abstract steps of the algorithm are indeed taken by many existing ordinal ranking algorithms. For instance, in the GPOR algorithm of Chu and Ghahramani [35], the potential function $f(\mathbf{x})$ is assumed to follow a Gaussian process, and the threshold vector θ is determined by Bayesian inference with respect to some noise distribution. In the PRank algorithm of Crammer and Singer [33], the potential function $f(\mathbf{x})$ is taken to be a linear function and the pair $\langle f(\mathbf{x}), \theta \rangle$ are updated simultaneously. Some other algorithms are based on SVM, and they work on potential functions of the form $f_v(\mathbf{x}) = \langle v, \phi(\mathbf{x}) \rangle$, where $\phi(\mathbf{x})$ maps $\mathbf{x} \in \mathbb{R}^k$ to some Hilbert space [36].

In our proposal the thresholds are fixed a priori and they are not modified during the training procedure. Therefore, each threshold is defined as: $\theta_j = \theta_1 + (j-1)\vartheta$, where ϑ represents the width of the intervals, $j = 2, \dots, J-1$ and $\theta_1 \in \mathbb{R}$ (in this approach, the parameter ϑ takes a constant value for simplicity). That is due to the fact that to measure the diversity in the proposed model, it becomes necessary that all individuals in the ensemble project the patterns using the same thresholds setting (to measure the diversity in a common space). Furthermore, a nonlinear model was applied (ensemble of ANNs). Considering this kind of models, it is relatively easy to converge towards an acceptable solution without modifying the thresholds. In addition, an optimal projection generated with adaptive thresholds can be transformed into an equivalent one with fixed thresholds by linearly scaling the patterns class by class.

As previously stated, a standard multilayer perceptron MLP

(non-linear ranking function) is considered as the potential function $f(\mathbf{x})$ of each individual. Finally, the potential ensemble function is determined by simple averaging:

$$\bar{f}(\mathbf{x}) = \frac{1}{M} \sum_{i=1}^M f_i(\mathbf{x}) \quad (4)$$

where $\bar{f}(\mathbf{x})$ is the average output of the whole ensemble of M networks. In this way, the threshold ensemble model determines the class of each pattern as:

$$r_{(\bar{f}(\mathbf{x}), \theta)} = \min\{k : \bar{f}(\mathbf{x}) \leq \theta_k\} = \max\{k : \bar{f}(\mathbf{x}) > \theta_{k-1}\} \quad (5)$$

It is well known that a multi-criteria search for an ensemble that maximizes both accuracy and diversity leads to more accurate ensembles than optimizing a single criterion. Intuitively, we expect the potential value $f(\mathbf{x})$ to be in the desired interval $(\theta_{i-1}, \theta_i]$, and we want $f(\mathbf{x})$ to be far from the boundaries (thresholds).

Fig. 3 shows two threshold ensemble models. In the first case, the predictions are very close to the thresholds. In this case, if there is noise in the data, it could cause changes in the predictions. In the second case, the predictions are as far as possible from the thresholds. In this case, if there is noise in the data, the changes in predictions would be lower.

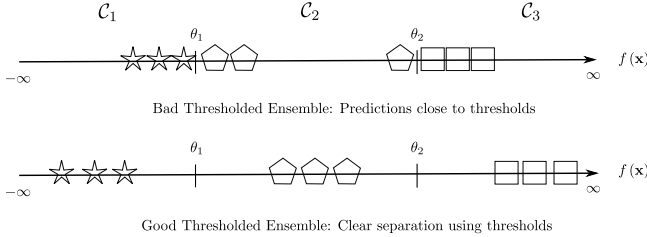


Fig. 3: Example of good and bad thresholded ensemble

However, in ensemble modelling, ensemble diversity is another key aspect. Fig. 4 shows three threshold ensemble models. All three models are perfectly accurate. However, as noted at the beginning of the section, these models assume that there is a latent variable which is defined on the real line where the patterns are projected. Taking this into account: which one of the three projections best represents the real projection of the patterns?. It is impossible to answer this question because we do not know the values for each of the patterns in the latent variable. If these values had been known, we would have considered the problem to be a standard regression problem.

In our opinion, it would be very interesting to combine individuals in the ensemble so that they meet two objectives: firstly, their projections must be as far as possible from the thresholds and secondly, their projections must be as different as possible from the projections of the remaining individuals in the ensemble. With this second objective, we ensure that the individuals in the ensemble are accurate even though their projections are different from the projections of other individuals. With the average of the projections, our aim is to better estimate the real values of the latent variable. This paper optimizes both objectives using the Negative Correlation Learning (NCL) framework [44], [58].

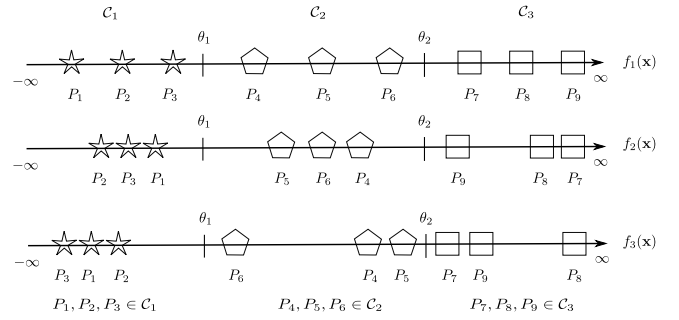


Fig. 4: Example of three different ranking functions: All of these ranking functions are diverse and accurate

NCL uses the following regularisation term to determine the amount of correlation in the ensemble [59], [60]:

$$\begin{aligned} R = p_i &= \frac{1}{N} \sum_{n=1}^N (f_i(\mathbf{x}_n) - \bar{f}(\mathbf{x}_n)) \left(\sum_{j \neq i} f_j(\mathbf{x}_n) - \bar{f}(\mathbf{x}_n) \right) \\ &= -\frac{1}{N} \sum_{n=1}^N (f_i(\mathbf{x}_n) - \bar{f}(\mathbf{x}_n))^2; i = 1, \dots, M \end{aligned}$$

Therefore, using this approach, the error of neural network i becomes:

$$e_i = \frac{1}{N} \sum_{n=1}^N (f_i(\mathbf{x}_n) - y_n)^2 + \lambda R; i = 1, \dots, M \quad (6)$$

where λ is a weighting parameter on the regularisation term R , y_n represent the rank of the n -th pattern and N the number of patterns. The λ parameter controls a trade-off between the two terms; with $\lambda = 0$ we would have an ensemble with each network trained with backpropagation, and as λ increases more and more emphasis would be placed on minimising the regularisation term.

In this work, the model parameters are optimized using the *iRprop+* local improvement procedure. In this case, the gradient vector is given by the following equation:

$$\nabla e_i = \left(\frac{\partial e}{\partial \beta_1}, \dots, \frac{\partial e}{\partial \beta_S}, \frac{\partial e}{\partial \mathbf{w}_1}, \dots, \frac{\partial e}{\partial \mathbf{w}_S} \right). \quad (7)$$

where β are the connections between hidden and output layer and \mathbf{w} the connections between input and hidden layer of the MLP neural network model $f_i(\mathbf{x}_n)$.

Let η be any of the parameters of β or \mathbf{w} . Therefore:

$$\begin{aligned} \frac{\partial e_i}{\partial \eta} &= \left(\frac{2}{N} \sum_{n=1}^N (f_i(\mathbf{x}_n) - y_n) \frac{\partial f_i(\mathbf{x}_n)}{\partial \eta} \right) - \\ &\quad \left(\frac{2\lambda}{N} \left(1 - \frac{1}{M} \right) \sum_{n=1}^N (f_i(\mathbf{x}_n) - \bar{f}(\mathbf{x}_n)) \frac{\partial f_i(\mathbf{x}_n)}{\partial \eta} \right) \end{aligned}$$

The following expressions include the derivatives of the parameters for the f_i MLP model ($f_i = \sum_{s=1}^S \beta_s B_s(\mathbf{x}, \mathbf{w}_s)$):

$$\frac{\partial f_i}{\partial \beta_s} = B_s(\mathbf{x}, \mathbf{w}_s); 1 \leq s \leq S$$

The gradient for the hidden layer depends of the kind of basis function. In this paper, $B_s(\mathbf{x}, \mathbf{w}_s)$ are sigmoidal nodes:

$$B_s(\mathbf{x}, \mathbf{w}_s) = \sigma \left(\sum_{i=1}^k w_{is} x_i \right), \sigma(x) = \frac{1}{1 + e^{-x}}$$

$$\frac{\partial f_i}{\partial \mathbf{w}_s} = \left(\frac{\partial f_i}{\partial w_{s1}}, \dots, \frac{\partial f_i}{\partial w_{sk}} \right)$$

$$\frac{\partial f_i}{\partial w_{st}} = \beta_s \sigma' \left(\sum_{i=1}^k w_{is} x_i \right) x_t$$

where $s = 1, 2, \dots, S$ and $t = 1, 2, \dots, k$.

V. COMPUTATIONAL EXPERIMENTS AND RESULTS

A. Data and variables involved in country risk detection

The analysis covers twenty-seven EU sovereign borrowers during the period from 2007 to 2010, encompassing the worldwide economic downturn and the beginnings of the on-going EU debt crisis: Austria, Belgium, Bulgaria, Czech Republic, Cyprus, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and the United Kingdom.

The output variable for the NCL is the long-term foreign currency rating assigned to each sovereign state by S&P included in the analysis as of December 31st of each year. The S&P's scale of measurement is a 22-point risk-rating scale, which is replaced by a numerical equivalent grade into broad categories maintaining the ordinal ranking of creditworthiness [21] (see Table I). In the case under study, the S&P ratings of EU countries are located on the top five broadest categories (see Table II). As shown in Table III, the number of sovereign states having their credit rating downgraded has outnumbered the number of those that have had their rating upgraded, leading some to conclude that the credit quality of the EU sovereign debt has declined over time. However, empirical studies have emphasized that CRAs have become more conservative over time, tightening the requirements that issuers must fulfil in order to achieve higher credit ratings [61], [62]. Regarding sovereign issuers, Ferri et al. [63] found that given their economic fundamentals, some countries were downgraded excessively during the East Asian financial crisis. In the same way, Gartner et al. [64] pointed out that the rating for the so-called PIGS countries (Portugal, Ireland, Greece and Spain) during the ongoing EU debt crisis is 2.30 notches lower than that of a hypothetical country, which has the same economic fundamentals but does not belong to this group. It is important to note that the dependent variable is the credit rating assigned to the country by analysts within S&P. Therefore, the model in this paper attempts to replicate the S&P sovereign ratings, rather than reproducing the decision process undertaken by the analyst or providing a more accurate estimate of sovereign rating.

Rating agencies, due to their business practice, do not officially disclose the precise models used for their rating methodologies. A common practice among rating agencies is to assign qualitative scores to several criteria and then arrive at a weighted average score. Beers and Cavanaugh

S&P		Rating	
Highest quality	AAA	C_1	
High quality	AA+	C_2	
	AA	C_2	
Strong payment capacity	AA-	C_2	
	A+	C_3	
Adequate payment capacity	A	C_3	
	A-	C_3	
Likely to fulfil obligations	BBB+	C_4	
	BBB	C_4	
High Credit Risk	BBB-	C_4	
	BB+	C_5	
Very High Credit Risk	BB	C_5	
	BB-	C_5	
Near default with possibility of recovery	B+	C_6	
	B	C_6	
Default	B-	C_6	
	CCC+	C_7	
Default	CCC	C_7	
	CCC-	C_7	
Default	CC	C_8	
	SD	C_9	
Default	D	C_9	

TABLE I: S&P credit rating measures

S&P		Ratings			
Rating	Class	2007	2008	2009	2010
AAA	C_1	11	11	9	9
AA	C_2	3	3	4	3
A	C_3	9	8	8	9
BBB	C_4	4	5	4	3
BB	C_5	0	0	2	3

TABLE II: S&P ratings, 2007-2010

Rating movements			
	2007-2008	2008-2009	2009-2010
Upgrades	2	0	2
Downgrades	5	8	4
Rating movements in rating notches			
Upgrades	1.0	0.0	1.0
Downgrades	1.2	1.4	2.0

TABLE III: S&P rating movements, 2007-2010

[65] provide an excellent explanation of the criteria used by Standard and Poor's. They list 44 variables grouped into 10 categories—political risk, income and economic structure, economic growth prospects, fiscal flexibility, general government debt burden, off-budget and contingent liabilities, monetary flexibility, external liquidity, and public sector external debt burden. Nevertheless, many researchers have found that the ratings by major agencies are largely explained by a handful of macroeconomic variables ([66]). To our knowledge, Feder and Uy [67] were the first to identify the determinants of country risk ratings. Many other studies investigate the determinants of sovereign ratings and show that sovereign ratings are mainly driven by economic fundamentals (e.g. [13], [20]). Therefore, in line with the theoretical and empirical literature, ten economic indicators as well as one political indicator have been selected as explanatory variables:

- 1) *Real GDP growth*: Indicator of the country's govern-

ment's ability to repay outstanding obligations. Unit of measurement: Rate (Eurostat).

- 2) *GDP per capita*: Targets total income of country's citizen—reflects cost of living. Unit of measurement: Euros per inhabitant (Eurostat).
- 3) *Government debt*: Indicates the total debt of government held by the public. Unit of measurement: Percent of GDP (Eurostat).
- 4) *Fiscal balance*: Indicates the demand (deficit)/offer (surplus) of external budgetary financing. Unit of measurement: Percent of GDP (Eurostat).
- 5) *External debt*: Indicates the outstanding amount of those current, and not contingent liabilities owed to non-residents by residents of an economy that requires payment(s) either of principal and/or interest by the debtor at some point(s) in the future. Unit of measurement: Percent of exports (World Bank and EU Member Central Banks).
- 6) *Level of external reserves*: Held by a country's central bank as defense against withdrawal of foreign credit. Unit of measurement: Percent of imports (Eurostat).
- 7) *Current account balance*: Indicate how much net import of capital a country requires. Unit of measurement: Percent of GDP (Eurostat).
- 8) *Inflation*: Shows change in the level of price index for a basket of commonly used goods. Unit of measurement: Index (2005=100) (Eurostat).
- 9) *Unemployment rate*: Indicator of size of output gap and of underutilisation of resources. Unit of measurement: Rate (Eurostat).
- 10) *Unit labor cost*: Indicator of a country's competitiveness in international trade. Unit of measurement: Index (2005=100) (Eurostat).
- 11) *Government effectiveness*: Indicator of the quality of policy formulation and implementation, and the credibility of the government's commitment to such policies. Unit of measurement: Percentile (World Bank)

The relationship between these variables and credit ratings is such that it is expected that higher levels of external and government debt, higher rates of inflation and unemployment, and higher labour costs and current account deficits will result in lower ratings; on the other hand, higher credit ratings are associated with higher levels of fiscal balance (surplus), higher levels of income and external reserves, higher rates of GDP growth and finally higher scores in government effectiveness. Furthermore, the variables of inflation, unemployment, GDP growth, fiscal balance and current accounts are entered as a 3 year average, reflecting the agencies' approach to eliminating the effect of the business cycle when deciding on a sovereign rating. Recently, during 2009 and 2010, CRAs have started to define new methodologies that shift the criteria from a "through-a-cycle" to a "through-a-crisis" approach. In particular, S&P have established a new credit stability criterion that uses hypothetical stress scenarios as benchmarks for calibrating the criteria over time. However, neither those scenarios nor other further developments are considered in our experimental design, as they will not be effective until the

second semester of 2011.

In order to assess the ability of the models in an out-of-time dataset, the dataset has been split into two subsequent time periods, holding the later for evaluation of the model only. Consequently, the training information table consisted of 81 samples from 27 countries in the period 2007-2009 (annual data) described by the variables explained. The test or generalization information table consisted of 27 data described by the same variables in the year 2010.

B. Machine learning methods used for comparison purposes and experimental design

For comparison purposes, different state-of-the-art methods have been included in the experimentation. These methods are the following:

- *Nominal Classifiers*:

- The **Multi Logistic Regression (MLR)** algorithm. It is based on applying the LogitBoost algorithm with simple regression functions and determining the optimum number of iterations by a five fold cross-validation [68].
- A **Gaussian Radial Basis Function Network (RBFN)** [69], deriving the centres and width of hidden units using k -means, and combining the outputs obtained from the hidden layer using logistic regression.
- A **Multilayer Perceptron (MLP)** [69] with sigmoid units as hidden nodes, obtained by means of the back-propagation algorithm.
- **Support Vector Machine (SVM)** [70] nominal classifier is included in the experiments in order to validate our proposal contributions. Cost Support Vector Classification (SVC) available in libSVM 3.0 [71] is used as the SVM classifier implementation.

- *Ordinal Regression approaches*:

- The **Proportional Odd Model (POM)** [31] is an extension of the binary logistic regression model for ordinal multi-class categorization problems. This is one of the first models specifically designed for ordinal regression, and it arose from a statistical background. The model is based on the assumption of stochastic ordering of the space X , i.e. for all \mathbf{x}_1 and \mathbf{x}_2 , too:

$$P(y \preceq C_j | \mathbf{x}_1) \succeq P(y \preceq C_j | \mathbf{x}_2) \quad \forall C_j \in Y,$$

or:

$$P(y \preceq C_j | \mathbf{x}_1) \preceq P(y \preceq C_j | \mathbf{x}_2) \quad \forall C_j \in Y. \quad (8)$$

In the POM model, the cumulative probability of the rating y is given by:

$$P(y \preceq C_j) = \frac{1}{1 + \exp(\theta_j + \beta_1 x_1 + \dots + \beta_k x_k)} \quad (9)$$

where $\mathbf{x} = (x_1, x_2, \dots, x_k)$ represents the array with k inputs variables, $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_k)$ the corresponding coefficient vector and $j = 1, \dots, J$.

Because $P(y \leq C_J) = 1$, the parameter θ_J is equal to ∞ . The latent variable is the linear combination of the input variables

$$z = -\beta_1 x_1 - \beta_2 x_2 - \dots - \beta_k x_k \quad (10)$$

and summarizes the financial information of the risk entity.

- **SVMRank** [72] applies the Extended Binary Classification (EBC) method to SVM. The EBC method can be summarized in the following three steps. First, all training samples are transformed into extended samples, weighting these samples by using the absolute cost matrix. Second, all the extended examples are jointly learned by a binary classifier with confidence outputs, aiming at a low weighted 0/1 loss. The last step is used to convert the binary outputs to a rank.
- **Support Vector Ordinal Regression (SVOR)** by Chu and Keerthi [73], [74], proposes two new support vector approaches for ordinal regression. Here, multiple thresholds are optimized in order to define parallel discriminant hyperplanes for the ordinal scales. The first approach with explicit inequality constraints on the thresholds derives the optimality conditions for the dual problem, and adapts the SMO algorithm for the solution: we will refer to it as SVOR(EX). In the second approach, the samples in all the categories are allowed to contribute errors to each threshold, which is why there is no need to include inequality constraints in the problem. This approach is called a SVOR with implicit constraints (SVOR(IM)).
- Ensemble approaches for Ordinal regression:
 - * **A Simple Approach to Ordinal Regression (ASAOR)** [41] is a meta classifier that allows standard classification algorithms to be applied to ordinal class problems. This methodology was already described in the Introduction Section. In the current work, the C4.5 method available in Weka [75] is used as the underlying classification algorithm, since this is the one initially employed by the authors of ASAOR.
 - * **Multi-Class Ordinal Support vector machines (MCOSvm)** [42] is an enhanced ensemble method for ordinal regression. It is closely related to the ASAOR method. In this case, weighted SVMs are used as base classifiers. Specific weights are assigned to each pattern in such a way that errors of more than one rank are more penalized. Therefore the weight of a training pattern differs for each binary SVM.
- *Regression approaches:*
 - **Regression neural network model (rNN)**: as stated in the introduction, regression models can be applied to solve the classification of ordinal data. A common technique for ordered classes is to estimate by regression any ordered scores $s_1 < \dots < s_J$ by

replacing the target class C_i with the score s_i . The simplest case would be setting $s_i = i$; $i = 1, \dots, J$. A neural network with a single output was trained to estimate the scores. Furthermore, this model is particularly interesting because our ensemble model is composed of M rNN-type individuals.

Regarding the hyper-parameters of different algorithms, the following procedure has been applied. For the Support Vector algorithms, i.e. SVC, SVMRank, SVOR(EX), SVOR(IM), and MCOSvm the corresponding hyper-parameters (regularization parameter, C , and width of the Gaussian functions, γ), were adjusted using a grid search with a 5-fold cross-validation, considering the following ranges: $C \in \{10^3, 10^1, \dots, 10^{-3}\}$ and $\gamma \in \{10^3, 10^0, \dots, 10^{-3}\}$.

For the Neural Network algorithms, i.e. MLP, and rNN, the corresponding hyper-parameters (number of hidden neuron, H , and number of iterations of the local search procedure, iterations), were adjusted using a grid search with a 5-fold cross-validation, considering the following ranges: $H \in \{5, 10, 15, 20, 30, 40\}$ and iterations $\in \{25, 50, \dots, 500\}$. In the case of the RBFN methodology, the number of hidden neuron, H , was adjusted using a grid search with a 5-fold cross-validation, considering the following ranges: $H \in \{5, 10, 15, 20, 30, 40\}$. Finally, the NCL selects the λ parameter by cross validation within the range $\{0.0, 0.1, \dots, 1.0\}$. We notice that λ could be a little greater than 1 [$\lambda \leq \frac{M}{M-1}$] to guarantee the positive definite of the Hessian matrix [59]. Since we use $M = 25$ in this paper, the up-bound of λ (1.0417) is close to 1 and we will not use the λ values which are greater than 1.

Regarding the experimental design, the rNN and NCL approaches are non-deterministic methodologies because the neural network weight vectors are initialized randomly. For these methodologies, we run the procedure 30 times for the holdout considered.

C. Ordinal classification evaluation metrics

Four evaluation metrics have been considered which quantify the accuracy of N predicted ordinal labels for a given dataset $\{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_N\}$, with respect to the true targets $\{y_1, y_2, \dots, y_N\}$:

- 1) Correct Classification Rate (C) is simply the fraction of correct predictions on individual samples:

$$C = \frac{1}{N} \sum_{i=1}^N I(\hat{y}_i = y_i), \quad (11)$$

where $I(\cdot)$ is the zero-one loss function and N is the number of patterns of the dataset.

- 2) Mean Absolute Error (MAE) is the average deviation of the prediction from the true targets, i.e.:

$$MAE = \frac{1}{N} \sum_{i=1}^N |\mathcal{O}(\hat{y}_i) - \mathcal{O}(y_i)|, \quad (12)$$

where $\mathcal{O}(C_k) = k$, $1 \leq k \leq K$, i.e. $\mathcal{O}(y_i)$ is the order of class label y_i .

- 3) τ_b : The Kendall's τ is a statistic used to measure the association between two measured quantities. Specifically, it is a measure of rank correlation [76]:

$$\tau_b = \frac{\sum \hat{c}_{ij} c_{ij}}{\sqrt{\sum \hat{c}_{ij}^2 \sum c_{ij}^2}} \quad (13)$$

where \hat{c}_{ij} is +1 if $\mathcal{O}(\hat{y}_i)$ is greater than $\mathcal{O}(\hat{y}_j)$, 0 if $\mathcal{O}(\hat{y}_i)$ and $\mathcal{O}(\hat{y}_j)$ are the same, and -1 if $\mathcal{O}(\hat{y}_i)$ is less than $\mathcal{O}(\hat{y}_j)$, and similar for $\mathcal{O}(y)$. τ_b values are between -1 and 1.

- 4) MMAE: The Maximum MAE value of all the classes. MMAE is the MAE value of the class with higher distance from the true values to the predicted ones [77]:

$$\text{MMAE} = \max\{\text{MAE}_j; j = 1, \dots, J\} \quad (14)$$

where MAE_j is the MAE value for the j -th class. MMAE values are between 0 and $J - 1$.

These measures aim to evaluate different aspects that can be taken into account when an ordinal regression problem is considered: accuracy measures, that patterns are generally well classified, that the MAE measures that the classifier tends to predict a class as close to the real class as possible, that the τ_b measures the correlation between predicted and real target pairs and finally, that the MMAE measures the order in the class which is worst ordered. The τ_b measure is independent of the values chosen for the ranks representing the classes.

D. Results

The NCL method has been compared to the well-known nominal classification, ordinal regression and regression techniques given in Section V-B, using the ordinal regression metrics defined in Section V-C. Tables IV and V show the results obtained with the different techniques tested using the training and generalization set. It is important to note that rNN and NCL are non-deterministic methods because they are based on randomly generated numbers. For this reason, these methods were run 30 times and the best individual in the training set was extracted (rNN_{Best} and NCL_{Best}). Tables IV and V include the average, the standard deviation, and the best values of the C, MAE, τ_b and MMAE on the training (C_T , MAE_T , τ_{b_T} and MMAE_T) and the generalization (C_G , MAE_G , τ_{b_G} and MMAE_G) sets of these 30 models, together with the results of the remaining methods.

A descriptive analysis of the results leads to the following remarks: the NCL method obtained the best result in terms of C, MAE, τ_b and MMAE both in training and generalization sets comparing all techniques. As can be observed, the POM model is not able to reflect nonlinear relationships among input variables, necessary for performing a realistic classification task. The results obtained for the MMAE metric show that the classifiers tested tend to predict a class quite close to the real class for all the classes.

Figure 5 shows the box plot obtained with the results of the different algorithms in C_G , MAE_G , τ_{b_G} and MMAE_G . Boxplots depict algorithm results according to the smallest observation, lower quartile, median, upper quartile and largest

	Training Results			
	C_T	MAE_T	τ_{b_T}	MMAE_T
MLR	87.654	0.123	0.922	0.600
RBFN	96.296	0.037	0.977	0.100
MLP	98.765	0.012	0.992	0.083
SVC	92.592	0.074	0.953	0.300
POM	82.719	0.185	0.898	0.666
SVMRank	96.296	0.037	0.977	0.100
SVOR(EX)	95.061	0.049	0.969	0.100
SVOR(IM)	100.000	0.000	1.000	0.000
ASAOR	97.530	0.022	0.987	0.080
MCOSvm	100.000	0.000	1.000	0.000
rNN	97.121 _{0.341}	0.021 _{0.012}	0.974 _{0.016}	0.092 _{0.009}
NCL	100.000 _{0.000}	0.000 _{0.000}	1.000 _{0.000}	0.000 _{0.000}
rNN _{Best}	100.000	0.000	1.000	0.000
NCL _{Best}	100.000	0.000	1.000	0.000

The best result is in bold face and the second best result in italics

TABLE IV: Training results of the C_T , MAE_T , τ_{b_T} and MMAE_T of the method proposed compared to those obtained using different statistical and artificial intelligence methods.

	Generalization Results			
	C_G	MAE_G	τ_{b_G}	MMAE_G
MLR	70.370	0.333	0.851	0.666
RBFN	74.074	0.296	0.836	0.444
MLP	70.370	0.296	0.860	0.333
SVC	70.370	0.370	0.821	1.333
POM	62.965	0.407	0.816	0.666
SVMRank	70.370	0.296	0.868	0.666
SVOR(EX)	70.370	0.296	0.866	0.333
SVOR(IM)	74.074	0.259	0.879	0.333
ASAOR	66.667	0.370	0.821	0.666
MCOSvm	77.777	0.259	0.890	0.444
rNN	73.898 _{1.316}	0.256 _{0.033}	0.863 _{0.045}	0.402 _{0.062}
NCL	83.465 _{0.919}	0.159 _{0.0190}	0.912 _{0.019}	0.431 _{0.093}
rNN _{Best}	77.777	0.222	0.873	0.333
NCL _{Best}	85.185	0.148	0.926	0.333

The best result is in bold face and the second best result in italics

TABLE V: Generalization results of the C_G , MAE_G , τ_{b_G} and MMAE_G of the method proposed compared to those obtained using different statistical and artificial intelligence methods.

observation. As seen in Table V and Figure 5, the NCL method obtains the best result in terms of C_G , MAE_G , τ_{b_G} out of all the techniques compared. The differences in C_G , MAE_G , τ_{b_G} are really important with respect to techniques such as MLP, SVC, POM, SVOR(EX) or ASAOR. In general, these results show that the proposed approach based on rNN ensembles is robust enough to classify the sovereign ratings reported by the S&P international credit rating agency, obtaining better results than the rNN base method and state-of-the-art classification algorithms.

E. Discussion

In order to perform this section, we have evaluated the projections of the best proposed model (NCL) both in training and test sets. As shown in Fig. 6, most of the patterns (85%) have been correctly classified by the NCL model proposed. On analysing the four errors committed in the classification test, they are seen to correspond to these four countries: Germany, the United Kingdom, France and Greece. The first three countries have been classified into a lower category than S&P's rating, while Greece has been classified into a higher category than the benchmark.

These differences may be due to the fact that quantitative data employed in our model only provide information on the

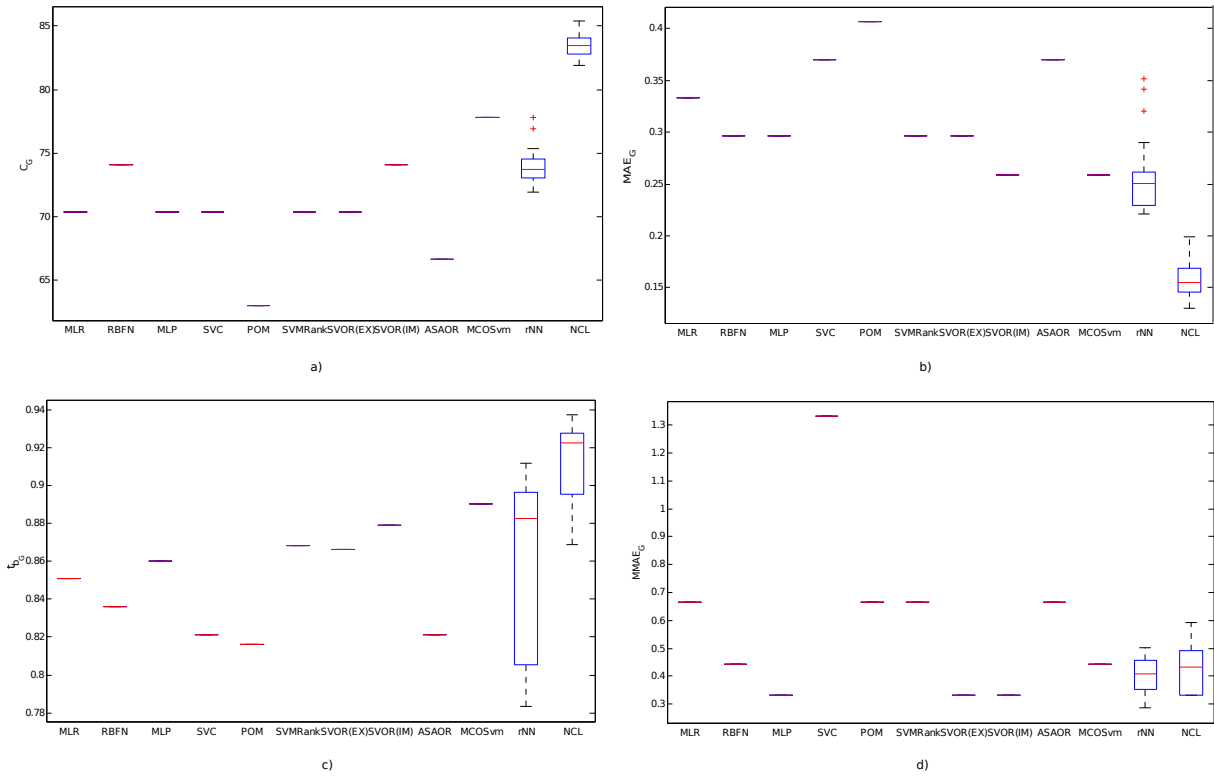


Fig. 5: Box plots: Results of the MLR, RBFN, MLP, SVC, POM, SVMRank, SVOR(EX), SVOR(IM), ASAOR, MCOSvm, rNN and the NCL method proposed

historical performance of the economy and on its fundamental structural features. They are basically backward-looking, while sovereign rating analysis requires forward-looking evaluations of the risk default over a medium to long-term time horizon. Therefore, examination of past experience has to be supplemented by medium-term projections and by the construction of a range of scenarios that test the vulnerability of a country's economic, political, and financial situation to a variety of shocks generated both internally and externally. In addition, qualitative and judgmental aspects of analysis are unavoidable even in the interpretation of quantitative indicators [78]. Furthermore, in the case of Greece, the reliability of the data employed [79] may be the cause of its better performance.

Finally, in order to make sure that the crisis has not a huge impact in the performance of the model, one additional experiment has been carried out. In this experiment, the performance of the best individual in the training set of the NCL method (NCL_{Best}) has been tested with two datasets. In these datasets, the training set has been built with the latest years and the generalization set with the earliest years. Two possible configuration has been included: (i) training the model with the 2008, 2009 and 2010 years and validate it with the 2007 year and (ii) training the model with the 2009 and 2010 years and validate it with the 2007 and 2008 years.

Table VI includes the results of the experiment. As can be observed, the performance of the model in the first case has been slightly reduced (considering the generalization set). In the second case, the degradation of the model performance is higher. The reason for such underperformance is not so

	Training Results (NCL_{Best})			
	C_T	MAE_T	τ_{b_T}	$MMAE_T$
{2008, 2009, 2010}	100.000	0.000	1.000	0.000
{2009, 2010}	100.000	0.000	1.000	0.000
	Generalization Results (NCL_{Best})			
	C_G	MAE_G	τ_{b_G}	$MMAE_G$
{2007}	81.481	0.222	0.918	0.500
{2007, 2008}	79.630	0.240	0.893	0.833

TABLE VI: Training and generalization results of the C , MAE , τ_b and $MMAE$ of the NCL_{Best} method considering training sets with the latest years and generalization sets with the earliest years.

much due to the crisis in itself, but to the fact that the second model has been trained using a set which does not include the data of 2008 (included in the generalization set). As shown in Table III, the year 2008 registered a higher number of rating movements than other years and hence it provided valuable information for training process.

VI. CONCLUSIONS

In this paper we have presented a new ensemble model for ordinal regression problems based on Negative Correlation Learning. The models are trained taking into account the correlation in the projections of different individuals in the ensemble. As a result, many accurate and diverse projections are obtained. This paper has tested this approach in a sovereign credit-rating dataset, formed by macroeconomic fundamentals of the twenty-seven EU countries in the periods 2007-2009

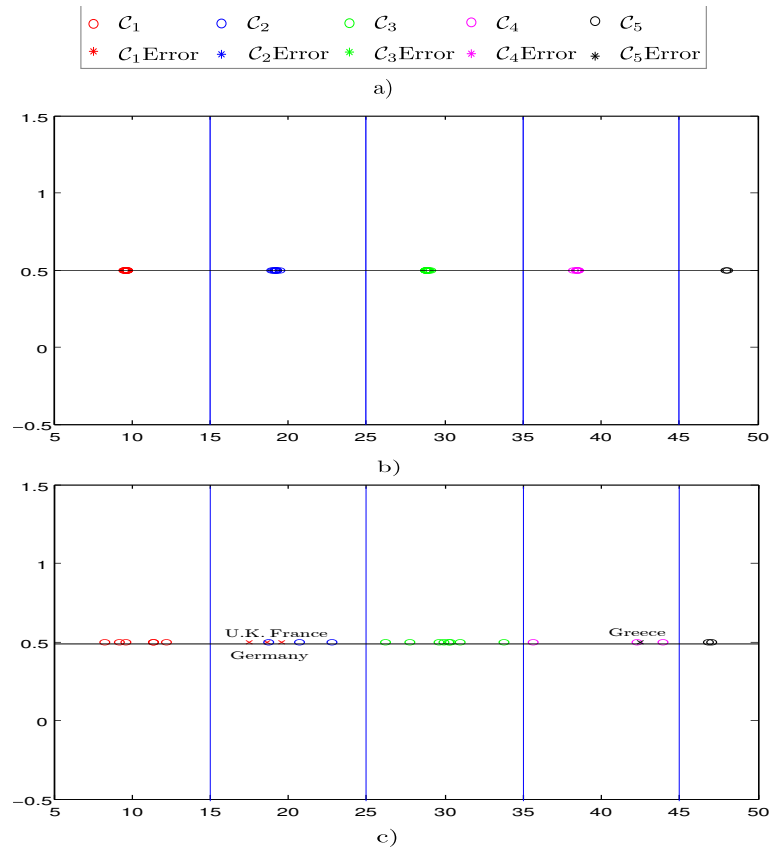


Fig. 6: NCL Projections on country risk ordinal classification datasets. Classes are shown as crosses and dots depending if the pattern has been correctly classified or not: (a) Class Naming (b) Projection patterns on the real line generated by NCL on the training set (c) Projection patterns on the real line generated by NCL on the test set

(training data) and 2010 (test data), and the corresponding rating variable provided by S&P. Furthermore, an additional experiment has been carried out in order to assess the possible impact of crisis on the best model performance.

The results obtained show that the approaches proposed, which are based on ensembles of MLPs trained with NCL, are robust enough to tackle the ordinal classification of sovereign credit-rating provided by the S&P rating agency, and obtain better results than the majority of existing alternative methods. Therefore, it appears that NCL could be a useful tool for informing and supporting the analyst in this decision process, entailing less time and cost. Despite this, such models cannot replace the analyst's role and the partly subjective process of assigning a credit rating.

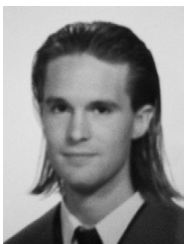
In future work, this study can be extended by comparing other new methods of ensemble learning for ordinal regression, and by adding the ratings assigned by the other two well-established credit rating agencies, namely Moody's and Fitch, in order to get a broader view of sovereign credit-rating assessment, since this information is especially valuable in a context of uncertainty in global financial and economic crises and keeping in mind the increasing relevance these agencies are acquiring in influencing the financial market.

REFERENCES

- [1] R. Cantor and F. Packer, "Differences of opinion and selection bias in the credit rating industry," *Journal of Banking and Finance*, vol. 21, no. 10, pp. 1395 – 1417, 1997.
- [2] Y. Jafry and T. Schuermann, "Measurement, estimation and comparison of credit migration matrices," *Journal of Banking and Finance*, vol. 28, no. 11, pp. 2603 – 2639, 2004.
- [3] P. Behr and A. Güttler, "The informational content of unsolicited ratings," *Journal of Banking and Finance*, vol. 32, no. 4, pp. 587 – 599, 2008.
- [4] R. Arezki, B. Candelon, and A. Sy, "Sovereign rating news and financial markets spillovers: Evidence from the European debt crisis," Tech. Rep., 2011.
- [5] E. I. Altman and H. A. Rijken, "How rating agencies achieve rating stability," *Journal of Banking and Finance*, vol. 28, no. 11, pp. 2679 – 2714, 2004.
- [6] J. D. Amato and C. H. Furfine, "Are credit ratings procyclical?" *Journal of Banking and Finance*, vol. 28, no. 11, pp. 2641 – 2677, 2004.
- [7] A. Duff and S. Einig, "Understanding credit ratings quality: Evidence from UK debt market participants," *The British Accounting Review*, vol. 41, pp. 107–119, 2009.
- [8] R. Al-Sakka and O. ap Gwilym, "Split sovereign ratings and rating migrations in emerging economies," *Emerging Markets Review*, vol. 11, no. 2, pp. 79 – 97, 2010.
- [9] M. A. Ferreira and P. M. Gama, "Does sovereign debt ratings news spill over to international stock markets?" *Journal of Banking and Finance*, vol. 31, no. 10, pp. 3162 – 3182, 2007.
- [10] S. Kim and E. Wu, "Sovereign credit ratings, capital flows and financial sector development in emerging markets," *Emerging Markets Review*, vol. 9, no. 1, pp. 17–39, 2008.
- [11] R. Brooks, R. W. Faff, D. Hillier, and J. Hillier, "The national market impact of sovereign rating changes," *Journal of Banking and Finance*, vol. 28, no. 1, pp. 233 – 250, 2004.
- [12] V. Hooper, T. Hume, and S.-J. Kim, "Sovereign rating changes—do they provide new information for stock markets?" *Economic Systems*, vol. 32, no. 2, pp. 142 – 166, 2008.

- [13] R. M. Cantor and F. Packer, "Determinants and impact of sovereign credit ratings," *Economic Policy Review*, no. Oct, pp. 37–53, 1996.
- [14] E. Durbin and D. Ng, "The sovereign ceiling and emerging market corporate bond spreads," *Journal of International Money and Finance*, vol. 24, no. 4, pp. 631–649, 2005.
- [15] J. R. M. Hand, R. W. Holthausen, and R. W. Leftwich, "The effect of bond rating agency announcements on bond and stock prices," *Journal of Finance*, vol. 47, no. 2, pp. 733–752, 1992.
- [16] A. Afonso and R. M. Sousa, "The macroeconomic effects of fiscal policy," *Applied Economics*, vol. 44, no. 34, pp. 4439–4454, 2012.
- [17] I. D. Dichev and J. Piotroski, "The long-run stock returns following bond ratings changes," *Journal of Finance*, vol. 56, no. 1, pp. 173–203, 2001.
- [18] A. Gande and D. C. Parsley, "News spillovers in the sovereign debt market," *Journal of Financial Economics*, vol. 75, no. 3, pp. 691–734, 2005.
- [19] R. Alsakka and O. ap Gwilym, "Rating agencies' signals during the european sovereign debt crisis: Market impact and spillovers," *Journal of Economic Behavior and Organization*, vol. Corrected Proof, no. 0, 2011.
- [20] A. Afonso, "Understanding the determinants of sovereign debt ratings: Evidence for the two leading agencies," *Journal of Economics and Finance*, vol. 27, pp. 56–74, 2003.
- [21] A. W. Butler and L. Fauver, "Institutional environment and sovereign credit ratings," *Financial Management*, vol. 35, no. 3, pp. 53–79, 2006.
- [22] N. Sargen, "Economic indicators and country risk appraisal," *Economic Review*, vol. 1, pp. 19–35, 1977.
- [23] Y.-T. Hu, R. Kiesel, and W. Perraudin, "The estimation of transition matrices for sovereign credit ratings," *Journal of Banking and Finance*, vol. 26, no. 7, pp. 1383–1406, 2002.
- [24] E. Bissoondoyal-Bheenick, "Rating timing differences between the two leading agencies: Standard and Poor's and Moody's," *Emerging Markets Review*, vol. 5, no. 3, pp. 361–378, 2004.
- [25] A. Afonso, P. Gomes, and P. Rother, "Ordered response models for sovereign debt ratings," *Applied Economics Letters*, vol. 16, no. 8, pp. 769–773, 2009.
- [26] J. A. Bennell, D. Crabbe, S. Thomas, and O. ap Gwilym, "Modelling sovereign credit ratings: Neural networks versus ordered probit," *Expert Systems with Applications*, vol. 30, no. 3, pp. 415–425, 2006.
- [27] J. Yim and H. Mitchell, "Comparison of country risk models: hybrid neural networks, logit models, discriminant analysis and cluster techniques," *Expert Systems with Applications*, vol. 28, no. 1, pp. 137–148, 2005.
- [28] K.-J. Kim and H. Ahn, "A corporate credit rating model using multi-class support vector machines with an ordinal pairwise partitioning approach," *Computers and Operations Research*, vol. 39, no. 8, pp. 1800–1811, 2012.
- [29] C. Hung and J.-H. Chen, "A selective ensemble based on expected probabilities for bankruptcy prediction," *Expert Systems with Applications*, vol. 36, no. 3, Part 1, pp. 5297–5303, 2009.
- [30] L. Yu, S. Wang, and K. K. Lai, "Credit risk assessment with a multistage neural network ensemble learning approach," *Expert Systems with Applications*, vol. 34, no. 2, pp. 1434–1444, 2008.
- [31] P. McCullagh, "Regression models for ordinal data," *Journal of the Royal Statistical Society, Series B*, vol. 4, pp. 109–142, 1980.
- [32] A. Agresti, *Analysis of ordinal categorical data*. Wiley Series in Probability and Statistics, 1984.
- [33] K. Crammer and Y. Singer, "Online ranking by projecting," *Neural Computation*, vol. 17, pp. 145–175, 2005.
- [34] A. Shashua and A. Levin, "Ranking with large margin principle: Two approaches," in *Advances in Neural Information Processing Systems 15 [Neural Information Processing Systems, NIPS 2002, December 9-14, 2002, Vancouver, British Columbia, Canada]*, S. Becker, S. Thrun, and K. Obermayer, Eds. MIT Press, 2002, pp. 937–944.
- [35] W. Chu and Z. Ghahramani, "Gaussian processes for ordinal regression," *Journal of Machine Learning Research*, vol. 6, pp. 1019–1041, 2005.
- [36] W. Chu and S. S. Keerthi, "Support vector ordinal regression," *Neural Computation*, vol. 19, pp. 792–815, 2007.
- [37] J. S. Cardoso and J. F. P. da Costa, "Learning to classify ordinal data: The data replication method," *Journal of Machine Learning Research*, vol. 8, pp. 1393–1429, 2007.
- [38] B.-Y. Sun, X.-M. Zhang, and W.-B. Li, "An improved ordinal regression approach with sum-of-margin principle," in *Sixth International Conference on Natural Computation, ICNC 2010, Yantai, Shandong, China, 10-12 August 2010*. IEEE, 2010, pp. 853–857.
- [39] E. K. Tang, P. N. Suganthan, and X. Yao, "An analysis of diversity measures," *Machine Learning*, vol. 65, no. 1, pp. 247–271, 2006.
- [40] M. Galar, A. Fernandez, E. Barrenechea, H. Bustince, and F. Herrera, "A review on ensembles for the class imbalance problem: Bagging, boosting, and hybrid-based approaches," *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, vol. 42, no. 4, pp. 463–484, 2011.
- [41] E. Frank and M. Hall, "A simple approach to ordinal classification," in *ECML'01*, 2001, pp. 145–156.
- [42] W. Waegeman and L. Boullart, "An ensemble of weighted support vector machines for ordinal regression," *International Journal of Computer Systems Science and Engineering*, vol. 3, no. 1, pp. 47–51, 2009.
- [43] Y. Liu and X. Yao, "Ensemble learning via negative correlation," *Neural Networks*, vol. 12, pp. 1399–1404, 1999.
- [44] Y. Liu, X. Yao, and T. Higuchi, "Ensembles with negative correlation learning," *IEEE Transactions on Evolutionary Computation*, vol. 4, no. 4, pp. 380–387, 2000.
- [45] G. Kaminsky and S. L. Schmukler, "Emerging market instability: Do sovereign ratings affect country risk and stock returns?" *World Bank Economic Review*, vol. 16, no. 2, pp. 171–195, 2002.
- [46] C. J. Frank and W. R. Cline, "Measurement of debt servicing capacity: An application of discriminant analysis," *Journal of International Economics*, vol. 1, no. 3, pp. 327–344, 1971.
- [47] W. T. Carleton and E. M. Lerner, "Statistical credit scoring of municipal bonds," *Journal of Money, Credit and Banking*, vol. 1, no. 4, pp. 750–64, November 1969.
- [48] F. Fernández-Navarro, C. Hervás-Martínez, P. A. Gutiérrez, and M. Carbo-reño, "Evolutionary q-gaussian radial basis functions neural networks for multi-classification," *Neural Networks*, vol. 24, no. 7, pp. 779–784, 2011.
- [49] A. Castaño, F. Fernández-Navarro, C. Hervás-Martínez, and P. A. Gutiérrez, "Neuro-logistic models based on Evolutionary Generalized Radial Basis Function for the microarray gene expression classification problem," *Neural Processing Letters*, vol. 34, no. 2, pp. 117–131, 2011.
- [50] F. Fernández-Navarro, C. Hervás-Martínez, J. Sánchez-Monedero, and P. A. Gutiérrez, "MELM-GRBF: A modified version of the extreme learning machine for generalized radial basis function neural networks," *Neurocomputing*, vol. 74, no. 16, pp. 2502–2510, 2011.
- [51] J. C. Singleton and A. J. Surkan, "Bond rating with neural networks," in *Proceedings of Neural Networks in the Capital Markets*. London, 1993.
- [52] Z. Huang, H. Chen, C.-J. Hsu, W.-H. Chen, and S. Wu, "Credit rating analysis with support vector machines and neural networks: a market comparative study," *Decision Support Systems*, vol. 37, pp. 543–558, 2004.
- [53] D. Trigueiros and R. Taffler, "Neural networks and empirical research in accounting," *Accounting and Business Research*, vol. 26, no. 4, pp. 347–355, 1996.
- [54] C. P. Lim and R. F. Harrison, "Online pattern classification with multiple neural network systems: An experimental study," *IEEE Transactions on Systems, Man and Cybernetics Part C: Applications and Reviews*, vol. 33, no. 2, pp. 235–247, 2003.
- [55] W. Lin, Y. H. Hu, and C.-F. Tsai, "Machine learning in financial crisis prediction: A survey," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol. 42, no. 4, pp. 421–436, 2012.
- [56] K. Tumer and J. Ghosh, "Analysis of decision boundaries in linearly combined neural classifiers," *Pattern Recognition*, vol. 29, pp. 341–348, 1996.
- [57] —, "Error correlation and error reduction in ensemble classifiers," *Connection Science*, vol. 8, no. 3, pp. 385–404, 1996.
- [58] H. Chen, P. Tino, and X. Yao, "Probabilistic classification vector machines," *IEEE Transactions on Neural Networks*, vol. 20, no. 6, pp. 901–914, 2009.
- [59] G. Brown, J. L. Wyatt, and P. Tiño, "Managing diversity in regression ensembles," *Journal of Machine Learning Research*, vol. 6, pp. 1621–1650, 2005.
- [60] S. Wang, H. Chen, and X. Yao, "Negative correlation learning for classification ensembles," in *The 2010 International Joint Conference on Neural Networks (IJCNN)*. IEEE, 2010, pp. 1–8.
- [61] M. E. Blume, F. Lim, and A. C. Mackinlay, "The declining credit quality of u.s. corporate debt: myth or reality?" *Journal of Financial Economics*, vol. 53, pp. 458–472, 1998.
- [62] E. Gonis and P. Taylor, "Changing credit rating standards in the UK: empirical evidence from 1999 to 2004," *Applied Financial Economics*, vol. 19, no. 3, pp. 213–225, 2009.
- [63] G. Ferri, L.-G. Liu, and J. E. Stiglitz, "The procyclical role of rating agencies: Evidence from the east asian crisis," *Economic Notes*, vol. 28, no. 3, pp. 335–355, 1999.

- [64] M. Gartner, B. Griesbach, and F. Jung, "Pigs or lambs? the European sovereign debt crisis and the role of rating agencies," *International Advances in Economic Research*, vol. 17, pp. 288–299, 2011.
- [65] D. Beers and M. Cavanaugh, "Sovereign credit ratings: A primer. new york," *Standard & Poor's*, 2005.
- [66] D. Ratha, P. K. De, and S. Mohapatra, "Shadow sovereign ratings for unrated developing countries," *World Development*, vol. 39, no. 3, pp. 295–307, 2011.
- [67] G. Feder and L. V. Uy, "The determinants of international creditworthiness and their policy implications," *Journal of Policy Modeling*, vol. 7, no. 1, pp. 133–156, 1985.
- [68] N. Landwehr, M. Hall, and E. Frank, "Logistic model trees," *Machine Learning*, vol. 59, no. 1-2, pp. 161–205, 2005.
- [69] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd ed., ser. Data Management Systems. Morgan Kaufmann, 2005.
- [70] V. N. Vapnik, *The Nature of Statistical Learning Theory*. Springer, 1999.
- [71] C.-C. Chang and C.-J. Lin, *LIBSVM: a library for support vector machines*, 2001, software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [72] L. Li and H.-T. Lin, "Ordinal Regression by Extended Binary Classification," in *Advances in Neural Information Processing Systems 19*, 2007, pp. 865–872.
- [73] W. Chu and S. S. Keerthi, "New approaches to support vector ordinal regression," in *In ICML '05: Proceedings of the 22nd international conference on Machine Learning*, 2005, pp. 145–152.
- [74] —, "Support vector ordinal regression," *Neural Computation*, vol. 19, pp. 792–815, 2007.
- [75] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The weka data mining software: an update," *SIGKDD Explorations Newsletter*, vol. 11, no. 1, pp. 10–18, 2009.
- [76] M. Kendall, *Rank correlation methods*. London: Griffin, 1948.
- [77] M. Cruz-Ramirez, C. Hervás-Martínez, J. Sánchez-Monedero, and P. A. Gutiérrez, "A Preliminary Study of Ordinal Metrics to Guide a Multi-Objective Evolutionary Algorithm," in *11th International Conference on Intelligent Systems Design and Applications (ISDA 2011)*, Cordoba, Spain, 2011.
- [78] E. Bissoondoyal-Bheenick, "An analysis of the determinants of sovereign ratings," *Global Finance Journal*, vol. 15, no. 3, pp. 251–280, 2005.
- [79] B. Rauch, M. Gottsche, G. Braehler, and S. Engel, "Fact and fiction in EU-governmental economic data," *German Economic Review*, vol. 12, no. 3, pp. 243–255, 2011.



Francisco Fernández-Navarro was born in Córdoba, Spain, in 1985. He received the B.S. degree in Computer Science from the University of Córdoba, Spain, in 2008 and the Ph. D. degree in Computer Science and Artificial Intelligence from the University of Málaga, Spain, in 2011. Currently, he is a Research Fellow in Computational Management at the European Space Agency (ESA), ESTEC. His current interests include radial basis functions neural networks, evolutionary computation and hybrid algorithms.



Pilar Campoy-Muñoz (Almería, 1982) received the B.A. degree in Business Administration in 2005 and M.A. degree in Research Methods on Economics and Business Sciences in 2011, both from ETEA, Business Administration Faculty in the University of Córdoba, Spain. She is currently pursuing the PhD degree and working as a research assistant in the Department of Economics at ETEA. Her current research interests include computational economics and financial economics, specially regarding to European Union countries.



Mónica-de la Paz was born in Córdoba (Spain) in 1973. She received a B.S. degree in Business Administration from the University of Córdoba (ETEA, Faculty of Economic and Business Sciences) in 1996, she has the equivalent to a Bachelor of Law Degree (B.L.) from the University of Córdoba (Faculty of Law) and she is currently a PhD candidate in the European Union Doctoral Programme (ETEA-LOYOLA, Department of Quantitative Methods) as well as she is cursing a M.S degree in Research Methods on Economics and Business Sciences at the Loyola Andalusia University (ETEA-LOYOLA, Faculty of Economic and Business Sciences). She initiated and continues her professional activities at the University of Córdoba and as Adviser in the Regional Government of Andalusia (Regional Ministry of Economy, Innovation and Science, Provincial Delegation). Her research interests are focused on the assessment of Research, Technological, Development and Innovation (RTDI) performance of European Public Policies through a variety of techniques as Data Development Analysis methodology, Artificial Neural Networks and modelling Bayesian Network causal models as a tool for public policy decision-making support.



César Hervás-Martínez was born in Cuenca, Spain. He received the B.S. degree in statistics and operating research from the Universidad Complutense, Madrid, Spain, in 1978 and the Ph.D. degree in mathematics from the University of Seville, Seville, Spain, in 1986. He is a Professor with the University of Córdoba in the Department of Computing and Numerical Analysis in the area of computer science and artificial intelligence and an Associate Professor in the Department of Quantitative Methods in the School of Economics. His current research interests include neural networks, evolutionary computation, and the modelling of natural systems.



Xin Yao (M'91-SM'96-F'03) received the B.Sc. degree from the University of Science and Technology of China (USTC), Hefei, Anhui, China, in 1982, the M.Sc. degree from the North China Institute of Computing Technology, Beijing, China, in 1985, and the Ph.D. degree from USTC in 1990, all in computer science.

He was an Associate Lecturer and Lecturer from 1985 to 1990 at USTC, while working towards his Ph.D on simulated annealing and evolutionary algorithms. He took up a Postdoctoral Fellowship in

the Computer Sciences Laboratory, Australian National University (ANU), Canberra, Australia, in 1990, and continued his work on simulated annealing and evolutionary algorithms. He joined the Knowledge-Based Systems Group, CSIRO (Commonwealth Scientific and Industrial Research Organisation) Division of Building, Construction and Engineering, Melbourne, Australia, in 1991, working primarily on an industrial project on automatic inspection of sewage pipes. He returned to Canberra in 1992 to take up a lectureship in the School of Computer Science, University College, University of New South Wales (UNSW), Australian Defence Force Academy (ADFA), where he was later promoted to a Senior Lecturer and Associate Professor. He moved to the University of Birmingham, U.K., as a Professor (Chair) of Computer Science in 1999. Currently, he is the Director of the Centre of Excellence for Research in Computational Intelligence and Applications (CERCIA) and a Changjiang (Visiting) Chair Professor (Cheung Kong Scholar) at the USTC. His major research interests include evolutionary artificial neural networks, automatic modularization of machine learning systems, evolutionary optimization, constraint handling techniques, computational time complexity of evolutionary algorithms, coevolution, iterated prisoner's dilemma, data mining, and real-world applications. He has more than 300 refereed publications.

Dr. Yao was the Editor-in-Chief of the IEEE TRANSACTIONS ON EVOLUTIONARY COMPUTATION (2003–2008), an associate editor or editorial board member of 12 other journals, and the Editor of the World Scientific Book Series on Advances in Natural Computation. He has given more than 50 invited keynote and plenary speeches at conferences and workshops worldwide. He was awarded the President's Award for Outstanding Thesis by the Chinese Academy of Sciences for his doctoral work on simulated annealing and evolutionary algorithms in 1989. He won the 2001 IEEE Donald G. Fink Prize Paper Award for his work on evolutionary artificial neural networks.

**EVOLUTIONARY NEURAL NETWORK CLASSIFIERS FOR
MONITORING RESEARCH, DEVELOPMENT AND
INNOVATION PERFORMANCE IN EUROPEAN UNION
MEMBER STATES ***

MÓNICA DE LA PAZ-MARÍN

*Department of Management and Quantitative Methods, ETEA, University of Córdoba,
C/ Escritor Castilla Aguayo 4, Córdoba, 14005, Spain*

PILAR CAMPOY-MUÑOZ

*Department of Economics, University of Córdoba, C/ Escritor Castilla Aguayo 4
Córdoba, 14005, Spain*

CÉSAR HERVÁS-MARTÍNEZ

*Department of Computing and Numerical Analysis, University of Córdoba, Campus of
Rabanales, Edificio Einstein, 3th floor, Córdoba, 14071, Spain*

The present work deals with the classification of the Research, Development (R&D) and innovation performance in 25 European Union (EU) Member States that follows the linear R&D model. As proxy indicators, expenditure in R&D, the personnel involved in these activities (inputs or enablers), patents as well as scientific publications (outputs), are taken into account, together with variables related to education and economy in order to classify R&D performance in 25 European Union (EU) Member States. This study classifies these countries from 2005 to 2008 employing a set of variables that characterize them and finding the most relevant ones for this classification. The Multilayer Perceptron Model (MLP) trained by Evolutionary Algorithms (also called ESUNN) and the Product-Unit Neural Network models trained by Evolutionary Algorithms (EPUNN) classified yearly country patterns in clusters, which had been previously obtained employing unsupervised algorithm k-means clustering that detected behavioural patterns among countries in the same cluster. Four different classes of national R&D performance are found through this algorithm: Low, Moderate, High and Innovation-driven countries. Finally, in order to analyse the appropriateness of our methodology, it is compared to other classification methods normally used in machine learning. The results show that while various methods of classification exist (like the one presented in this paper), our methodology obtains models with a significantly lower number of coefficients, without decreasing their accuracy and could be employed as a complementary tool to monitor R&D performance in the EU.

* This work was partially subsidized by the Spanish Inter-Ministerial Commission of Science and Technology under Project TIN2011-22794, the European Regional Development fund, and the “Junta de Andalucía”, Spain, under Project P08-TIC-3745.

1. Introduction.

Research, Development (R&D) and innovation have become, especially in these days of economic downturn and budget constraint, the keys for economic and smart growth in a knowledge-based society and the driving forces for national competitive advantage. Europe's competitiveness, our future standard of living, depends on our ability to incorporate the outputs of the innovation process in products, services, business and social processes and models. This is why R&D and innovation have been at the heart of, first, the Lisbon Strategy (2000), re-launched in Barcelona (2008) and, today, in the Europe 2020 Strategy and the European Research Area.

The R&D effort is a very complex structure with multiple factors to resource allocation strategies and to convert them into innovations. In response to competitive international pressure, firms' survival and competitive advantage rely upon R&D ability, and hence, upon innovation. The underlying argument for R&D public and private expenditure, is that the more resources a country invests in innovation, the higher its expected productivity [8]; it will be more competitive, leading to a more inclusive, knowledge-based and sustainable economy.

Two important underlying assumptions must be considered in this paper. The first is the technical change, which is found in Romer's model of economic growth [24], his most important work. This model is based on the underlying assumptions that: technical change lies at the heart of economic growth, that this change arises because of intentional actions taken by people who respond to market incentives and that instructions for working with raw materials are different from other economic goods.

Regarding the second premise, the linear innovation model, one of the first important steps towards the understanding of the innovation process was taken by Kline and Rosenberg in 1986 [13]. But although one of the key elements of that model of the innovation process is the understanding that research and innovation are linked in a non-linear and interactive manner, we assume the linear-model because, as Godin remarks, this model functions as a social fact. Rival and optional models can do little because of the lack of statistics [7].

All these factors result in an increasing demand for and interest in finding overall measures of efficiency and productivity in R&D investment and innovation. Scholars and experts started investigating the topic from different albeit complementary perspectives. Prior studies generally tested and evaluated the efficiency of R&D investments using parametric and non-parametric techniques (mainly Data Envelopment Analysis and Stochastic Frontier

Analysis) at macro [15, 20] and micro level (country) and micro level (firms and R&D programs and projects) and at different moments in the innovation process (largely divided in the literature in ex-post and ex-ante types). Although many researchers have studied the subject of competitiveness through innovation, most of the studies have been focused at firm level and seek to explain patterns of international competition, emphasizing the importance of the home country's characteristics in determining the competitive position of its firms in international markets.

At the macro-level, the European Innovation Scoreboard examines differences in R&D and innovation policies' efficiency by assuming that efficiency to be the ratio of outputs over inputs. Here it is measured by comparing the ratio between a composite indicator score for one or more input dimensions and one or more output dimensions. A wide range of evaluation methodologies and indicators for evaluating the socio-economic impact of R&D have been also presented as a toolbox document by the European Commission [6] but, to date, few works have been based on Artificial Neural Networks. Data Mining and Cluster Analysis techniques have been used on regional innovation systems [10] and for the competitiveness of nations [30].

The most popular neural network model could be the Multilayer Perceptron (MLP) due to its simple architecture yet powerful problem-solving ability, where neurons are grouped in layers and only forward connections exist [26].

However, alternatives to MLP have appeared in the last few years: Product Unit Neural Network (PUNN) models are an alternative to MLPs and are based on multiplicative neurons instead of additive ones. They correspond to a special class of feed-forward neural network introduced by Durbin and Rumelhart [5].

In many cases, neural networks that use sigmoid-unit basis functions[†] (SUNN) and networks that use product-unit basis functions (PUNN) are trained by using evolutionary algorithms (EA), obtaining significant advantages with this method compared to traditional training approaches [29]. One of the most important aspects of using of Evolutionary Programming (EP) algorithms as a modelling methodology is their ability to make the complexity of the network more flexible during training through the use of structural mutation operators that add and delete nodes and add and delete connections.

The aim of this work focuses on the prediction of the classification of the R&D performance in European countries thanks to their assignment into clusters, which will help to monitor European Strategies for R&D and innovation. First, k-means clustering (unsupervised algorithm) is applied to

[†] Called MLP or SUNN in this paper.

detect behavioural patterns among 25 EU Member States between 2005 and 2008. As a result of clustering, a number of classes are set to define the characteristics of each one. Then a multiclass classifier is built to assign each country-year observation to its corresponding cluster according to a set of specific features that characterize that country. To build that classifier, we develop neural network models using Sigmoid Unit Neural Networks (SUNN) and Product Unit Neural Networks (PUNN) trained with an EA. Thus, Evolutionary Algorithms for Artificial Neural Networks (EANN) are used to classify 25 European countries according to their R&D performance in one of the classes obtained thanks to the clustering procedure. These EANN use structural operation mutations that allow us to build models with a reduced number of basis functions and connections, which determine a lower number of coefficients in the model.

The study has the following structure: first, the methodology section is devoted to clustering methodology and class descriptions, neural networks, multiclass classification and evolutionary algorithm methodologies; second, the experimental study is carried out with the description of the database, the results and the analysis of the best model. Finally, conclusions are drawn about experimental results.

2. Methodology

2.1. *K-means clustering and classes description*

A simple K-means clustering algorithm [12] was used, implemented in WEKA software [10], due to its ease in implementation and low computational cost. Given that there was a significant decrease in the cluster sum of the squared errors when we increase k , a value of $k=4$ was chosen.

Using the economic analysis of those patterns and cluster centroids, the profile of each group was drawn up and labeled by classes as follows (for descriptions of variables, see Table I).

- Moderate innovation countries (cluster 1): countries in which R&D plays a less than central role. R&D intensity is not well developed either in the public and private sectors or, on average, in the number of researchers and human resources in Science and Technology (S&T). So, these countries have lagged behind in performance with very low patenting intensity and numbers of scientific publications. These countries are not economically strong (in terms of employment rates, GDP and real labor productivity growth) and give little emphasis to PhD graduates and life-long learning,

although they present good performance in community trademarks and license and patent revenues.

- Innovation-driven countries (cluster 2): this class involves the most developed countries. They are economically prosperous, and they present the largest proportion of investment in R&D, researchers, human resources in S&T and doctoral graduates. We also find emphasis on tertiary and life-long education, which is representative of their education systems. This has provided the workforce with the skills needed for a changing environment and has laid the groundwork for high levels of technological adoption and innovation. They present their best performances in scientific publications and patent revenues as well.
- Low innovation countries (cluster 3): mainly represented by the transition economies of the New Member States (except Greece and Portugal). As new economies, they involve considerable growth in GDP and in Real Labour Productivity growth-per-hour-worked (RLPGH) due to a high domestic demand following a period of socialism, among other reasons. Nonetheless, their economies do not seem to be based either on education or on R&D investment and could be an obstacle for future growth.
- High innovation countries or followers (cluster 4): countries in this cluster fit well into the model of technological competitiveness with high average values in most variables and are very close to cluster 2 in the number of PhD graduates, the employment rate, number of patents by residents, and medium and high-technology exports (these two latter cases hold the highest position). However, they present lower average values in terms of the profitability of their trademarks and license and patent revenues as well as in the number of persons enrolled in tertiary education. Lifelong learning is not as good as one might suppose in this group of countries. Thus, their performance economically and in education is worse than in the innovation-driven countries cluster, although they fall close behind.

Mapping these clusters (see figures 1-3), it can be observed that only two country-year patterns change cluster throughout the time span: Malta and Estonia. Both change from low to a moderate innovation cluster, but while the former moves in 2006, the latter moves in 2008.

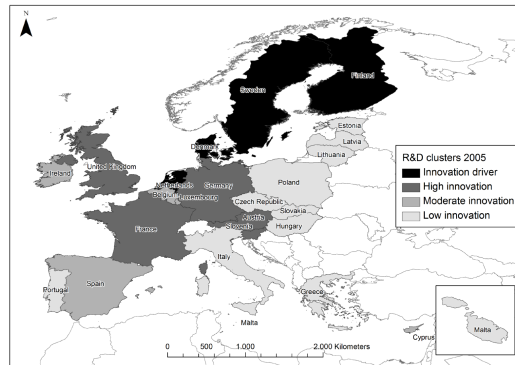


Figure 1. Clusters with country-year patterns 2005.

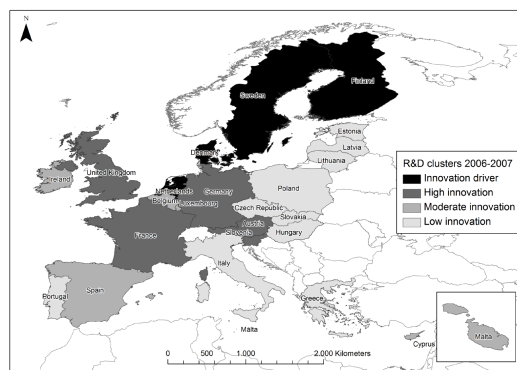


Figure 2. Clusters with country-year patterns 2006-2007

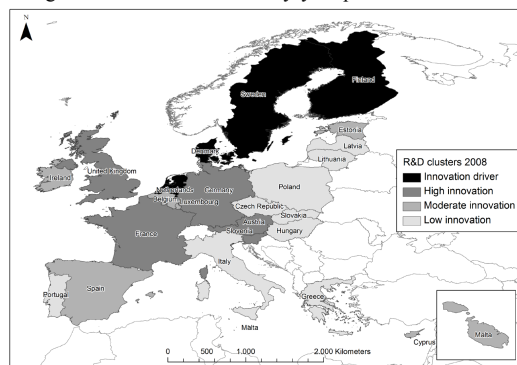


Figure 3. Clusters with country-year patterns 2008.

2.2. Neural Networks

Neural Networks (NNs) have become an important tool for classification since recent research activities have identified them as a promising alternative to conventional classification methods, such as the linear discriminant analysis or decision trees. Different types of NNs are now being used for classification purposes [16]: multilayer perceptron neural networks (MLP), where the transfer functions are Sigmoidal-Unit basis functions, SU; Radial Basis Functions (RBF), kernel functions where the transfer function are usually Gaussian [3]; the General Regression Neural Networks (GRNN) proposed by Specht [28]; a class of multiplicative NNs which comprises such types as sigma-pi networks; and Product-Unit networks, PU [5,19], where a multiplicative node is given as

$$y_i = \prod_{j=1}^k x_j^{w_{ji}}, \text{ where } k \text{ is the number of inputs.}$$

In order to solve the problem of selecting the proper size for the hidden layer of the net, a new category of algorithms has been introduced to automatically determine the structure of the network. These include the orthogonal least-squares algorithm; individual training of each hidden unit based on functional analysis; constructive methods, where the structure of the network is built incrementally; pruning methods that start with an initial selection of a large number of hidden units which is reduced as the algorithm proceeds; and the simultaneous selection of network structure and parameters by employing optimization methods based on genetic programming algorithms, EP. This training methodology will be applied in this research study on Evolutionary Sigmoidal-Units Neural Networks (ESUNN) and Evolutionary Product-Unit Neural Networks (EPUNN) models.

2.3. Multiclass classification

Let a training sample $D = \{(\mathbf{x}_n, \mathbf{y}_n); n = 1, 2, \dots, N\}$ be where $\mathbf{x}_n = (x_{1n}, \dots, x_{kn})$ is the vector of measurements taking values in $\Omega \subset \mathbb{R}^k$ and y_n is the class level of the n -th individual. We adopt the common technique of representing the class levels using a “1-of- Q ” encoding vector $\mathbf{y} = (y^{(1)}, y^{(2)}, \dots, y^{(Q)})$ such as $y^{(l)} = 1$ if \mathbf{x} corresponds to an example belonging to class l and $y^{(l)} = 0$ otherwise. Based on the training sample we wish to find a ANN ($\hat{\theta}$) that can result in a good classification of the patterns in the generalization set. This would be the same as obtaining a decision function $C: (\Omega, \theta) \rightarrow \{1, 2, \dots, Q\}$ to classify the individuals. A misclassification occurs when C assigns an individual (based on measurements vector) to class j when it is actually coming from a class $l \neq j$.

For a g ANN ($\hat{\theta}$) classifier, a classification problem with Q classes and N training or testing patterns is considered, obtaining a $Q \times Q$ contingency or confusion matrix $M(g)$:

$$M(g) = \left\{ n_{ij}; \sum_{i,j=1}^Q n_{ij} = N \right\} \quad (1)$$

where n_{ij} represents the number of times the patterns are predicted by classifier g to be in class j when they really belong to class i . Let us denote the number of patterns associated with class i by

$$f_i = \sum_{j=1}^Q n_{ij}, i = 1, \dots, Q \quad (2)$$

To evaluate the performance of the classifiers three scalar measures are defined that take the elements of the confusion matrix into consideration from different points of view. Let $S_i = n_{ii}/f_i$ be the number of patterns correctly predicted to be in class i with respect to the total number of patterns in i (sensitivity for class i). Therefore the sensitivity for class i estimates the probability of correctly predicting a class i example. From the above quantities the first comparison measure of a multiclass classifier define the sensitivity S of the classifier as the minimum value of sensitivities for each class $S = \min \{S_i; i=1, \dots, Q\}$. The second measure is the correct classification rate or accuracy,

$$CCR = (1/N) \sum_{i=1}^Q n_{ii} \quad (3)$$

that is the rate of all the correct predictions. The third measure is Cohen's Kappa coefficient, as an association measure between the class that *a priori* belongs to a pattern and the assignation that *a posteriori* assigns the classifier to that class.

Thus, we consider the three-dimensional measure (CCR, MS, K) associated with classifier g . The measure tries to evaluate three features of a classifier throughout the generalization set: global performance in the whole dataset, the performance in each class and the degree of association between pattern distributions in classes before and after classifier application.

The output layer of the ANN ($\hat{\theta}$) classifier is interpreted from the point of view of probability, which considers the soft-max activation function:

$$g_l(\mathbf{x}, \boldsymbol{\theta}_l) = \frac{\exp f_l(\mathbf{x}, \boldsymbol{\theta}_l)}{\sum_{i=1}^Q \exp f_i(\mathbf{x}, \boldsymbol{\theta}_i)} \text{ for } l = 1, \dots, Q \quad (4)$$

where $g_l(\mathbf{x}, \boldsymbol{\theta}_l)$ is the probability a pattern \mathbf{x} has of belonging to class l , $\boldsymbol{\theta}_l = (\boldsymbol{\beta}_l, \mathbf{w}_1, \dots, \mathbf{w}_M)$, $\boldsymbol{\beta}_l = (\beta_1^l, \dots, \beta_M^l)$ is the l -th vector of weights of the output node, M is the number of hidden nodes, $\mathbf{w}_j = (w_0^j, \dots, w_k^j)$ for $j = 1, \dots, M$, is the vector of input weights of the hidden node j , and $f_l(\mathbf{x}, \boldsymbol{\theta}_l)$ is the output of the l -th output node for pattern \mathbf{x} given by:

$$f_l(\mathbf{x}, \boldsymbol{\theta}_l) = \beta_0^l + \sum_{j=1}^M \beta_j^l \sigma_j \left(w_0^j + \sum_{i=1}^k w_i^j x_i \right) \text{ for } l = 1, \dots, Q-1$$

$$f_Q(\mathbf{x}, \boldsymbol{\theta}_Q) = 0 \quad (5)$$

where $\sigma(\bullet)$ is the sigmoidal activation function of the nodes in the hidden layer.

The classification rule coincides with the optimal Bayes rule. Thus the classification rule makes an individual be assigned to that class which has the maximum probability, given vector measurement \mathbf{x}

$$C(\mathbf{x}) = \hat{l}, \text{ where } \hat{l} = \arg \max_l g_l(\mathbf{x}, \hat{\boldsymbol{\theta}}_l), \text{ for } l=1, \dots, Q \quad (6)$$

2.4. Evolutionary Algorithm

The evolutionary algorithm for SUNN (ESSUN) is detailed in Figure 4, where p^B is the best-optimized ESUNN returned by the algorithm. The main characteristics of the algorithm are the following:

1. Representation of the individuals. The algorithm evolves architectures and connection weights simultaneously, each individual being a fully specified ESUNN. ESUNNs are represented using an object-oriented approach and the algorithm deals directly with the ESUNN phenotype. Each connection is specified by a binary value indicating if the connection exists along with a real value representing its weights.
2. Error and Fitness Functions. Our fitness function is a decreasing transformation of cross-entropy error because the Entropy function is continuous and helps the algorithm to enhance the classifiers more gradually than with the use of CCR. However, the relationship between CCR and Entropy error strongly depends on the dataset structure. Hence, regarding experimental results, using Entropy elitism is more suitable for some datasets to get higher test accuracy, but maintaining the best

CCR individual can be more appropriate for some other datasets [9]. In general, the relationship between CCR and cross-entropy error strongly depends on the data base structure. Hence, regarding experimental results, in general, using cross-entropy elitism is more suitable to obtain higher generalization accuracy. For this reason, the EP algorithm returns the best cross-entropy individual as solutions. Then, the error function given by cross-entropy error for N observations associated with the ESUNN model, is:

$$L^*(\boldsymbol{\theta}) = \frac{1}{N} \sum_{n=1}^N \left[-\sum_{l=1}^{Q-1} y_n^{(l)} f_l(\mathbf{x}_n, \boldsymbol{\theta}_l) + \log \sum_{l=1}^{Q-1} \exp f_l(\mathbf{x}_n, \boldsymbol{\theta}_l) \right] \quad (7)$$

where $y_n^{(l)}$ is equal to 1 if the pattern \mathbf{X}_n belongs to the l -th class and equal to 0 otherwise. The fitness measure needed for evaluating the individuals (Figure 4, steps 2, 7 and 16) is a strictly decreasing transformation of the error function

$$L^*(\boldsymbol{\theta}) \text{ given by } A(f) = \frac{1}{1 + L^*(\boldsymbol{\theta})} \text{ where } 0 < A(f) \leq 1.$$

3. Initialization of the Population. The initial population is generated trying to obtain ESUNNs with the maximum possible fitness. First, 5,000 random ESUNNs are generated (Figure 4, step 1), where the number of SUNNs M is a random value in the interval $[M_{\min}, M_{\max}]$. The number of connections between all SUNNs of an individual and the input layer is a random value in the interval $[1, k]$ and all of them are connected with the same randomly chosen input variables. In this way, all the SUNNs of each individual are initialized in the same random subspace of the input variables. A random value in the $[-1, 1]$ interval is assigned for the weights between the input layer and the hidden layer and in the $[-0, 0]$ interval for those between the hidden layer and the output layer. The individuals obtained are evaluated using the fitness function and the initial population is finally obtained by selecting the best 500 SUNNs (Figure 4, steps 2-4).

4. Parametric Mutation. This operator is accomplished for each coefficient of the model with Gaussian noise and applying a standard simulated annealing process for accepting or rejecting modifications (Figure 4, step 10) alters the value of the coefficients of the model. The connections are modified by adding Gaussian noise, $w(t+1) = w(t) + \boldsymbol{\xi}(t)$, where $\boldsymbol{\xi}(t) \in N(0, T(g))$ and $N(0, T(g))$ represent a one-dimensional normally distributed random variable with mean 0 and with variance in the network temperature ($T(g) = 1 - A(g)$) [18].

5. Structural Mutation. It implies a modification in the structure of the ESUNNs (Figure 4, step 11) and allows the exploration of different regions in the search space, helping to maintain the diversity of the population. There are four

different structural mutations similar to the mutations in the GNARL model [1]: hidden node addition, hidden node deletion, connection addition and connection deletion, which are applied sequentially to each network. The number of nodes added or deleted in hidden node addition and hidden node deletion is calculated as $\Delta_{\min} + uT(g)[\Delta_{\max} - \Delta_{\min}]$, u being a random uniform variable in the interval $[0, 1]$, Δ_{\min} and Δ_{\max} a minimum and maximum number of nodes specified as parameters. The severity of mutations depends on the temperature $T(g)$ of the neural network model, defined by $T(g) = 1 - A(g)$ $0 < T(g) < 1$. Structural connection mutations are performed as follows:

- Connection addition. Connection addition mutations are first performed in the hidden layer and then in the output layer. When adding a connection from the input layer to the hidden layer, a node from each layer is selected randomly, and then the connection is added with a random weight. A similar procedure is performed from the hidden layer to the output layer.
- Connection deletion. In the same way, connection deletion mutation is first performed in the hidden layer and then in the output layer, choosing randomly the origin node randomly from the previous layer and the target node from the mutated layer.

We apply connection mutations sequentially for each mutated neural net, first, adding (or deleting) $1 + u_0\Delta_0$ connections from the hidden layer to the output layer and then, adding (or deleting) $1 + u_h\Delta_h$ connections from the input layer to the hidden layer, u being a random uniform variable in the interval $[0, 1]$, Δ_0 and Δ_h previously defined ratios of the number of connections in the hidden and the output layer, and n_0 and n_h the current number of connections in the output and the hidden layers. Parsimony is also encouraged evolving networks by attempting the four structural mutations sequentially, where node 1 or connection deletion is always attempted before addition. Moreover, the deletion operations are made with higher probability. If a deletion operation is successful, no other mutation will be made. If the probability does not select any mutation, one of the mutations is chosen at random and applied.

ESUNN Algorithm:

Input: Training dataset (D)

Output: Best optimized SUNN (p^B)

1: $P^l \leftarrow \{p_1^l, \dots, p_{5000}^l\}$ p^l is a randomly generated SUNN

2: $\forall p_i^l \in P^l f_i^l \leftarrow A(p_i^l)$ {Evaluate fitness}

3: $P \leftarrow \{p_{(1)}, \dots, p_{(5000)}\}$, $(p_{(i)} \prec p_{(j)}) \Leftrightarrow (f_i^l > f_j^l)$ {Sort individuals in P^l by increasing f_i^l }

4: $P \leftarrow \{p_{(1)}, \dots, p_{(500)}\}$ {Retain the best 500 SUNNs}

6: **while not** Stop Condition **do**

7: $\forall p_i \in P$ $f_i \leftarrow A(p_i)$ {Evaluate fitness}

8: $P \leftarrow \{p_{(1)}, \dots, p_{(500)}\}$ $(p_{(i)} \prec p_{(j)}) \Leftrightarrow (f_i > f_j)$ {Sort individuals in P by increasing f_i }

9: $p^B \leftarrow p_{(1)}$ {Store Best Individual}

10: $P^P \leftarrow \{p_{(1)}, \dots, p_{(50)}\}$ {Parametric mutation parents (best 10% of individuals)}

11: $P^S \leftarrow \{p_{(1)}, \dots, p_{(449)}\}$ {Structural mutation parents (best 90% of individuals minus one)}

12: $\forall p_i^P \in P^P$, $p_i^P \leftarrow$ parametric Mutation (p_i^P) {Apply parametric mutation}

13: $\forall p_i^S \in P^S$, $p_i^S \leftarrow$ structural Mutation (p_i^S) {Apply structural mutation}

14: $P \leftarrow P^P \cup P^S \cup \{p^B\}$ {Offspring including the elite}

15: **end while**

16: $\forall p_i \in P$, $f_i \leftarrow A(p_i)$ {Evaluate fitness}

17: $P \leftarrow \{p_{(1)}, \dots, p_{(500)}\}$ $(p_{(i)} \prec p_{(j)}) \Leftrightarrow (f_i > f_j)$ {Sort individuals in P by increasing f_i }

18: $p^B \leftarrow p_{(1)}$

19: **return** p^B

Fig. 4. ESUNN training algorithm framework.

Although this section focuses on the ESUNN algorithm, for the EPUNN model we employ the same training algorithm framework but in the third step (initialization of the population) the random value in the $[-1, 1]$ interval assigned for the weights between the input layer and the hidden layer is much lower, usually $[-1, 1]$. More details about the EP algorithm can be consulted in [17] and [18].

3. Experimental study: predicting a country's R&D performance

3.1. Dataset description

The insights gained from the use of the EANN's algorithms were applied to a dataset of 25 EU Members for R&D performance classification. Due to data availability (until 2008 in most official databases), 100 country-year observations have been considered since no changes between clusters of country-year items are expected, similar way to some other studies [27]. The incorporations of Bulgaria and Romania are dated in 2007, thus scientific and technological European policy and strategy were not directly applicable and so their inclusion has not been considered appropriate for this classification.

The difficulty in finding data about R&D and innovation [2] as well as the missing data for a number of countries forces us to select proxy variables in most cases. Thus, fifteen main variables were selected, as shown in table 1, for characterizing R&D performance at country level following the input/output orientation in their interpretation. We ascribe to the idea that the production of knowledge requires specific investments in R&D and that the role of human capital is of great importance, thus the amount of funds and the number of researchers were selected as a proxy for labor input and capital input of R&D intensity, respectively. The R&D expenditure was broken down into public and business sectors.

Data were obtained from various sources (Eurostat's official website, OECD Main Science and Technology Indicators, World Bank Database and SCImago Journal & Country Rank in the case of scientific publications per country), adjusting when necessary to 1000 per population.

Despite this, an imputation has been carried out for missing data when necessary (it occurs in up to 1.3% of the cases); this imputation was made by linear regression [4]. Six models were fit for the tasks at hand, one for each variable with missing data, (PHD06, GERDGO, GERDBU, TPATRE, PATREV and RDPERS). The dependent variable is the year for which there is no data. The covariates are the remaining years without missing data. All models were fitted to $\alpha=0.05$ and a determination coefficient $R^2 > 0.90$.

Among the resulting set of variables of R&D activity [22], scientific publications have been considered as the major output of research and are widely used to evaluate the performance of researchers. Patents were also selected since they have often been used as the direct output of innovation process because technical advance and technological innovation are difficult to measure.

Table 1. List of variables

R&D enablers		
Code	<i>Variables</i>	<i>Source</i>
HUMLF	Human Resources in Science and Technology as a share of the labour force	Eurostat/ OECD
RDPER	R&D personnel, as a share of the labour force.	Eurostat
GERDBU	Gross Domestic Expenditure on R&D, Business Sector as % of GDP	Eurostat
GERDGO	Gross Domestic Expenditure on R&D, Government Sector as % of GDP	Eurostat
R&D outputs		
SPUBLI	Number of scientific publications per 1000 population	SCImago JCR
TPATRE	Worldwide patent applications filed through the Patent Cooperation Treaty procedure or with a national patent office (residents) per 1000 population	World Bank Database
Education		
PHD06	Number of PhD graduates (ISCED 6) per 1000 population	Eurostat
TERTIT	Number of persons who are enrolled in tertiary education (including university and non-university studies) in the regular educational system in each country per 1000 population	Eurostat
LLEARN	Lifelong learning (% of persons aged 18 to 64)	Eurostat
Economy		
GDPGRO	Growth rate of GDP volume - percentage change with respect to previous year	Eurostat
RLPGH	Real labour productivity growth per hour worked (previous year)	Eurostat
EMPLO	Total employment rate	Eurostat
PATREV	License and patent revenues from abroad as % of GDP	Eurostat
TRADEM	Community trademarks per billion GDP (in PPS€)	Innometrics /Eurostat
MHTEXP	Medium and High-technology exports (% of manufactured exports)	Innometrics /Eurostat

R&D performance includes the perspective of education with two variables, PhD graduates and tertiary education students, which is one of the main objectives of government R&D policies because human resources, with their managerial and organizational skills, play an ever-increasing role in the performance of R&D activities. Finally, economic variables like Gross Domestic Product (GDP) growth, per-hour-growth of real labour productivity

and medium and high technology exports, are taken into account to locate patterns among countries.

3.2. Results

We have compared Evolutionary Sigmoidal-Unit Neural Networks (ESUNNs) and Evolutionary Product Unit Neural Networks (EPUNNs) with nine state-of-the-art methods that are well-known in the literature. Eight of them have been configured and run in WEKA [11] and the Simple Linear Discriminant Analysis (SLDA) method is available in SPSS [21]. The methods used for comparison are [14]: MLP: a neural network classifier that uses back propagation to calculate weights. SLDA: a multivariate statistical procedure that derives equations to classify instances [25]. C4.5: a classifier tree for generating a pruned or unpruned decision tree. AdaBoost100: a classifier tree based on the AdaBoost.M1 algorithm with a maximum of 100 iterations. LMT: a classifier for building logistic model trees, which are classification trees with logistic regression functions at the leaves. NaiveBayes: a classifier whose numeric estimator precision values are chosen based on the analysis of the training data. SLogistic: a classifier for building linear logistic regression models. MLogistic: a classifier for building a multinomial logistic regression model with a ridge estimator. SVM: a classifier for building a linear model (maximum-margin-hyperplane) over a kernel space resulting in a nonlinear classifier. A Gaussian kernel is used with the selection of the SVM hyperparameters (regularization parameter C , and width of the Gaussian functions γ); a grid search algorithm a 10- fold cross validation, using the following ranges: $C \in \{2^{-5}, 2^{-3}, \dots, 2^{15}\}$ and

$$\gamma \in \{2^{-15}, 2^{-13}, \dots, 2^3\}$$

A comparison of the accuracy of eleven methods in predicting the R&D performance of each country is shown in Table 2. From a descriptive point of view, Best ESUNN, Best EPUNN, Best MLP, LMT, SLogistic, MLogistic and SVM models obtain equal results in CCR_G (100), MS_G (100) and K_G (1) for the generalization set. Nevertheless, concerning the number of connections of each model (see last column of Table II), the Best ESUNN model presents the lowest number of coefficients (20); thus this is the model selected, which will be further explained.

With respect to the results obtained in the 30 runs for the three stochastic algorithms (see Table 3), it is clear that the best results come from the MLP neural network, although the ESUNN methodology provides the second best results with a high mean (98% for CCR_G) and a low standard deviation, SD, (2.73% for CCR_G). The principal difference between these methodologies is the number of connections or model coefficients, because for MLP the mean is 44 and for ESUNN the mean is 26.57 with a SD= 3.47.

In this manner, considering accuracy and the lower number of connections, we recommend ESUNN methodology for country-year classification.

Table 2. Statistical results of the CCR_G , MS_G , K_G and number of connections (#conn.)

Method ^a	$CCR_G(\%)$	$MS_G(\%)$	K_G	#conn.
Best ESUNN	100	100	1	20
Best EPUNN	100	100	1	26
Best MLP	100	100	1	44
SLDA	96	80	0.94	24
C4.5	84	77	0.78	13
AdaBoost100	52	0	0.31	24
LMT	100	100	1	27
NaiveBayes	92	80	0.89	124
SLogistic	100	100	1	27
MLogistic	100	100	1	48
SVM	100	100	1	96

Table 3. Mean and Standard Deviation (SD) for stochastic methods

Method ^a	$CCR_G(\%)$	$MS_G(\%)$	K_G	# conn.
	Mean±SD	Mean±SD	Mean±SD	Mean±SD
ESUNN	98.00±2.73	92.33±9.60	0.97±0.05	26.57±3.47
EPUNN	94.00±4.79	84.55±11.13	0.92±0.07	29.93±4.66
MLP	100.00±0.00	100.00±0.00	1.00±0.00	44.00±0.00

a. The best quantitative result method is represented in bold face.

In the case of the stochastic algorithms shown in table 3, the experimental design was conducted using a holdout procedure with 30 runs. Approximately 75% of the patterns were randomly selected for the training set and the remaining 25% for the test set [23].

3.3. Best model

Once the algorithm has been run 30 times, the 30 best models are obtained, both for ESUNN and EPUNN. Then, the following ESSUN model is chosen: this model consists of nine variables in the input layer, two hidden nodes with sigmoid transfer function units and three nodes in the output layer which return the probability a pattern has of belonging to each class. From them, the probability of membership in the fourth class is calculated. The selected input variables were TERTIT, LLEARN, GERDGO, TRADEM, RLPGH, GPGROW, TPATRE, PATREV and HUMLF.

Interest in the proposed best ESUNN model is due to the fact that it improves the results of other approaches (see Table 3) and because it provides a (non-linear) interpretable model. Table 4 shows the expression of the probability of this best ESUNN model as well as its performance:

Table 4. Probability expression and performance of the best ESUNN model.

Best ESUNN model
$f_1(x, \theta_1) = 4.69 - 18.69 * SU_1$
$f_2(x, \theta_2) = 9.3 - 13.62 * SU_2$
$f_3(x, \theta_3) = 0.49 + 13.49 * SU_1 - 23.59 * SU_2$
$f_4(x, \theta_4) = 0$
$SU_1 = \frac{1}{(1 + EXP - (-9.32 + 8.07 * TERTIT + 9.31 * TRADEM - 4.08 * RLPGH + 5.03 * PATREV + 3.68 * HUMLF))}$
$SU_2 = \frac{1}{(1 + EXP - (-12.44 + 3.42 * TERTIT - 9.61 * LLEARN - 9.04 * GERDGO - 3.07 * GDPGRO + 7.32 * TPATRE + 4.47 * HUMLF))}$
CCR _G = 100% ; MS _G = 100%; K _G =1

Table 4 also shows the construction of sigmoid unit SU_1 as a function of variables TERTIT, HUMLF (labour force skills), PATREV, TRADEM (the profitability of innovation through patents and trademarks) and RLPGH (real labour productivity growth per-hour-worked). The same table shows the construction of SU_2 as a function of variables TERTIT, HUMLF (as well as SU_1) and LLEARN (which denotes the importance of education for labour force skills), GERDGO (public policy support for R&D investment that stimulates the innovation process in the private sector), TPATRE (main proxy indicator for measuring a country's innovation capability) and GDPGRO (the most common indicator of an economy's well-being).

On the other hand, if one looks at the coefficients of the sums of the exponents that belong to the sigmoidal units, along with the importance of variables according to the probability each one has of belonging to class (g_i) in Table 5, the variables which have the strongest effects for measuring R&D performance at country level are those that present (++) and (--). Variables with (+) or (-) have a lower effect and those with (=) have no significant effects. This means that an increase in variable value implies mainly an increase (+), a

decrease (-) or non-significant changes (=) in the probability of belonging to the corresponding cluster.

Table 5. Influence of each variable on a higher or lower probability of appearing in a given cluster

Variable	Probability Cluster 1	Probability Cluster 2	Probability Cluster 3	Probability Cluster 4
TERTIT	(--)	(--)	(+)	(++)
LLEARN	(--)	(++)	(+)	(--)
GERDGO	(-)	(++)	(-)	(--)
TRADEM	(--)	(--)	(++)	(++)
RLPGH	(+)	(=)	(=)	(--)
GDPGRO	(=)	(=)	(=)	(=)
TPATRE	(--)	(++)	(+)	(--)
PATREV	(--)	(--)	(++)	(++)
HUMLF	(-)	(=)	(-)	(=)

Thus, the variables that exert a major influence on the probability of belonging to cluster 2 (Innovation driven economies) are LLEARN, GERDGO and TPATRE. While these variables increase the probability of being included in cluster 2, TERTIT, TRADEM and PATREV tend to reduce it. The variables that have no effects on such a probability are GDPGRO, RLPGH and HUMLF.

Interestingly, the number of tertiary students (TERTIT) tends to strongly increase a country's R&D performance in cluster 4 (high innovation). Along with variable lifelong learning, the EU strategy could therefore continue to support education, the key to employment, as it is the most important base to increase necessary skills in knowledge-based economies.

4. Conclusions

A classification of European countries according to their R&D and innovation performance (based on the traditional R&D linear model) can be useful to assess their position compared to other countries with respect to reaching competitiveness and long-term sustainability objectives. In this study, after the use of unsupervised algorithm k-means clustering with four resulting clusters, we examined nine machine learning and two evolutionary algorithms as well as neural networks classifiers to predict the R&D performance

classification in 25 EU Member States, to identify the best model and major variables out of 15 that can influence the probability of belonging to each class.

Our analyses pinpointed the variables that exert a major influence on the R&D performance of a country belonging to “innovation-driven countries” (cluster 2), which are TERTIT, TRADEM and PATREV. While these variables improve R&D performance classification, LLEARN, GERDGO and TPATRE tend to reduce it. Interestingly, the number of tertiary students seems to tend to increase a country’s R&D performance since education is the base for increasing necessary skills in a knowledge-based economy. Nations will also need to consider investing in education in addition to R&D since these higher skills are basic in the application of new industry and service technologies.

This technique has to be considered a complementary tool to manage the extensive information provided by the number of official report indicators and methodologies. Our aim is to help experts to classify countries or predict their R&D performance in future years with innovation scores and publicly available country data, and then using the best model obtained, to identify the major contributing factors that could estimate the above-mentioned R&D and innovation performance with the fewest variables possible.

The present research demonstrates that nine out of fifteen variables like those in our dataset could be used to predict the R&D performance of a country reasonably well. Although our set of 15 variables resulted in very reasonable proxies, overlooked variables must be considered in future studies. If we have the appropriate numbers are available, the implications of our model would appear to be worthwhile in future comparative research performance on the cross-country performance, using nine out of a number of variables used in similar studies.

References

1. P.J. Angeline, G.M. Saunders and J.P. Pollackm, *IEEE Trans. Neural Netw.* **5**, 54 (1994).
2. B.C. Rao, *Int. J. Innov. Manage.* **14**, 823 (2010).
3. C.M. Bishop, *Neural Comput.* **3**, 579 (1991).
4. S.F. Buck, *J. R. Stat. Soc.* **B22**, 302 (1960).
5. R. Durbin and D. Rumelhart, *Neural Comput.* **1**, 133 (1989).
6. G. Fahrenkrog, W. Polt, J. Rojo, A. Tübke and K. Zinöcker, *Technical Report EUR20382EN*. (2002).
7. B. Godin, *Technol. Hum. Values* **31**, 639 (2006).
8. Z. Griliches, *R&D and Productivity: The Econometric Evidence*. University of Chicago Press (1998).

9. P.A. Gutiérrez, C. Hervás-Martínez and M. Lozano, *Soft Comput.* **14**, 599 (2009).
10. V. Hájková and P. Hájek, Proceedings of Communication and Management in Technological Innovation and Academic Globalization (COMATIA'10), 46 (2010).
11. M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann and I. Witten, *ACM SIGKDD Explorations Newsl.* **11**, 10 (2009).
12. J.A. Hartigan and M.A. Wong, *J. R. Stat. Soc.* **C28**, 100 (1979).
13. S.J. Kline and N. Rosenberg, The Positive Sum Strategy: Harnessing Technology for Economic Growth. National Academy Press, 275 (1986).
14. N. Landwehr, M. Hall and E. Frank, *Mach. Learn.* **59**, 161 (2005).
15. M. Kim, H. Lee, S. R. Yee and K. Choe, *Int. J. Innov. Technol. Manage.* **8**, 295 (2011).
16. R.P. Lippmann, *IEEE Commun. Mag.* **27**, 47 (1989).
17. A.C. Martínez-Estudillo, C. Hervás-Martínez, F.J. Martínez-Estudillo and N. García-Pedrajas, *IEEE Trans. Syst. Man Cybern.* **B36**, 534 (2006).
18. A.C. Martínez-Estudillo, C. Hervás-Martínez, F.J. Martínez-Estudillo and N. García-Pedrajas, *Neural Netw.* **19**, 447 (2006).
19. F.J. Martínez-Estudillo, C. Hervás-Martínez, P.A. Gutiérrez and A.C. Martínez-Estudillo, *Neurocomputing* **72**, 548 (2008).
20. W. Nasierowski, *Int. J. Innov. Technol. Manage.* **7**, 389 (2010).
21. M. Norusis, SPSS 15.0 Advanced Statistical Procedures Companion. Prentice Hall Press (2007).
22. OECD. Main Science and Technology Indicators. OCDE (2010).
23. L. Prechelt, Tech. Rep. Karlsruhe University **21/94**, (1994).
24. P.M. Romer, *J. Polit. Econ.* **98**, S71 (1990).
25. D.J. Sheskin, Handbook of parametric and nonparametric statistical procedures. Chapman and Hall/CRC (2011).
26. K.A. Smith and J.N.D. Gupta, *Comput. Oper. Res.* **27**, 1023 (2000).
27. F. Solt, *Social Sci.* **90**, 231 (2009).
28. D.F. Specht, *IEEE Trans. Neural Netw.* **2**, 568 (1989).
29. X. Yao and Y. Liu, *IEEE Trans. Neural Netw.* **8**, 694 (1997).
30. S. H. Zanak and I. Becerra-Fernandez, *Eur. J. Oper. Res.* **166**, 185(2005).

2. TRABAJOS ENVIADOS Y EN FASE DE REVISIÓN

A continuación, se adjuntan los artículos derivados de la presente tesis y que se encuentran en fase de revisión en las correspondientes revistas:

Revista: *Research Policy* (Índice de Impacto 2012: 2,520)

Evaluation of Research, Technological Development and Innovation public policies in the European Union: a Bayesian network model and the Two-Linked-Steps Data Envelopment Analysis approach.

Autores: Mónica de la Paz-Marín; Carlos R. García-Alonso.

Revista: *European Journal of Operational Research* (IF:2,420)

Evaluation of countries' progress toward a Knowledge Economy based on machine learning classification techniques.

Autores: Mónica de la Paz-Marín; Pedro A. Gutiérrez-Peña; César Hervás-Martínez.

Manuscript Number: EJOR-D-13-00427

Title: Prediction of countries' progress toward a Knowledge Economy based on machine learning classification techniques.

Article Type: Innovative Application of OR

Section/Category: Decision support systems

Keywords: Decision support systems; machine learning; Knowledge Economy; hierarchical clustering; nominal classification; ordinal classification.

Corresponding Author: Mrs. Mónica de la Paz-Marín,

Corresponding Author's Institution: University of Córdoba

First Author: Mónica de la Paz-Marín

Order of Authors: Mónica de la Paz-Marín; Pedro Antonio Gutiérrez-Peña, PhD; César Hervás-Martínez

Abstract: Knowledge is still a key factor for competitive advantages in the current economic crisis and uncertain environment. There are many indicators to measure knowledge advances at both micro and macro levels but the benefits for stakeholders and policy makers are still limited due to the absence of predicting models. This paper introduces an approach to predict with a high degree of accuracy the progress toward a Knowledge Economy (KE) in 54 countries (years 2007-2009). Three nominal and three ordinal classifiers from machine learning field were compared in order to select the corresponding classifier and to verify the nature of the problem. Previously, the groups for classification purposes were obtained by means of a hierarchical clustering (unsupervised algorithm): advanced, followers, moderate, and early knowledge economies. Then, an ordinal model based on the Support Vector Ordinal Regression with Implicit constraints (SVORIM) was selected as the best performing one due to its ability to classify the patterns into one of four clusters obtained with a high degree of accuracy, and we conclude the problem presents an ordinal nature. The clusters reflect the overall KE position of countries. The proposed ordinal model that predicts their classification could be used as a decision support tool for monitoring the progress or stage of transition to a KE once the data of new countries (or a new year) are available.

Cover Letter

Dear Editor-in-Chief of European Journal of Operational Research,

On behalf of my co-authors, I am pleased to submit the enclosed material for possible publication in your journal. It has not been submitted for other publication nor has it been published in whole or in part elsewhere by any of the authors. I can attest to the fact that all authors listed on the title page have read the manuscript, vouch for the validity and legitimacy of the data and its interpretation, and agree to its submission.

Mónica de la Paz Marín

February 20th, 2013

DETAIL OF SUBMITTED ARTICLE

Manuscript title:	Evaluation of countries' progress toward a Knowledge Economy based on machine learning classification techniques
Corresponding author:	Mónica de la Paz-Marín
Co-authors:	Pedro Antonio Gutiérrez-Peña César Hervás-Martínez
Type of manuscript:	Innovative Application in OR
Appropriate for this journal:	This article deals with the stage of transition toward a Knowledge Economy in 54 countries, showing practical implications and a decision support tool for policy makers at country level. Through hierarchical clustering technique, behavior patterns of a Knowledge Economy have been identified in those countries. Then three nominal and ordinal classifiers were compared to build a model that classifies countries by those patterns, once the ordinal nature of the problem was determined. In this manner, this tool aims to be useful to policy makers in order to identify key features in the transition to a Knowledge Economy, for monitoring such transition, and to achieve smart growth and competitive advantages of countries.
Contact information (corresponding author)	Mónica de la Paz-Marín Department of Management and Quantitative Methods ETEA, Business Administration Faculty Escritor Castilla Aguayo 4, 14004-Córdoba, Spain Telephone number: +34.957.222100 Fax: +34.957.218030 e-mail: mpaz@uco.es
Possible conflicts of interest:	To the date, there are no conflicts of interest.
Detail of Financial Support	This work was partially subsidized by the Spanish Inter-Ministerial Commission of Science and Technology under Project TIN2011-22794, the European Regional Development fund, and the "Junta de Andalucía", Spain, under Project P08-TIC-374.

Highlights

This paper is focused on transition toward a Knowledge Economy (KE) of countries.>Behaviour patterns on KE of these countries are identified and defined.>Nominal and ordinal classifiers are compared to determine the nature of the problem.>The best ordinal model according its best accuracy is chosen.>The model could be used to predict transition toward KE of countries.

Prediction of countries' progress toward a Knowledge Economy based on machine learning classification techniques

Mónica de la Paz-Marín (corresponding author)^a, Pedro Antonio Gutiérrez Peña^b, César Hervás-Martínez^b

^a *Department of Management and Quantitative Methods ETEA, Business Administration Faculty, Escritor*

Castilla Aguayo 4, 14004 Córdoba, Spain. Telephone number: +34.957.042226; Fax:

+34.957.218030{mpaz@uco.es}

^b *Department of Computing and Numerical Analysis, University of Córdoba, Albert Einstein Building, 14071*

Córdoba, Spain {pagutierrez@uco.es;chervas@uco.es}

Abstract

Knowledge is still a key factor for competitive advantages in the current economic crisis and uncertain environment. There are many indicators to measure knowledge advances at both micro and macro levels but the benefits for stakeholders and policy makers are still limited due to the absence of predicting models. This paper introduces an approach to predict with a high degree of accuracy the progress toward a Knowledge Economy (KE) in 54 countries (years 2007-2009). Three nominal and three ordinal classifiers from machine learning field were compared in order to select the corresponding classifier and to verify the nature of the problem. Previously, the groups for classification purposes were obtained by means of a hierarchical clustering (unsupervised algorithm): advanced, followers, moderate, and early knowledge economies. Then, an ordinal model based on the Support Vector Ordinal Regression with Implicit constraints (SVORIM) was selected as the best performing one due to its ability to classify the patterns into one of four clusters obtained with a high degree of accuracy, and we conclude the problem presents an ordinal nature. The clusters reflect the overall KE position of countries. The proposed ordinal model that predicts their classification could be used as a decision support tool for monitoring the progress or stage of transition to a KE once the data of new countries (or a new year) are available.

Keywords: *Decision support systems; machine learning; Knowledge Economy; hierarchical clustering; nominal classification; ordinal classification.*

1. INTRODUCTION

1 Academics, policy-makers, stakeholders, consultants and the media have shown a
2 growing interest in the relevance of knowledge creation as a key factor in production of,
3 enabling an increase in the competitive advantages of firms and, consequently, of national
4 economies (von Krogh et al., 2000). That is especially crucial in the uncertain, changing,
5 ambiguous and complex environment that characterizes nations today (Johannessen and
6 Olsen, 2010). Thus, it can be said that we are merged into the so-called knowledge economy
7 or knowledge-based economy (KE), seen as the stage that follows the industrial era, which
8 has become almost an imperative for nations, stressing even more the role of innovation in the
9 efforts to achieve competitiveness and a sustainable economic development. The fact that
10 governance bodies hold knowledge to be a central strategy also reveals the relevance of
11 achieving this type of growth model.

12 In the last two decades, literature and research related to the KE have merged (Aghion
13 and Howitt, 1998; David and Foray, 2002; Drucker, 1993; Grossman and Helpman, 1991;
14 Leydesdorff, 2006; OECD, 1996; Thurow, 1999), focusing mainly on the important role of
15 knowledge or human capital as a source of the long-term economic growth. The relevance of
16 knowledge is clearly linked to the new growth theory, which considers knowledge (or human
17 capital) to be an endogenous variable of the economic growth. Knowledge is regarded as the
18 basic form of capital, and economic growth is driven by the accumulation of knowledge
19 (Romer, 1990; Lucas, 1988). Other economic theories have appeared to examine this
20 phenomenon: the evolutionary theory of economic change (Nelson and Winter, 1986), the
21 triple helix theory (Etzkowitz and Leydesdorff, 2000), the national innovation systems theory
22 (Freeman, 1987) or the knowledge gap theory (Abramovitz, 1986).

23 There are several real examples of the influence of knowledge on current economic
24 growth: (1) progress in information and communications technology (ICT) that enables cheap
25 and rapid access to knowledge and information, (2) the ever-increasing speed of scientific and
26 technological advances, (3) global competition and (4) the new demands, tastes and customs
27 of citizens. Based on all these, the World Bank Institute emphasizes that most of the countries
28 that have made rapid progress have staged nationwide KE-inspired programs of change
29 (IBRD/World Bank, 2007).

30 The concept of knowledge economy (KE), although it could have its roots in Adam
31 Smith's work, it was possibly first used by Machlup (1962) and coined by Peter Drucker
32 (1969) and an object of special attention in the KE report of the Organization for Economic
33 Co-operation and Development (OECD). According to this report, knowledge economies are
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 those “which are directly based on the production, distribution and use of knowledge and
2 information” (OECD, 1996, pp. 7). The definition of the World Bank is also a widely used
3 setting that is “essentially an economy where knowledge is the main engine of economic
4 growth” (Chen and Dahlman, 2006, pp. 1). In these economies, the emphasis is placed on
5 intellectual capabilities rather than on physical factors (Powell and Snellman, 2004).
6
7

8
9 In this context, governments have to plan investments, and develop strong education
10 systems to train highly-skilled workers for high-skilled jobs if they want to achieve a
11 knowledgeable society (Hsu et al, 2008). Measurement tools, frameworks, models and
12 methodologies help stakeholders to analyze and benchmark the capabilities of countries as
13 knowledge-based economies. Such assessments facilitate the adoption of policies as well as
14 the creation of national knowledge systems for holistic development. However, these
15 indicators yield different scores and rankings depending on the nature and type of
16 assessments, they report on past performance (Al Shami et al, 2012), which involves many
17 questions that have to be answered subjectively (Booyesen, 2002) and they do not predict
18 where a certain KE is heading (or could head) if all variables employed in the model were
19 known.
20
21

22
23 Classification, in general, is one of the most frequent decision-making tasks in human
24 activity. A classification problem occurs when an object needs to be assigned into a
25 predefined (supervised) class based on a number of observed attributes related to that object.
26 Most classification algorithms focus on predicting data labels from nominal non-ordered
27 classes and real problems are tackled as nominal classification. However, many multi-criteria
28 classification problems involve classifying data into classes that have a natural order (ordinal
29 problem) (Zopounidis and Doumpos, 2002).
30
31

32
33 Ordinal classification techniques have broad applications where it is natural to rank
34 instances such as information retrieval (Herbrich et al., 1999; Chu and Keerthi, 2007),
35 econometric modeling (Mathieson, 1995), credit risk (Doumpos, 2002; Xu et al., 2009) or gen
36 analysis (Pyon and Li, 2009), to name a few.
37
38

39
40 In this study, the problem has been addressed by comparing both nominal and ordinal
41 classifiers in order to test which model performs the best, because the nominal or ordinal
42 nature of the clusters obtained is not obvious or “natural”. A priori, the dependent variable
43 (the cluster or class previously obtained) would have an ordinal consideration as it can be
44 seen in the myriad of examples that imply the ranking of countries in socioeconomic issues:
45 the current Rating Agencies (i.e., Moody’s, Standard and Poor’s or Fitch), the Global
46 Competitiveness Index of the International Monetary Fund, the Knowledge Economy Index
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 of the World Bank, the Innovation Union Scoreboard of the European Commission, the
2 Human Development Index of the United Nations, the ranking of universities, and so on.

3
4 In accordance with all the above mentioned, the aim of this work focuses on obtaining
5 a model to predict the classification of 162 country-year observations (54 countries in the time
6 span of 2007, 2008 and 2009) thanks to their previous assignment into clusters. These clusters
7 were previously defined in relation to their progress to a KE by using eleven major KE
8 indicators that represent four KE pillars. Thus, a hierarchical clustering (an unsupervised
9 algorithm) was applied to detect behavioral patterns. As a result, a number of clusters were set
10 and the characteristics of each one were defined. Then, three nominal and three ordinal
11 classifiers were built to assign each country–year observation to its corresponding cluster. The
12 results of ordinal algorithms were compared to those of nominal classifiers to evaluate which
13 of them performed the best and to determine the nominal or ordinal nature of the problem.
14
15

16
17 The main conclusion is that ordinal classification algorithms perform better than
18 nominal ones if an order in the clusters is taken into account, so the Support Vector SVORIM
19 ordinal model is selected to predict the classification of countries into one of the clusters
20 obtained. This model will help to monitor national strategies and some key features related to
21 knowledge creation and innovation, in general terms.
22

23
24 After the introductory Section 1, we briefly review the relevant literature on the
25 knowledge economy assessment of countries and classification methodologies in Section 2.
26 Then, the methodology applied in this study is set out in detail in Section 3; in Section 4, the
27 experimental study is carried out with the description of the dataset and variables employed as
28 well as the main results, and these are discussed. Finally, main conclusions are drawn in
29 Section 5.
30

31
32
33
34
35
36
37
38
39
40
41
42 **2. REVIEW OF METHODOLOGIES FOUND IN THE LITERATURE**

43
44 In relation to the KE assessment literature to date, empirical studies form two main
45 groups: statistical techniques and composite indicators or indices. Composite indicators are
46 increasingly being recognized as a useful tool for policy-making at national level in setting
47 policy priorities, in benchmarking or monitoring performance. A considerable number of
48 them have been proposed by international bodies (Nardo et al., 2005).
49
50

51
52 The modeling and measurement of a KE follow prior efforts to measure related
53 components such as regional human development and the building of intellectual-capital
54 frameworks (Hanley and Malafsky, 2002). Several knowledge assets measurement models
55 have also been proposed by reputable organizations, i.e., the United Nations Economic
56
57
58
59
60
61
62
63
64
65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Commission for Europe (UNECE) or the eEurope National Knowledge Assets Measurement models.

Of these, the World Bank Knowledge Assessment Methodology (KAM) is one of the main knowledge economy initiatives in the world, providing a country Knowledge Economy Index (KEI). Briefly, the methodology consists of a set of 148 structural and qualitative variables; they note that it can be used for benchmarking (World Bank, 2012; Chen and Dahlman, 2006). A reduced number of variables serve as proxies for the four relevant pillars in the development of a knowledge economy (Economic Regime, Innovation System, Education and the Information and Communication Infrastructure) providing of a basic scoreboard.

Regarding to the methodology proposed in this study, multi-class pattern recognition in general has become increasingly popular. One of the most widely employed methodologies is multi-class nominal or ordinal classification. On one hand, in nominal classification learning problems, no order is found between the classes. The most popular approach is to decompose the problem into multiple two-class classification problems.

On the other hand, in ordinal classification problems¹, target categories present an order. Due to the various synergies between data mining and operational research domains (Corne et al., 2012), a great deal of effort has been devoted to the problem of ordinal classification in operations research and computational intelligence fields, where the learning of ordinal classification models has been seen as a generalization of some multi-criteria nominal techniques and, more recently, has led to new theoretical developments². This problem appears as different from standard regression because a distance between the labels cannot be established.

The present study employs a group of algorithms known by the term of *threshold models* that group the majority of the proposals for ordinal regression. They focus on two main issues (Fouad and Tino, 2012): i) how to find the optimal projection line, representing the assumed linear order of classes, onto which the input data will be projected; ii) how to optimally position thresholds defining the label intervals so that the margin of separation between neighboring classes is maximized.

Arguably, the most well-known group of ordered response models is based on the estimation of cumulative probabilities (Fullerton and Xu, 2012). The Proportional Odd Model (POM), is one of the first models specifically designed for ordinal regression. This is one of

¹ This kind of problems is also called *ordinal regression*.

² For a review of multicriteria models for learning ordinal data, see Sousa et al., 2013.

1 the extended methods for analyzing binary data to models with ordered responses. Various
2 authors, especially McCullagh (1980), are often credited with the idea of using the logit link
3 for ordered response models, and thereby the POM.
4

5 In the context of Support Vector Machines (SVMs), a class of models has been
6 developed under the name of support vector ordinal regression (SVOR). Shashua and Levin
7 (2003) proposed a generalized formulation of SVMs applied to ordinal data based on two
8 large-margin principles: i) the fixed-margin principle, in which the margin of the closest pair
9 of classes is maximized, leading to equal margins between two neighboring classes and ii) the
10 sum of margins principle, which allows for different margins and only the sum of all $Q-1$
11 margins is maximized (assuming there are Q ordered categories). However, the order on the
12 $Q-1$ class thresholds was not imposed, this work being extended in the SVOR with EXplicit
13 ordering constraints (SVOREX) formulation where the order of class thresholds is considered
14 explicitly (Chu and Keerthi, 2007). Furthermore, the authors also presented an alternative
15 SVOR model named SVOR with IMplicit ordering constraints (SVORIM). Both SVOREX
16 and SVORIM are the two additional methods selected in this work for comparison.
17

28 **3. METHODOLOGY**

29 *3.1. Hierarchical Clustering and class description*

30 In many real problems, decision makers have to group objects into homogeneous
31 classes. Cluster analysis allows the assigning of a set of data into groups, so that the data in
32 the same cluster are more similar to each other than to those in other clusters and each data
33 element belongs to one cluster. In order to identify an appropriate number of clusters, a
34 variety of clustering algorithms have merged in the last years, but the most relevant could be
35 grouped in *partitional* and *hierarchical* clustering (Lin, 2012; Meyer, 2013).
36

37 Our initial task is to employ a cluster technique to detect patterns in the selected
38 countries with respect to their stage of transition to a KE. In the absence of class labels or
39 knowledge about the clusters present, it is very difficult to select a criterion to judge whether
40 one subset of features is better than another. The technique adopted in this study is
41 hierarchical clustering (an unsupervised machine learning technique), which is used in many
42 fields and often preferred to non-hierarchical clustering when the number of clusters Q is not
43 known or when the interest lies in the relationships between the objects, such as in
44 taxonomical studies (Questier et al., 2002).
45

46 Hierarchical clustering provides insight into the data by assembling all the objects into
47 a dendrogram, so that each sub-cluster is one of its nodes, and the combinations of sub-clusters
48 create a hierarchy—a structure that is more informative than unstructured set of clusters in
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

partitioned clustering (Wu et al, 2009). As other clustering techniques, one set of attributes is considered for both partitioning the data space and measuring the similarity between objects (Chen et al., 2006).

The hierarchical clustering technique was used along with the COBWEB algorithm due to its low computational cost with WEKA software (Hall et al., 2011). The COBWEB algorithm was developed by machine learning researchers to cluster objects in an object-attribute data set yielding a clustering dendrogram called a classification tree that characterizes each cluster with a probabilistic description. This algorithm operates based on the so-called category utility function (CU) that measures clustering quality. The resulting classification tree of hierarchical clustering or dendrogram can be seen in Figure 1. For each node, the identifier of the cluster corresponds to the number without brackets and the number in brackets is the number of country-year patterns included in the corresponding cluster. Although many clusters were derived from the Cobweb algorithm, we decided to group them in four final classes, with the help of the hierarchical structure found by the algorithm.

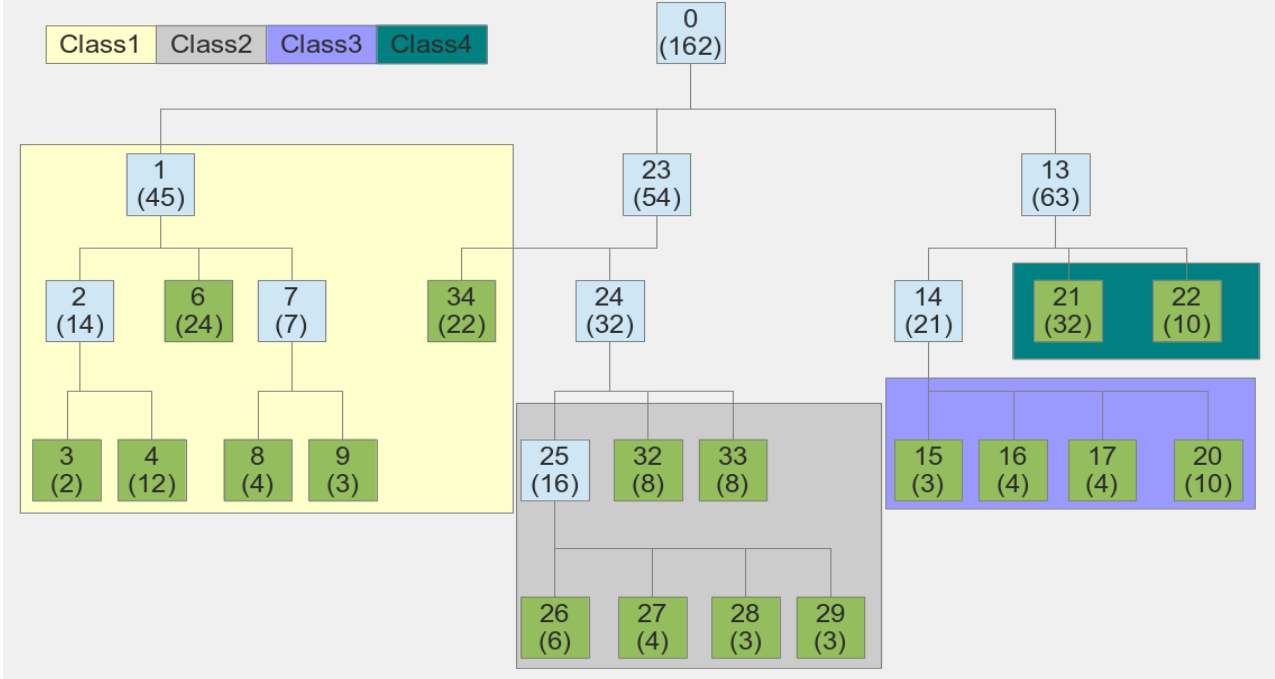


Figure 1: Dendrogram from hierarchical clustering

Table 1. Description of clusters (mean, standard deviation and country-year pattern of reference).

CODE	VARIABLE	Mean±SD (country-year pattern of reference for each cluster and variable)			
		Cluster 1 ^(a)	Cluster 2 ^(b)	Cluster 3 ^(c)	Cluster 4 ^(d)
Pillar 1: Economic and institutional regime					
<i>TNTBA</i>	Tariff and Non-tariff Barriers	85.149±4.575 (Japan-09)	85.400±1.975 (Italy-07)	83.324±5.271 (Cyprus-07)	75.638±9.605 (Chile-08)
<i>REGQU</i>	Regulatory Quality	1.523±0.291 (Korea Rep.-08)	1.109±0.262 (Croatia-09)	0.721±0.550 (Maced. FYR-08)	0.009±0.539 (Tunisia-09)
<i>RULEL</i>	Rule of Law	1.640±0.298 (Latvia-09)	0.912±0.383 (Hungary-09)	0.300±0.568 (Greece-07)	-0.394±0.502 (Chile-07)
Pillar 2: Education and skills					
<i>SCHOO</i>	Primary enrollment (% gross)	102.379±3.976 (Germany-09)	103.585±6.220 (Hungary-09)	99.155±4.520 (Bulgaria-07)	103.323±7.803 (Romania-07)
<i>SECON</i>	Secondary Enrollment (% gross)	105.982±11.667 (United Kingdom-07)	100.271±8.094 (Lithuania-09)	92.691±5.255 (Cyprus-07)	87.525±5.333 (Belarus-07)
<i>TERTI</i>	Tertiary Enrollment (% gross)	66.122±17.479 (Japan-07)	64.153±12.894 (Czech Rep.-08)	56.251±13.201 (Chile-09)	42.699±19.073 (Romania-07)
Pillar 3: Information and communication infrastructure					
<i>TELEP</i>	Telephones	1,619.142±239.961 (Korea Rep.-07)	1,587.687±160.324 (Hungary-07)	1,387.418±207.976 (Belarus-09)	1,035.833±288.396 (Ukraine-07)
<i>FIXBI</i>	Fixed broadband Internet subscribers	280.205±49.153 (Estonia-07)	177.438±29.787 (Latvia-08)	107.771±16.798 (Cyprus-08)	41.149±28.340 (Russian Fed.-09)
<i>INTERN</i>	Internet Users	768.438±82.969 (Malta-09)	534.872±73.289 (Spain-07)	457.583±114.678 (Poland-08)	235.764±94.148 (Albania-09)

Pillar 4: Innovation system					
<i>PATEN</i>	Patent Applications	1.630±1.224 (United Kingdom-09)	0.372±0.413 (Ireland-07)	0.099±0.093 (Cyprus-08)	0.075±0.079 (Belarus-08)
<i>STJOU</i>	Scientific and Technical Journal Articles	1.864±0.706 (Estonia-08)	1.104±0.509 (Czech Rep.-08)	0.591±0.308 (Greece-07)	0.169±0.100 (Turkey-09)

^(a) Cluster 1 groups Australia, Austria, Belgium, Canada, Denmark, Estonia, Finland, France, Germany, Hong Kong SAR, China, Iceland, Ireland-09, Japan, Korea Rep., Latvia-09, Luxembourg, Malta-09, Netherlands, New Zealand, Norway, Slovenia-09, Sweden, United Kingdom and the United States.

^(b) Cluster 2 groups Croatia-09, Cyprus-09, Czech Republic, Greece-09, Hungary, Ireland-07-08, Israel, Latvia-07-08, Lithuania, Malta-07-08, Portugal, Slovenia-07-08 and Spain.

^(c) Cluster 3 groups Belarus-09, Bulgaria, Chile-09, Croatia-07-08, Cyprus-07-08, Greece-07-08, Macedonia FYR-08-09, Poland, Romania-08-09 and the Slovak Republic.

^(d) Cluster 4 groups Albania, Armenia, Azerbaijan, Belarus-07-08, Bosnia and Herzegovina, Chile-07-08, China, Georgia, Macedonia FYR-07, Mexico, Moldova, Romania-07, Russian Federation, Tunisia, Turkey and Ukraine.

The groups are referred as classes in the figure, and they were decided observing the country-year observations trying to reflect a degree of transition to a KE economy, and also trying to obtain representative groups with a reasonable number of elements in each class or group. A further analysis of these clusters is made based on their centroids (Table 1).

Using an economic analysis carried out by experts as complementary to the present hierarchical clustering, four clusters were selected and the profile of each group was drawn and labeled. Regarding the number of clusters selected, four groups achieve greater differentiation; this avoids the need for massive information that a very fine-grained measure would require if a greater number of clusters were chosen.

The main findings of the clustering analysis are that, based on the four main pillars (Economic Incentive, Innovation, Education and ICTs) and across 11 indicators, the countries selected fall into four groups or clusters according to the stage of their transition to a knowledge-based economy (for a detailed description of variables see Table 2 in Section 4.1):

Cluster 1 (*advanced KEs*). It groups the country-year patterns that perform the best in the four pillars. They are economies with high-quality legal frameworks, which provide effective protection for property rights and strong support for the rule of law regulations, intolerance for corruption, regulatory efficiency and openness to strong global commerce with

1 an overall macroeconomic stability that minimizes inherent uncertainty. They show excellent
2 levels of patent applications and scientific publications (i.e., Switzerland, Japan and Korea)
3 and a strong educational system, especially in the tertiary education level, a key element for a
4 highly-skilled future labor force. They are very strong in information and communication
5 technologies (ICT) with values far above the cluster centroid in relation to the number of
6 telephone lines, Internet users, and fixed broadband Internet subscribers variables.
7
8

9
10
11 **Cluster 2 (*followers*).** These country-year patterns follow those previously mentioned
12 in cluster 1 according to the mean of each variable in this cluster with respect to the
13 corresponding cluster centroid, and they are far behind the first cluster, particularly at the
14 level of patents and scientific publications. They have a legal, institutional and commercial
15 environment with values very close to the overall average and with gross enrollment rates in
16 primary, secondary and tertiary levels that are close to the values in the first cluster, except in
17 the primary education variable where cluster 2 has a slightly higher average. Although they
18 follow in second place, they are far from the patenting, scientific publications and ICT levels
19 of the first cluster.
20
21

22
23
24
25
26
27 **Cluster 3 (*moderated KEs*).** It groups the country-year patterns that show low
28 performance in innovation (low number of patent applications and half the scientific
29 publications), with a lower gross ratio of enrollment in each level of education (primary,
30 secondary and tertiary). In relation to the law-rule variable that measures the extent to which
31 agents have confidence in and abide by the rules of society, it shows very small values on
32 average, but they are not negative. The countries grouped in this cluster also show low values
33 in the variables representing the pillar of ICT. In spite of this, it includes the countries with
34 the greatest increases in broadband connection rates (i.e. Greece, Romania and Bulgaria had
35 high average annual growth in the period 2007-2009).
36
37

38
39
40
41
42
43
44 **Cluster 4 (*early KEs*).** This cluster groups the country-year patterns that show a very
45 low performance in each knowledge economy pillar. Along with Chile and Mexico, this
46 cluster mainly represents transition economies in Europe and China. Thus, in relation to their
47 economic incentive regime, it must be said that they have a far from favorable framework due
48 to the transition from a soviet-type institutional environment to a capitalist one. Hence, the
49 special socioeconomic situation should explain the wide range of differences compared to the
50 other clusters. With respect to the innovation pillar, the level of patenting and publishing is
51 very low (almost nonexistent in some cases) and, finally, the number of telephone lines and
52 mobiles, as well as the number of Internet subscribers, are far from the cluster centroid.
53
54
55
56
57
58
59
60
61
62
63
64
65

3.2. Evaluated nominal and ordinal classifiers

For comparison purposes and to determine the nature of the problem, different well-known nominal and ordinal classifiers in Machine Learning literature (see Section 2 for a brief literature review) have been included in the experimentation. The general formal framework is the following: the problem consists of predicting the label y of an input vector \mathbf{x} , where $x \in X \subseteq \mathbb{R}^K$ and $y \in Y = \{C_1, C_2, \dots, C_Q\}$, i.e. \mathbf{x} is in a K -dimensional input space and y is in a label space of Q different categories. The objective is to find a classification rule or function $f: X \rightarrow Y$ to predict the labels of new patterns, given a training set of N points, $D = \{(\mathbf{x}_i, y_i), 1 \leq i \leq N\}$ (clusters, in our case). If the labels are ordered, so $C_1 \prec C_2 \prec \dots \prec C_Q$, it is an additional restriction in ordinal problems. The symbol \prec denotes the ordering relationship between the different labels. Many ordinal regression measures and algorithms consider the rank of the label, that is, the position of the label in the ordinal scale, what can be expressed by the function $\mathcal{O}(\cdot)$, in such a way that $\mathcal{O}(C_q) = q, 1 \leq q \leq Q$.

3.2.1. Nominal classifiers.

The *nominal classifiers* employed are:

- *MLogistic*: It is an algorithm for building a multinomial logistic regression model with a ridge estimator to guard against overfitting by penalizing large coefficients, based on the work by le Cessie and van Houwelingen (1992).

- *K-NN*: The K nearest neighbor rule is one of the most successful and simplest techniques used to resolve classification tasks. It consists of taking the training patterns, and assigning the class value by a majority vote of its neighbors, with the pattern being assigned to the most common class amongst its K nearest neighbors. If $K = 1$, then the pattern is simply assigned to the class of its nearest neighbor (Cover and Hart, 1967).

- *SimpleLogistic*: The SLogistic algorithm builds multinomial logistic regression models by using the LogitBoost algorithm (a boosting algorithm that performs forward stagewise fitting) that was proposed by Friedman et al. (2000) for fitting additive logistic regression models by maximum likelihood. These models are a generalization of (linear) logistic regression models. This version of the algorithm is based on controlling the number of variables in the model to avoid over-fitting.

3.2.2. Ordinal classifiers.

In the case of ordinal methods, *threshold methods* are the most common approach. Their main idea is to model ordinal regression problems from a regression perspective in the

sense that some underlying real-valued outcomes are assumed to exist, although they are unobservable. In this way, two different things are usually estimated:

- A function $f(\mathbf{x})$, which is able to predict these real-valued outcomes, uncovering the nature of the assumed underlying outcome.
- A vector $\mathbf{b} \in \mathbb{R}^{Q-1}$ (Q is the number of classes) of free parameters (thresholds) where $\mathbf{b} = (b_0, b_1, \dots, b_{Q-1}, b_Q)$, which represent intervals in the range of satisfying the constraints $b_1 \leq b_2 \leq \dots \leq b_{Q-1}$. Each interval is associated with a class, allowing possible different scales around the different classes.

All the ordinal classifiers selected for the experimentation follow this structure and they are briefly summarized now:

- *The Proportional Odd Model (POM)*. This model is based on the following general form: $g^{-1}(P(y \leq C_q | \mathbf{x})) = b_q - \mathbf{w}^T \mathbf{x}$, $1 \leq q \leq Q$, which satisfies stochastic ordering. In this way, $g^{-1}: [0, 1] \rightarrow (-\infty, +\infty)$ is a monotonic function (the inverse link function), and b_q is the threshold for class C_q . A latent variable is the main motivation of the POM, which has the linear form $f(\mathbf{x}) = \mathbf{w}^T \mathbf{x}$. Then, a probability density function over the class labels for a given feature vector \mathbf{x} is assumed. Thus, the label C_i of the training set is observed if and only if $f(\mathbf{x}) \in [b_{q-1}, b_q]$, where the function f (latent utility) and $\{b_0, b_1, \dots, b_{Q-1}, b_Q\}$ are the parameters to be learnt from the data, with $b_0 = -\infty$ and $b_Q = +\infty$. Once the model of the latent variable is defined, $f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + \varepsilon$, where ε is the random component with zero expectation, $E[\varepsilon] = 0$, and distributed according to $F\varepsilon$, the key is to select a proper distribution assumption, $F\varepsilon$, for ε . The cumulative model is obtained by choosing the inverse distribution $F\varepsilon^{-1}$ as the inverse link function g^{-1} . The most common choice for $F\varepsilon$ is the standard logistic function, where the logit is modeled in the following way:

$$\text{logit}(y \leq C_q | \mathbf{x}) = \ln(\text{odds}(y \leq C_q | \mathbf{x})) = \ln\left(\frac{P(y \leq C_q | \mathbf{x})}{1 - P(y \leq C_q | \mathbf{x})}\right) = \mathbf{w}^T \cdot \mathbf{x} + b_q. \quad (1)$$

This model form allows direct estimation of the probabilities

$$P(y = C_q | \mathbf{x}) = P(y \leq C_q | \mathbf{x}) - P(y \leq C_{q-1} | \mathbf{x}).$$

- *Support Vector Ordinal Regression (SVOR)*.

The Support Vector Machine (SVM) (Cortes and Vapnik, 1995) may be the most common kernel learning method for classification. It can be thought of a basis function model with a kernel computing the inner product on transformed input vectors $\phi(\mathbf{x})$. Here, $\phi(\mathbf{x})$ denotes the input pattern \mathbf{x} in a very high dimensional space, which is related to \mathbf{x} by a

1 specific transformation (Cortes and Vapnik, 1995). All computations are done using only the
2 reproducing kernel function, which is defined as:

$$3 \quad k(\mathbf{x}, \mathbf{x}') = \langle \phi(\mathbf{x}) \cdot \phi(\mathbf{x}') \rangle \quad (2)$$

4
5 where \cdot denotes the inner product in the high dimensional space. The idea is that linear models
6 in this space will be nonlinear in the original one.

7
8
9 SVMs are based on the idea of separating the two different classes — they are firstly
10 defined for two classes and then extended to the multiclass case — using a hyperplane
11 specified by its normal vector \mathbf{w} and the bias b , $\mathbf{w} \cdot \phi(\mathbf{x}) + b = 0$.

12
13
14
15 The optimal separating hyperplane is the one that maximizes the distance between the
16 hyperplane and the nearest points of both classes (called margin) for a bi-class problem,
17 resulting in the best prediction for unseen data. Additionally, hard margins are replaced by
18 soft margins, which allows the handling of noise and pre-labeling errors. *Slack-variables*, ξ_i ,
19 are used to relax hard-margin constraint. The optimal separating hyperplane with maximal
20 margin is a Quadratic Programming (QP) optimization problem. To deal with the multiclass
21 case, a “1-versus-1” approach is usually considered (Hsu and Lin, 2002).

22
23
24
25
26
27 SVM formulation has been adapted to the ordinal regression setting (Support Vector
28 Machines for Ordinal Regression or SVOR), by defining a different threshold b_j for each
29 class, and adapting the QP problem. Instead of deciding the class by the sign of the projection
30 $\mathbf{w}^T \cdot \mathbf{x}$, the corresponding real line is divided into consecutive intervals with a threshold vector
31 \mathbf{b} , resulting in parallel hyperplanes with the same \mathbf{w} and different thresholds b_j . In this paper,
32 two different implementations for this idea are considered (Chu and Keerthi, 2007):

- 33 – SVOR with Explicit constraints (SVOREX). The QP problem includes explicitly a
34 set of constraints assuring the order between the thresholds, while the slacks for the j -
35 th parallel hyperplane consider patterns of class j and $j + 1$.
- 36 – SVOR with Implicit constraints (SVORIM). The following principle is used: instead
37 of considering only the errors from the samples of adjacent categories, samples in all
38 the categories are allowed to contribute errors for each hyperplane. It is shown that
39 ordinal inequalities on the thresholds are implicitly satisfied at the optimal solution.

40 41 42 43 44 45 46 47 48 49 50 51 **4. EXPERIMENTAL STUDY AND RESULTS**

52 53 *4.1. Dataset description*

54
55 We use a dataset of 54 countries to be classified according to their progress toward a
56 KE and for obtaining a classifier (nominal or ordinal) to predict such a classification. Due to
57 the data availability of the selected variables (until 2009 in most official databases), 162 items
58
59
60
61
62
63
64
65

1 were selected considering each country and year for the period 2007-2009 as a single
2 composed item (country-year observation). The countries selected were all European
3 countries (except Monaco, Andorra, Serbia and Montenegro)³, adding the remaining non-
4 European members of the OECD, as well as China, Hong Kong (due to their socioeconomic
5 and political relevance in the international arena) and Tunisia (due to its Free Trade Zone
6 Agreement with the EU and its geographic proximity).
7
8
9

10 To select the variables, we considered related literature, official reports and the World
11 Bank's Knowledge Assessment Methodology (World Bank, 2012). As mentioned in Section
12 2, the knowledge economy index (KEI) of the World Bank is built as a simple average of four
13 sub-indexes, which represent the following four pillars of the knowledge economy:
14
15
16
17
18
19

20 **Pillar 1: Economic and institutional regime.** An appropriate economic and institutional
21 regime provides incentives for the efficient use of existing and new knowledge and for
22 entrepreneurship.
23
24

25 **Pillar 2: Education and skills.** One of the most important pillars of the knowledge-based
26 economy is human capital. Thus, education appears as a critical element of socioeconomic
27 life enabling people to create and share knowledge.
28
29
30

31 **Pillar 4: Information and communication infrastructure.** Knowledge economy is heavily
32 based on information and communications technology (ICT). Many global changes that have
33 taken place since the early 1970s are driven by technology. The global use of ICTs has
34 radically modified the landscape of business, government, and social life.
35
36
37

38 **Pillar 4: Innovation System.** This is an efficient innovation system of firms, research centers,
39 universities, and other organizations, where knowledge diffusion and knowledge creation are
40 typically measured by the country's ability to patent and publish scientific articles.
41
42
43

44 These variables try to track the overall performance of the economy, which "illustrate
45 how well an economy is actually using knowledge for its overall economic and social
46 development" (World Bank, 2012). Data were obtained from various sources: the World Bank
47 Database, the SCImago Journal and Country Rank (in the case of scientific publications per
48 country) and the Heritage Foundation for Economic Regime pillar variables. They were
49 adjusted to 1,000 per population when necessary and an imputation has been carried out for
50 missing data when necessary (it occurred in up to 18 cases). The variables and their
51 descriptive statistics are the following:
52
53
54
55
56
57
58

59 ³ Monaco and Andorra were not selected due to their small population; Serbia and Montenegro due to the lack of statistics
60 because of their independence in 2006.
61
62
63
64
65

Table 2. Description of variables and descriptive statistics.

CODE	VARIABLE	DESCRIPTION	Mean±SD
Pillar 1: Economic and institutional regime			
<i>TNTBA</i>	Tariff and Non-tariff Barriers	Score assigned to each country based on the analysis of its tariff and non-tariff barriers to trade, such as import bans and quotas as well as strict labeling and licensing requirements.	82.038±7.296
<i>REGQU</i>	Regulatory Quality	Incidence of market-unfriendly policies such as price controls or inadequate bank supervision, as well as perceptions of the burdens imposed by excessive regulation in areas such as foreign trade and business development.	0.943±0.734
<i>RULEL</i>	Rule of Law	This indicator includes several indicators, which measure the extent to which agents have confidence in and abide by the rules of society.	0.796±0.933
Pillar 2: Education and skills			
<i>SCHOO</i>	Primary enrollment (% gross)	Gross enrollment ratio is the ratio of total enrollment, regardless of age, to the population of the age group that officially corresponds to the primary level of education.	101.855±5.818
<i>SECON</i>	Secondary Enrollment (% gross)	The ratio of total enrollment, regardless of age, to the population of the age group that officially corresponds to secondary level of education.	97.818±11.782
<i>TERTI</i>	Tertiary Enrollment (% gross)	The ratio of total enrollment, regardless of age, to the population of the age group that officially corresponds to the tertiary level of education.	58.142±19.179
Pillar 3: Information and communication infrastructure			
<i>TELEP</i>	Telephones	Telephone mainlines, per 1.000 population. Integrated services digital network channels and fixed wireless subscribers are included.	1425.006±340,150
<i>FIXBI</i>	Fixed broadband Internet subscribers	The number of broadband subscribers with a digital subscriber line, cable modem or other high-speed technology, per 1.000 population.	175.152±106,165
<i>INTERN</i>	Internet Users	Internet users are people with access to the worldwide network, per 1.000 population.	542.012±233,420

Pillar 4: Innovation system			
<i>PATEN</i>	Patent Applications	Patent grants by country of origin and patent office, per 1.000 population.	0.782±1.083
<i>STJOU</i>	Scientific Journal Articles	Scientific and engineering articles published by country, per 1.000 population.	1.108±0.878

4.2. Classification metrics

The two evaluation metrics selected quantify the accuracy of n predicted labels for a given dataset $\{y_1^*, y_2^*, \dots, y_n^*\}$ with respect to the true targets $\{y_1, y_2, \dots, y_n\}$. These measures evaluate two different facets to be taken into account in ordinal classification problems (Chu and Keerti, 2007): if the patterns are correctly classified (*CCR*) and if the classifier predicts a class as close to the true class as possible (*MAE*).

The first one, the Correct Classification Rate or accuracy (*CCR*), is the fraction of correct predictions in individual samples:

$$C = \frac{1}{N} \sum_{i=1}^N I(y_i^* = y_i) \quad (3)$$

where $I(\cdot)$ is the zero-one loss function ; n is the number of patterns in the evaluated dataset (generalization set or year 2009 in our case). A good classifier tries to achieve the highest possible *CCR* in a given problem.

The second metric is the mean absolute error (*MAE*) that takes into account the degree of misclassification. It is the average deviation in absolute value of the prediction from the true target,

$$MAE = \frac{1}{n} \sum_{i=1}^n |\mathcal{O}(y_i^*) - \mathcal{O}(y_i)| \quad (4)$$

where $|\mathcal{O}(y_i^*) - \mathcal{O}(y_i)|$ is the distance between the true and predicted labels.

4.3. Results and discussion

The set of experiments performed in this paper includes all the methods considered in methodology Section 3. Table 3 compares the accuracy of the six methods (three nominal and three ordinal) in predicting the stage or progress toward a KE for a given country due to its classification in the corresponding cluster. A first objective was to test if ordinal classifiers actually performed better than nominal ones, (as was to be expected due to the nature of the problem), and consequently to choose the best model for predicting.

The dataset has been split into two subsequent time periods. The first period was for training the model and the second one for evaluating or testing the model. Consequently, the training set consisted of 108 country-year observations in the period 2007-2008 (annual data). The test or generalization set consisted of 54 country-year observations in the year 2009.

Table 3. CCR_G and MAE_G for the generalization set of the different methods evaluated^a

Type of classifier	Classifier	Metrics	
		CCR_G	MAE_G
Nominal	Logistic	83.33%	0.167
	1-NN	81.48%	0.185
	Slogistic	83.33%	0.185
Ordinal	POM	85.19%	0.148
	SVOR-EXC	79.63%	0.204
	SVOR-IMC	87.04%	0.130

^a The best result is in bold face and the second best result in italics.

The results for the two different evaluation measures (CCR_G and MAE_G ; see Eqs. (4) and (5)) are included in Table 3. Based on these values, the performance of each method can be analyzed. The first conclusion is that high accuracies (CCR_G) are obtained in ordinal classifiers, what shows that considering the problem as an ordinal task can provide accurate information of the stage of transition toward a KE. Namely, the SVORIM is the one that perform the best ($CCR_G = 87.04%$). MAE_G values are also low and the lowest one can also be found in the SVORIM model. A MAE_G value of 0.130 means that the classifier predictions are, in average, 0.130 categories lower or higher than the target ones) with seven misclassified country-year patterns in 2009, which is the corresponding year of the generalization set.

The ordinal classifier SVORIM is chosen because it is more accurate. However, although one could say that the differences are not excessively great with respect to nominal ones, it would be interesting to take additional considerations into account. Firstly, in real situations like the one dealt with in this study, stakeholders and policy-makers need ordered classes both to monitor specific policies and strategies and to carry out benchmarking practices. Thus, the identification of ordered clusters might help the decision maker to identify profiles or patterns to be used later on in ordinal problems for monitoring,

1 benchmarking and offering a new perspective in a world where more and more rankings are
 2 coming into play (De Smet et al., 2012).

3 Other arguments in favor of the selection of the ordinal classifiers are:

4
 5 1. By doing an ordinal classification, we project observations/patterns into one
 6 dimension, while discriminant analysis has multiple projections. To interpret the model and
 7 study the projections, it is much better and more natural to use a single projection to order the
 8 patterns.
 9

10
 11 2. The correct classification rate (*CCR*) measures the proportion of correctly classified
 12 objects on a sample of data. For the ordinal problem this might not always be the most
 13 appropriate performance criterion. The *CCR* tacitly assumes equal misclassification costs for
 14 all incorrectly classified objects. The ordinal nature of the problem however implies that a one
 15 adjacent cluster misclassification is not equally costly that two adjacent clusters error in the
 16 classification. For example, the correct classification rate for cluster 1 is 92% (23 patterns
 17 correctly classified out of 25; see Table 4), being the cluster that groups the highest number of
 18 patterns (46.29% of the patterns in the generalization set), and only two patterns are not
 19 correctly classified in the adjacent cluster 2.
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30

31 Table 4. Confusion Matrix for the best ordinal classifier (SVORIM)

#patterns	Actual class	Predicted class (SVORIMC) ^a			
		Cluster 1	Cluster 2	Cluster 3	Cluster 4
25	Cluster 1	23	2	0	0
10	Cluster 2	0	10	0	0
7	Cluster 3	0	1	4	2
12	Cluster 4	0	0	2	10

32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44 ^a The seven misclassified country-year observations are: Belarus-09, Bosnia-09,
 45 Chile-09, Latvia-09, Poland-09, Russia-09 and Slovenia-09.
 46
 47
 48

49 The selected SVORIM model correctly classified most of the patterns (87.04%). On
 50 analyzing the errors committed in the classification test set, they correspond to the following
 51 seven countries in 2009: Belarus, Bosnia, Chile, Latvia, Poland, Russia and Slovenia. Latvia,
 52 Slovenia and Poland have been misclassified in both nominal and ordinal methods. Belarus,
 53 Chile, Latvia and Slovenia are classified into a lower category than in our cluster analysis,
 54 while Bosnia, Russia and Poland were classified in a higher category.
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

These differences may be due to the fact that the quantitative data employed in our model provide information on the historical data of the pillars selected and on their fundamental structural features. They are basically backward-looking, while in social sciences and at macro-level, there could be unexpected situations; a model is not expected to embody all possibilities of reality. Examination of future events by experts has to be a supplementary tool to test the vulnerability of a country to a variety of shocks generated both internally and externally. Thus, qualitative and judgmental aspects of analysis are unavoidable, even in the interpretation of quantitative indicators.

Continuing in this vein, these misclassifications could be explained with the help of the literature and experts' opinions. Thus, the first reason could lie in the fact that five out of the seven countries are post-socialist countries and three are new member states (NMS). Only fairly little importance is given to regions and regional policy in most of the new member states. After EU accession, it seems more difficult for EU NMS to build knowledge-based economies; the gaps in knowledge and innovation-related activities will widen, partly because of their weak technological capability (Kołodko, 2001; Orłowski, 2001). Poland is a special case in that its transition to capitalism was facilitated by the fact that a small number of private firms had been tolerated throughout the communist reign (Boettke et al., 2008).

In the case of Latvia, it moves from cluster 2 to cluster 1 in 2009 but classification models still consider it to belong to cluster 2. The analysis of variables shows the amazing increase in the performance of the ICT pillar along with the high gross rate of enrollment in tertiary education, which led to its classification in cluster 1. Nevertheless, a significant decrease in the number of schools is observed in previous years due to the demographic situation. This characteristic does not correspond to a country with a high educational level and a strong increase in the penetration of ICT in 2009. This is a special situation that our model considers atypical and not decisive enough for a cluster change.

The same case is observed for Slovenia: it moves from cluster 2 to cluster 1 but the best classifier still considers it to belong to the second cluster. It shows a good level of patent applications and scientific publications as well as an overall increase in the rate of enrollment at all educational levels; it has also continuously improved its ICT penetration rate. This is the reason for the change from cluster 2 to cluster 1. However, this increase has not been enough to consider it as a first cluster member for our model.

Russia belongs to cluster 4 in all the years under study, but the selected ordinal classifiers predicted that Russia would belong to cluster 3 in 2009. This could also be due to its special transition from a communist to a capitalist economy and its regulatory framework.

5. CONCLUSIONS

1 Knowledge economy has become the major global trend in international society in the
2 globally competitive twenty-first century. Exploring the characteristics of knowledge
3 economy and establishing an appropriate economic paradigm to accelerate technological
4 innovation are priority tasks for governments. In this context, main stakeholders need useful
5 tools to support decision-making, like the number of composite indicators provided by
6 institutions like the OECD, the World Bank or the EU.
7
8
9
10
11

12 This study examined three nominal and three ordinal classification methods from the
13 machine-learning field with a two-fold aim: i) to obtain a model to predict the classification of
14 countries according to their stage of transition to a Knowledge Economy (KE) as a monitoring
15 and benchmarking tool for policy makers and stakeholders (previously, a hierarchical
16 clustering was performed to obtain the clusters in which countries could be grouped according
17 to the variables that are usually employed to characterize a KE) ; ii) to test our hypothesis
18 about the ordinal nature of this economic issue.
19
20
21
22
23
24

25 To achieve this aim, a dataset of 162 patterns formed by 54 countries was employed
26 (European Countries, non-European OECD members plus China and Hong Kong as relevant
27 economies) in the period 2007-2009 and four clusters were obtained: advanced, followers,
28 moderate and early Knowledge Economies.
29
30
31
32

33 Once the clusters were obtained, we compared the accuracy of the proposed nominal
34 and ordinal methods from machine learning. The results of this preliminary research show
35 that ordinal classifiers perform better than nominal ones and that the best performing method
36 to predict the classification of countries was the SVORIM, with high accuracy (*CCR*) and low
37 *MAE* values. Thus, this work has also confirmed how the ordering of clusters can improve the
38 performance of nominal classifiers, yielding more accurate predictions and supporting the
39 initial assumption of the ordinal nature of clusters defined by hierarchical clustering.
40 Complementary economic and methodological reasons were provided to support this claim.
41
42
43
44
45
46

47 Monitoring is a relevant task for shedding light on the progress of a knowledge
48 economy. The classification of these countries according to their advances to a KE (based on
49 the four pillars) can also be useful to assess their position as compared to other countries with
50 respect to reaching competitiveness and sustainability objectives in the long-term. The
51 clusters obtained and the ordinal classification model selected reflect a global picture of the
52 KE stage of countries, which could enrich and complement the judgment of stakeholders
53 more than a single indicator score value or trying to find the KE readiness of a country
54 through separate indicators.
55
56
57
58
59
60
61
62
63
64
65

1
2 Therefore, the methodology presented can be useful as: i) a managerial tool for
3 supporting decision making since it allows a stage in the transition to a KE to be predicted; ii)
4 a benchmarking tool to compare results with other countries of reference and to learn from
5 their best practices. The classifier could help experts and the main stakeholders to predict the
6 classification of countries according to their stage of transition to a KE in future years or other
7 countries, in a similar way to scores, rankings, composite indexes and other methodologies
8 with publicly available country-data. It could even simulate scenarios that could help to
9 predict the classification of the country in a case where one or another variable increases or
10 decreases.
11

12
13 As appropriate data become available, the implications of our approach would appear
14 to be worthwhile for future comparative-research in cross-country performance. Additionally,
15 it would be of interest to amplify research with data since the year 2009 because the period
16 analyzed in the present study may not fully capture the impact of the economic and financial
17 crisis due to the delay in economic impacts and in data availability.
18 Finally, the choice of variables selected could bias both the results in clustering and in
19 classification methods, as they are proxy measures. The selection of alternative, additional
20 and/or a combination of variables for each pillar is proposed for further research.
21
22
23
24
25
26
27
28
29
30

31 32 33 **Acknowledgment**

34 This work was partially subsidized by the Spanish Inter-Ministerial Commission of
35 Science and Technology under Project TIN2011-22794, the European Regional Development
36 fund, and the “Junta de Andalucía”, Spain, under Project P08-TIC-374.
37
38
39
40

41 42 **References**

- 43 Al Shami, A., Lotfi, A. & Coleman, S. (2012). Unified Knowledge Based Economy Neural
44 Forecasting Map. In WCCI 2012 IEEE World Congress on Computational Intelligence
45 (IJCNN) (pp. 2008–2015). Brisbane, Australia.
46
47 Abramovitz, M. (1986). Catching up, forging ahead, and falling behind. *The Journal of*
48 *Economic History*, 46, 385–406.
49
50 Aghion, P. & Howitt, P. (1992). A model of growth through creative destruction.
51 *Econometrica*, 60, 323–351.
52
53 Boettke, P.J., Coyne, C.J. & Leeson, P.T. (2008). Institutional Stickiness and the New
54 Development Economics. *American Journal of Economics and Sociology*, 67, 1536–7150.
55
56
57
58
59
60
61
62
63
64
65

- 1 Booyesen, F. (2002). An Overview and Evaluation of Composite Indices of Development.
2 *Social Indicators Research*, 59, 115–151.
- 3 Buck, S.F. (1960). A method of estimation of missing values in multivariate data suitable for
4 use with an electronic computer. *Journal of the Royal Statistical Society Series B-Statistical*
5 *Methodology*, 22, 302–306.
- 6
7
8
9 Chen Derek, H.C. & Dahlman C. J. (2006). The Knowledge Economy, The Kam
10 Methodology and World Bank Operations. The World Bank, 1-24.
- 11
12 Chen, Y.-L., Hsu, W.-H. & Lee, Y-H. (2006). TASC: Two-attribute-set clustering through
13 decision tree construction. *European Journal of Operational Research*, 174, 930-944.
- 14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
- Chu, W. & Keerthi, S. S. (2007). Support vector ordinal regression. *Neural Computing*, 19,
792–815.
- Corne, D., Dhaenens, C., & Jourdan, L. (2012). Synergies between operations research and
data mining: The emerging use of multi-objective approaches. *European Journal of*
Operational Research, 221, 469-479.
- Cortes, C. & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20, 273–297.
- Cover T.M. & Hart P.E. (1967). Nearest neighbor pattern classification. *IEEE Transactions*
on Information Theory, 13, 21–27
- David, P. & Foray, D. (2002). An introduction to the economy of the knowledge society.
International Social Science Journal, 54, 9–24.
- De Smet Y., Nemery, P. & Selvaraj, R. (2012). An exact algorithm for the multicriteria
ordered clustering problem. *Omega*, 40, 861–869.
- Doumpos, M., Kosmidou, K., Baourakis, G. & Zopounidis, C. (2002). Credit risk assessment
using a multicriteria hierarchical discrimination approach: A comparative analysis. *European*
Journal of Operational Research, 138, 392–412.
- Drucker, P. (1969). *The Age of Discontinuity*. New York: Harper and Row.
- Drucker, P. (1993). *Postcapitalist Society*. New York: HarperBusiness.
- 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,
109–23.
- Fouad, S. & Tino, P. (2012). Adaptive Metric Learning Vector Quantization for Ordinal
Classification. *Neuronal Computing*, 24, 2825–2851.
- Frank, E. & Hall, M. (2001). A simple approach to ordinal classification. In L. de Raedt and
Flach (Eds.), *Proceedings of the Twelfth European Conference on Machine Learning* (pp.
145–156). New York: Springer-Verlag.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Freeman, C. (1987). *Technology Policy and Economic Performance: Lessons from Japan*. London: Frances Pinter.

Friedman, J., Hastie, T. & Tibshirani, R. (2000). Additive logistic regression: a statistical view of boosting. *Annals of Statistics*, 38, 337–374.

Fullerton, A. & Xu, J. (2012). The proportional odds with partial proportionality constraints model for ordinal response variables. *Social Science Research*, 41, 182–198.

Grossman, G.M. & Helpman, E. (1991). *Innovation and growth in the global economy*, Cambridge, MA: The MIT Press.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P. & Witten, I. (2009). The WEKA Data Mining Software: An update. *ACM SIGKDD Explorations News*, 11, 10–18.

Hanley, S. & Malafsky, G. (2002). A Guide for Measuring the Value of Km Investments. In C.W. Holsapple (Eds.), *Handbook on Knowledge Management 2* (pp. 369–394). Berlin, Heidelberg: Springer-Verlag.

Herbrich, R., Graepel, T. & Obermayer, K. (1999). Support vector learning for ordinal regression. In *Ninth international conference on artificial neural networks, ICANN'99* (pp. 97–102). Edinburgh, UK.

Hsu, C.W. & Lin, C.J. (2002). A comparison of methods for multi-class support vector machines. *IEEE Transaction on Neural Networks*, 13, 415–425.

Hsu, G.J.Y., Lin, Y.-H. & Wei, Z.-Y. (2008). Competition policy for technological innovation in an era of knowledge-based economy. *Knowledge Based Systems*, 21, 826–832.

International Bank for Reconstruction and Development, (2007). *Building Knowledge Economies: Advanced Strategies for Development*. World Bank Institute, Knowledge for Development Program, Available at: <http://go.worldbank.org/851HK6EUH0> (last accessed on 28 October 28 2012).

Johannessen, J.A. & Olsen, B. (2010). The future of value creation and innovations: Aspects of a theory of value creation and innovation in a global knowledge economy. *International Journal of Information Management*, 30, 502–511.

Kolodko, G.W. (2001). The 'New Economy' and the Old Problems. Prospects for Fast Growth in Postsocialist Countries. *Zagreb International Review of Economics & Business*, 2, 71–90.

Kotsiantis, S.B. & Pintelas, P.E. (2004). A Cost Sensitive Technique for Ordinal Classification Problems. In Vouros, G.A. & Panayiotopoulos, T. (Eds.), *SETN 2004. LNCS (LNAI)* (pp. 220–229). Berlin, Heidelberg: Springer-Verlag.

Le Cessie, S. & van Houwelingen, J. (1992). Ridge estimators in logistic regression. *Journal of The Royal Statistical Society, Series C-Applied Statistics*, 41, 191–201.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Leydesdorff, L. (2006). *The Knowledge-based Economy: Modeled, Measured, Simulated*. Boca Raton: Universal Publishers.

Lin, H.-Y. (2013). Feature selection based on cluster and variability analyses for ordinal multi-class classification problems. *Knowledge Based Systems*, 37, 94–104.

Lucas, R.E. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22, 3–42.

Machlup, F. (1962). *The Production and Distribution of Knowledge in the United States*. Princeton: Princeton University Press.

Mathieson, M. J. (1995). Ordinal models for neural networks. In A. Refenes, Y. Abu-Mostafa, Moody, J. (Eds.), *Neural networks in Financial Engineering* (pp. 523–536). Singapore: World Scientific.

McCullagh, P. (1980). Regression models for ordinal data. *Journal of Royal Statistical Society Series B- Methodological*, 42, 109–142.

Meyer, P. & Olteanu, A.-L., Formalizing and solving the problem of clustering in MCDA, *European Journal of Operational Research*, in press, <http://dx.doi.org/10.1016/j.ejor.2013.01.016>.

Nardo, M. et al. (2005). *Handbook on Constructing Composite Indicators: Methodology and User Guide*. OECD Statistics Working Papers. Paris: OECD Publishing.

Nelson, R.R. & Winter, S.G. (1982). *An Evolutionary Theory of Economic Change*. Cambridge, MA: The Belknap Press of Harvard University Press.

OECD, (1996). *The knowledge Economy*. Paris: OECD.

Orłowski, W. (2000). Knowledge Economy and Knowledge-Based Growth: Some Issues in a Transition Economy. In A. Kukliński (Eds.), *The Knowledge-Based Economy: The European Challenges of the 21st Century* (pp. 89–96). Warsaw: State Committee for Scientific Research of the Republic of Poland.

Powell, W.W. & Snellman, K. (2004). The Knowledge Economy. *Annual Review of Sociology*, 30, 199–220.

Pyon, Y.S. & Li, J. (2009). Identifying gene signatures from cancer progression data using ordinal analysis. In *BIBM '09 Proceedings of the 2009 IEEE International Conference on Bioinformatics and Biomedicine* (pp. 136–141). Washington, DC: IEEE Computer Society.

Questier, F., Walczaka, B., Massarta, D.L., Bouconb, C. & de Jong, S. (2002). Feature selection for hierarchical clustering. *Analytica Chimica Acta*, 466, 311–324.

Romer, P.M. (1990). Endogenous technical change. *Journal of Political Economy*, 98, 71–102.

1 Shashua, A. & Levin, A., (2003). Ranking with Large Margin Principle: Two Approaches.
2 *Advances in Neural Information Processing Systems*, 15, 937–944.

3 Sousa, R., Yevseyeva, I., Pinto da Costa, J.F. & Cardoso, J.S., (2013). Multicriteria Models
4 for Learning Ordinal Data: A Literature Review. In X.-S. Yang (Eds.). *Artificial Intelligence,
5 Evolutionary Computing and Metaheuristics, Studies in Computational Intelligence* (pp. 109–
6 138). Berlin, Heidelberg: Springer-Verlag.

7
8
9
10
11 Thurow, L. (1999). *Building wealth: The new rules for individuals, companies, and nations in
12 a knowledge-based economy*. New York: HarperBusiness.

13
14 Von Krogh, G., Ichijo, K. & Nonaka, I. (2000). *Enabling knowledge creation: How to unlock
15 the mystery of tacit knowledge and release the power of innovation*. New York: Oxford
16
17 University Press.

18
19
20 World Bank (2012). Knowledge for Development (K4D), Knowledge Assessment
21 Methodology 2012 (last access: November, 2012; www.worldbank.org/kam).

22
23 Wu, J., Xiong, H. & Chen, J. (2009). Towards understanding hierarchical clustering: A data
24 distribution perspective. *Neurocomputing*, 72, 2319–2330.

25
26
27 Xu, X., Zhou, C. & Wang, Z. (2009). Credit scoring algorithm based on link analysis ranking
28 with support vector machine. *Expert Systems with Applications*, 36, 2625–2632.

29
30 Zopounidis, C. & Doumpos, M. (2002). Multicriteria classification and sorting methods: A
31 literature review. *European Journal of Operational Research*, 138, 229–246.
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

Table 1. Description of clusters (mean, standard deviation and country-year pattern of reference).

CODE	VARIABLE	Mean±SD (country-year pattern of reference for each cluster and variable)			
		Cluster 1 ^(a)	Cluster 2 ^(b)	Cluster 3 ^(c)	Cluster 4 ^(d)
Pillar 1: Economic and institutional regime					
<i>TNTBA</i>	Tariff and Non-tariff Barriers	85.149±4.575 (Japan-09)	85.400±1.975 (Italy-07)	83.324±5.271 (Cyprus-07)	75.638±9.605 (Chile-08)
<i>REGQU</i>	Regulatory Quality	1.523±0.291 (Korea Rep.-08)	1.109±0.262 (Croatia-09)	0.721±0.550 (Maced. FYR-08)	0.009±0.539 (Tunisia-09)
<i>RULEL</i>	Rule of Law	1.640±0.298 (Latvia-09)	0.912±0.383 (Hungary-09)	0.300±0.568 (Greece-07)	-0.394±0.502 (Chile-07)
Pillar 2: Education and skills					
<i>SCHOO</i>	Primary enrollment (% gross)	102.379±3.976 (Germany-09)	103.585±6.220 (Hungary-09)	99.155±4.520 (Bulgaria-07)	103.323±7.803 (Romania-07)
<i>SECON</i>	Secondary Enrollment (% gross)	105.982±11.667 (United Kingdom-07)	100.271±8.094 (Lithuania-09)	92.691±5.255 (Cyprus-07)	87.525±5.333 (Belarus-07)
<i>TERTI</i>	Tertiary Enrollment (% gross)	66.122±17.479 (Japan-07)	64.153±12.894 (Czech Rep.-08)	56.251±13.201 (Chile-09)	42.699±19.073 (Romania-07)
Pillar 3: Information and communication infrastructure					
<i>TELEP</i>	Telephones	1,619.142±239.961 (Korea Rep.-07)	1,587.687±160.324 (Hungary-07)	1,387.418±207.976 (Belarus-09)	1,035.833±288.396 (Ukraine-07)
<i>FIXBI</i>	Fixed broadband Internet subscribers	280.205±49.153 (Estonia-07)	177.438±29.787 (Latvia-08)	107.771±16.798 (Cyprus-08)	41.149±28.340 (Russian Fed.-09)
<i>INTERN</i>	Internet Users	768.438±82.969 (Malta-09)	534.872±73.289 (Spain-07)	457.583±114.678 (Poland-08)	235.764±94.148 (Albania-09)
Pillar 4: Innovation system					
<i>PATEN</i>	Patent Applications	1.630±1.224 (United Kingdom-09)	0.372±0.413 (Ireland-07)	0.099±0.093 (Cyprus-08)	0.075±0.079 (Belarus-08)
<i>STJOU</i>	Scientific and Technical Journal Articles	1.864±0.706 (Estonia-08)	1.104±0.509 (Czech Rep.-08)	0.591±0.308 (Greece-07)	0.169±0.100 (Turkey-09)

^(a) Cluster 1 groups Australia, Austria, Belgium, Canada, Denmark, Estonia, Finland, France, Germany, Hong Kong SAR China, Iceland, Ireland-09, Japan, Korea Rep., Latvia-09, Luxembourg, Malta-09, Netherlands, New Zealand, Norway, Slovenia-09, Sweden, United Kingdom and the United States.

^(b) Cluster 2 groups Croatia-09, Cyprus-09, Czech Republic, Greece-09, Hungary, Ireland-07-08, Israel, Latvia-07-08, Lithuania, Malta-07-08, Portugal, Slovenia-07-08 and Spain.

^(c) Cluster 3 groups Belarus-09, Bulgaria, Chile-09, Croatia-07-08, Cyprus-07-08, Greece-07-08, Macedonia FYR-08-09, Poland, Romania-08-09 and the Slovak Republic.

^(d) Cluster 4 groups Albania, Armenia, Azerbaijan, Belarus-07-08, Bosnia and Herzegovina, Chile-07-08, China, Georgia, Macedonia FYR-07, Mexico, Moldova, Romania-07, Russian Federation, Tunisia, Turkey and Ukraine.

Table 2. Description of variables and descriptive statistics.

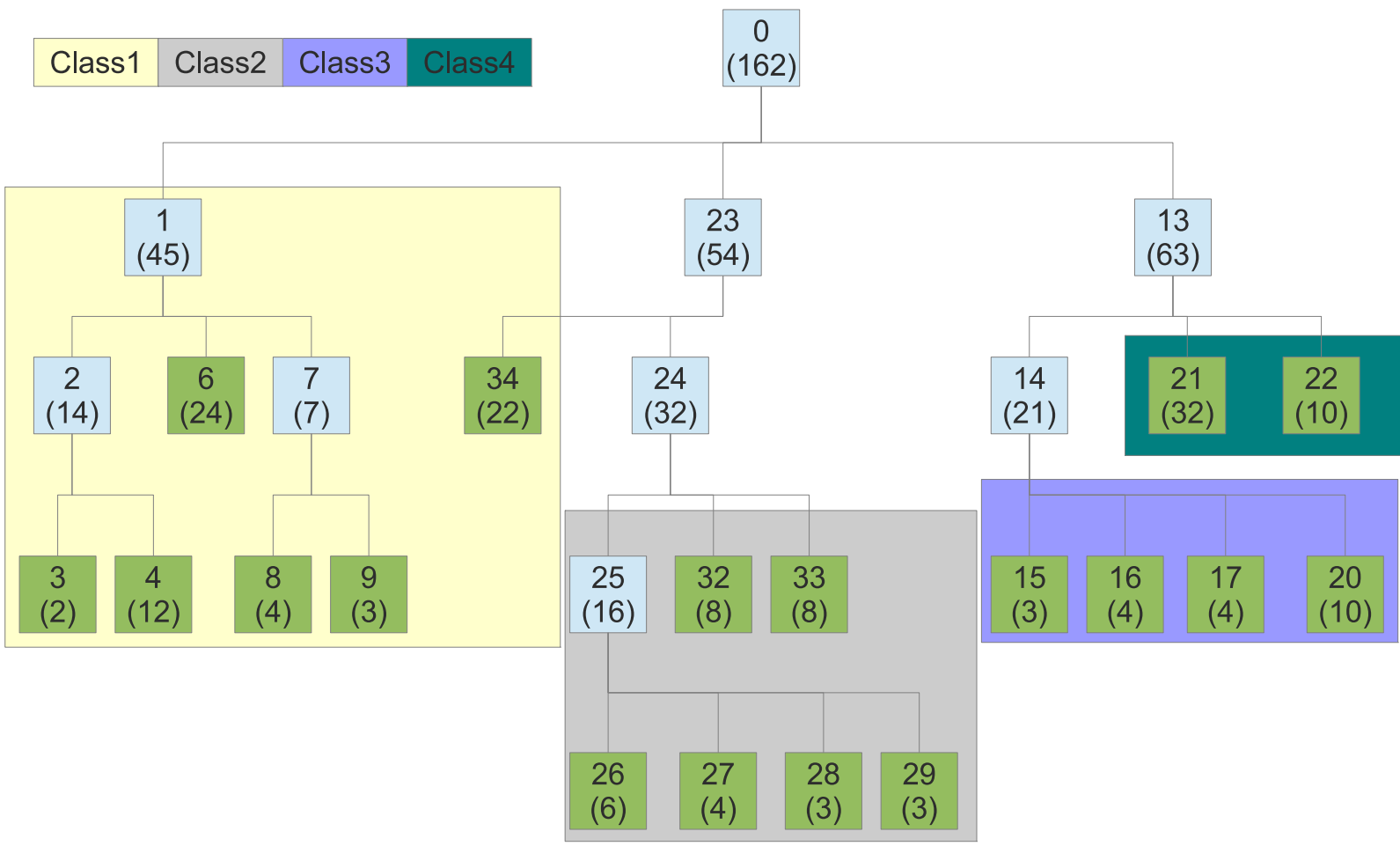
CODE	VARIABLE	DESCRIPTION	Mean±SD
Pillar 1: Economic and institutional regime			
<i>TNTBA</i>	Tariff and Non-tariff Barriers	Score assigned to each country based on the analysis of its tariff and non-tariff barriers to trade, such as import bans and quotas as well as strict labeling and licensing requirements.	82.038±7.296
<i>REGQU</i>	Regulatory Quality	Incidence of market-unfriendly policies such as price controls or inadequate bank supervision, as well as perceptions of the burdens imposed by excessive regulation in areas such as foreign trade and business development.	0.943±0.734
<i>RULEL</i>	Rule of Law	This indicator includes several indicators, which measure the extent to which agents have confidence in and abide by the rules of society.	0.796±0.933
Pillar 2: Education and skills			
<i>SCHOO</i>	Primary enrollment (% gross)	Gross enrollment ratio is the ratio of total enrollment, regardless of age, to the population of the age group that officially corresponds to the primary level of education.	101.855±5.818
<i>SECON</i>	Secondary Enrollment (% gross)	The ratio of total enrollment, regardless of age, to the population of the age group that officially corresponds to secondary level of education.	97.818±11.782
<i>TERTI</i>	Tertiary Enrollment (% gross)	The ratio of total enrollment, regardless of age, to the population of the age group that officially corresponds to the tertiary level of education.	58.142±19.179
Pillar 3: Information and communication infrastructure			
<i>TELEP</i>	Telephones	Telephone mainlines, per 1.000 population. Integrated services digital network channels and fixed wireless subscribers are included.	1425.006±340,150
<i>FIXBI</i>	Fixed broadband Internet subscribers	The number of broadband subscribers with a digital subscriber line, cable modem or other high-speed technology, per 1.000 population.	175.152±106,165
<i>INTERN</i>	Internet Users	Internet users are people with access to the worldwide network, per 1.000 population.	542.012±233,420
Pillar 4: Innovation system			
<i>PATEN</i>	Patent Applications	Patent grants by country of origin and patent office, per 1.000 population.	0.782±1.083
<i>STJOU</i>	Scientific Journal Articles	Scientific and engineering articles published by country, per 1.000 population.	1.108±0.878

Table 3. CCR_G and MAE_G for the generalization set of the different methods evaluated^a

Type of classifier	Classifier	Metrics	
		CCR_G	MAE_G
Nominal	Logistic	83.33%	0.167
	1-NN	81.48%	0.185
	Slogistic	83.33%	0.185
Ordinal	POM	<i>85.19%</i>	<i>0.148</i>
	SVOR-EXC	79.63%	0.204
	SVOR-IMC	87.04%	0.130

^a The best result is in bold face and the second best result in italics

Figure(s)



Elsevier Editorial System(tm) for Research Policy
Manuscript Draft

Manuscript Number: RESPOL-D-12-00697

Title: Evaluation of Research, Technological Development and Innovation public policies in the European Union: a Bayesian network model and the Two-Linked-Steps Data Envelopment Analysis approach.

Article Type: Research Article

Keywords: RTDI public policy; EU Strategies; technical efficiency; Bayesian network; sustainable development; two-linked-step DEA

Corresponding Author: Mrs Mónica de la Paz-Marín,

Corresponding Author's Institution: University of Córdoba

First Author: Mónica de la Paz-Marín

Order of Authors: Mónica de la Paz-Marín; Carlos García-Alonso, PhD

Abstract: The purpose of this work is to design a conceptual model through a Bayesian network (BN) to provide a useful tool for decision-making support in the evaluation of the relative technical efficiency of Research, Technological Development and Innovation (RTDI) public policies within a sustainable development framework. The efficiency of such policies in the 27 European Union (EU) Member States was assessed applying the two-linked-step Data Envelopment Analysis (2LS-DEA) approach. Providing stakeholders with this conceptual model could reduce the inherent uncertainty in both RTDI and public decision-making processes. Thus, the main contributions of this study are: i) the presentation of a Bayesian Network (BN) as a conceptual model for decision-making support; ii) the evaluation of the model through the 2LS-DEA approach, which assesses the relative technical efficiency of the RTDI public policies of the 27 EU Member State in various socioeconomic scenarios embodied in the BN (the original dataset was previously transformed by a rule-base that emerged from the BN) and iii) to obtain final efficiency scores of the RTDI public policies in the EU. The results suggest that the BN along with the 2LS-DEA approach could be a useful methodology for decision-support in monitoring the efficiency of EU RTDI public policies as well as for future better resource allocation.

Cover Letter

Dear Editor-in-Chief of Research Policy Journal,

On behalf of my co-authors, I am pleased to submit the enclosed material for possible publication in your journal. It has not been submitted for other publication nor has it been published in whole or in part elsewhere by any of the authors. I can attest to the fact that all authors listed on the title page have read the manuscript, and vouch for the validity and legitimacy of the data and its interpretation, and agree to its submission.

Mónica de la Paz Marín

October 22th, 2012

DETAIL OF SUBMITTED ARTICLE

Manuscript title:	Evaluation of Research, Technological Development and Innovation public policies in the European Union: a Bayesian network model and the Two-Linked-Steps Data Envelopment Analysis approach.
Corresponding author:	Mónica de la Paz-Marín
Co-authors:	Carlos García-Alonso
Type of manuscript:	Research Article.
Appropriate for this journal:	This article deals with the evaluation of RTDI public policies showing practical implications and support for policy makers at country level in the European Union. Through a Two-Linked-Steps Data Envelopment Analysis (2LS-DEA) the design of a Bayesian Network as a conceptual model was evaluated and final relative technical efficiency of such policies obtained. The raw data were transformed with a rule base that comes from de BN relationships in a way the BN was iteratively improved as implicit knowledge emerged. In this manner, this tool aims to be useful for stakeholders and policy makers in order to benchmark efficient RTDI public policies and for having a tested model (the BN) to performance future resource allocation.
Contact information (corresponding author)	Mónica de la Paz-Marín Department of Management and Quantitative Methods ETEA, Business Administration Faculty Escritor Castilla Aguayo 4, 14004-Córdoba, Spain Telephone number: +34.957.042226 Fax: +34.957.218030 e-mail: mpaz@uco.es
Possible conflicts of interest:	To the date, there are no conflicts of interest.
Detail of Financial Support	This work was partially subsidised by the Spanish Inter-Ministerial Commission of Science and Technology under Project TIN2011-22794, the European Regional Development fund, and the “Junta de Andalucía” (Spain), under Project P2011-TIC-7508.

Highlights

A Bayesian network (BN) as a conceptual model for evaluating RTDI EU policies is designed.>The Two-Linked-Steps DEA (2LS-DEA) is employed for the evaluation of its design.> Analyzing the results from 2LS-DEA, the BN is improved iteratively.>After evaluating the BN design, final RTDI policies efficiency scores are obtained. >The BN could be used as a useful tool for decision-making support.

Evaluation of Research, Technological Development and Innovation public policies in the European Union: a Bayesian network model and the Two-Linked-Steps Data Envelopment Analysis approach.

Mónica de la Paz-Marín^{1*}, Carlos García-Alonso²

1,2 Department of Management and Quantitative Methods. Loyola-Andalusia University, Business Administration Faculty, C/ Escritor Castilla Aguayo 4, 14004, Córdoba (Spain). Telephone number: +34.957.042226; Fax: +34.957.218030 {mpaz@uco.es*;cgarcia@etea.com}

* Corresponding author

Abstract

The purpose of this work is to design a conceptual model through a Bayesian network (BN) to provide a useful tool for decision-making support in the evaluation of the relative technical efficiency of Research, Technological Development and Innovation (RTDI) public policies within a sustainable development framework. The efficiency of such policies in the 27 European Union (EU) Member States was assessed applying the two-linked-steps Data Envelopment Analysis (2LS-DEA) approach. Providing stakeholders with this conceptual model could reduce the inherent uncertainty in both RTDI and public decision-making processes. Thus, the main contributions of this study are: i) the presentation of a Bayesian Network (BN) as a conceptual model for decision-making support; ii) the evaluation of the model through the 2LS-DEA approach, which assesses the relative technical efficiency of the RTDI public policies of the 27 EU Member State in various socioeconomic scenarios embodied in the BN (the original dataset was previously transformed by a rule-base that emerged from the BN) and iii) to obtain final efficiency scores of the RTDI public policies in the EU. The results suggest that the BN along with the 2LS-DEA approach could be a useful methodology for decision-support in monitoring the efficiency of EU RTDI public policies as well as for future better resource allocation.

Keywords: *RTDI public policies; EU Strategies; technical efficiency; Bayesian network; sustainable development; two-linked-steps DEA.*

1. Introduction

The relevance of efficiency assessment in public policies like those of Research, Technological Development and Innovation (RTDI) is highlighted in many political agendas and official reports, i.e. the Budget proposed by the U.S.A. government for 2013 or the European Union (EU) 2020 Strategy (European Commission, 2010). This evaluation and its potential socio-economic impacts are even more relevant due to the worldwide economic crisis, which implies budget cuts.

However, even in times of economic downturn, it is fundamental to continue supporting Research and Development (R&D) as well as innovation activities¹ (Filippettia and Archibugi, 2011). Numerous studies and economic growth theories state that R&D –including technological change, knowledge diffusion and innovation capacity- play a crucial role in the development, economic growth and well-being of a country (Aghion and Howitt, 1992; Furman et al., 2002; Grossman and Helpman, 1991; Romer, 1990) and have positive impacts in competitiveness (Freeman, 2004; Schumpeter, 1934; Solow, 1956; Fagerberg et al., 2004). From an empirical point of view, many authors also highlight the relevance of both the direct and socioeconomic returns of R&D expenditure (Archibal and Pereira, 2003; Baumol, 2003; Jones and Williams, 1998).

According to the above statements, the main objectives of European RTDI public policies are to increase economic growth through sustainable development and to create a knowledge-based economy in the EU by supporting R&D. In addition, the EU still faces the challenge of reducing its existing technological gap with other advanced economies (the well-known “European Paradox”). This means that, in spite of the amount of investment in R&D, Europe is weak in transforming new knowledge into innovation (Dosi et al. 2009). EU strategies like the Lisbon Strategy in 2000, re-launched in Barcelona’s European Council in 2002, and the current “European 2020” Strategy (European Commission, 2010) reflect this trend.

Under this general framework, the EU has established R&D expenditure intensity targets relying on their impact on economic performance. Sustainable development also appears as an overarching objective of the EU since 1997. The EU Strategy for Sustainable Development aims “for the continuous improvement of quality of life for current and future generations” (European Commission, 2009). R&D and innovation activities also appear as being fundamental to achieve this objective.

¹ It is important to stress that in this study the term “innovation” is understood in a narrow sense that does not include R&D activities because, while R&D plays a vital role in the innovation process, much innovation activity is not R&D-based. Innovation policies can include others like patent regulation, R&D support, etc.

R&D investments are very different from other investments. These differences are (Martin and Scott, 2000): i) indivisibility; ii) economies of scale; iii) its inherent uncertainty (it is not sure that R&D efforts will result in new knowledge and innovation); iv) asymmetric information; v) market failures; and, finally, vi) limited appropriateness of the results. This study is focused on the public side of R&D expenditure to evaluate the efficiency of such expenditure. At this point, it must be stressed that, although R&D public funds are mainly spent on basic research in universities and research institutes, this public support also has a relevant impact on R&D private expenditure (Cohen et al., 2002; Salter and Martin, 2001).

The main reason for assessing the impacts of RDTI public policies is to evaluate their contribution to general public objectives: economic growth, health-care and/or education, thus avoiding the draining of scarce resources dedicated to ineffective and inefficient programs. This evaluation allows the creation of a better-informed society that increasingly demands enhanced public accountability. This issue is still a source of great debate because there are different instruments and indicators to carry it out, and there have been as well different attempts to provide classification schemes as well as for exploring the character of innovation studies as a scientific field (Vedung, 1998; Fagerberg and Verspagen, 2009; Martin, 2012).

The methodologies for assessing public R&D impacts go from econometric studies, general Equilibrium Models (which rely on a large number of assumptions) and the use of composite indicators (Grupp and Schubert, 2010; Freeman and Soete, 2009) to survey-based studies, case studies and, more recently, non-linear classifiers for R&D performance (De la Paz-Marín et al., 2012). Nevertheless, few studies consider public R&D investment from a macroeconomic perspective (David and Hall, 2000). According to the latter, the design of a conceptual model for identifying the key variables involved in R&D and innovation processes and their relationships can be a useful decision-support tool for policy-makers and analysts.

The drawback for designing this conceptual model is the existence of multiple variables and the study of their interactions in an uncertain socio-economic environment. Moreover, classical approaches to modeling could find it quite difficult to capture the whole structure of the domain (Gibert et al., 2010), particularly when information is incomplete. Classic algebraic formalisms disregard the knowledge background necessary to design and develop a conceptual model that includes both quantitative and qualitative aspects of causal socio-economic phenomena (Pearl, 2000).

There are a number of methodologies to design conceptual models like system analysis, causal modeling, etc. In social sciences, causal modeling can be very interesting due to its

ability to combine cause-effect information with statistical data to provide a quantitative – stochastic- assessment of the phenomena under study (Anderson and Vastag, 2004).

The main approach for representing a causal model is a Bayesian Network (BN) (Pearl 2000, Darwiche 2010). Among a variety of fields and in issues related to R&D and innovation, BNs have also been used to assess innovation collaboration processes (Kim and Park, 2008), webometrics (Lamirel et al., 2004) and to map the internal configuration of the efficiency results of research groups (Ruiz et al., 2009). Nevertheless, to our knowledge they are not used as conceptual models for modeling the R&D and innovation processes and as a tool for the evaluation of RTDI public policies.

On the other hand, Data Envelopment Analysis (DEA) is a non-parametric technique for assessing the technical relative efficiency of a set of comparable Decision-Making Units (DMUs), which use multiple inputs to produce multiple outputs (Cooper et al., 2004). This analysis identifies the most efficient DMUs (Pareto-optimal) that are then used as the standard for comparison with the rest of the DMUs –which are inefficient– for analysing potential improvements for them. Rousseau and Rousseau (1997) were the first using a DEA approach to assess the relative efficiency of R&D processes in 18 developed countries. DEA has also been used to approximate the efficiency of universities (Bonaccorsi et al. 2006), research centers (Korhonen et al. 2001), research groups (Guan and Wang, 2004) and R&D activities (Wang and Huang, 2007).

In this study, the relative technical efficiency of RTDI public policies in the 27 EU Member States was assessed through a new approach called Two-Linked-Steps DEA (2LS-DEA) based on a BN that identifies the inputs and outputs. By analyzing the preliminary 2LS-DEA results, the BN was improved in an iterative way and final efficiency scores for the EU Member States were obtained. To achieve this, various scenarios (combinations of specific inputs and outputs) were designed to analyze the problem. The corresponding inputs and outputs for each scenario were chosen according to the causal relationships previously defined in the BN. Original input and output values were interpreted by a knowledge base (rule-base) that stores the explicit expert knowledge used to design input/output causal relationships. This interpretation process was carried out automatically by a linear monotone algebraic transformation.

The aim of the present research is twofold: i) to design a conceptual model (Bayesian Network) that identifies the key variables and their relationships for describing R&D and innovation processes and their socioeconomic impacts in a sustainable development framework and ii) the use of the 2LS-DEA in an iterative way to evaluate the model and to

obtain the relative technical efficiency of the 27 RTDI public policies of the 27 EU Member States.

In this way, after introductory section 1, the underlying assumptions involved in the BN, the variables and its structure are presented in section 2. In section 3, the methodology and the employed dataset are described in order to assess the technical relative efficiency of RTDI public policies in the 27 EU countries employing a Two-Linked-Steps DEA (2LS-DEA) approach. Empirical results of the application of 2LS-DEA and the evaluation of the model by these results are presented in section 4. Finally, conclusions as well as future lines of research are outlined in section 5.

2. The Bayesian Network: a conceptual model

A Bayesian Network (BN) can be used for modeling complex and stochastic domains under uncertainty (Pearl 2000, Darwiche, 2010). They are usually shown as graphs, which typically have the form of a network diagram (Jensen, 2001), and incorporate variables (nodes) and their probabilistic relationships (arcs). BNs capture uncertainties in terms of probabilities to perform reasoning and inference, motivated by the need to make rational decisions (Darwiche, 2010). In general, a BN has two parts. The first part is qualitative, where nodes represent variables, and directed arcs describe conditional causal relationships. The second part is quantitative, where causal relationships are expressed in terms of joint conditional-probability distributions (Nadkarni and Shenoy, 2001). This structure lets BNs be combined with analytic decision tools like, for example, operational models (Jensen, 2001).

The most common network diagram for a BN is the Direct Acyclic Graph (DAG) in which each arrow/arc/connection that joins two variables depicts causal dependence, and two-headed arrows and feedback loops are not permitted (Pearl, 2000). A DAG could be solved using statistical and artificial intelligence-based techniques.

Some of the disadvantages of a BN include: i) the difficulty in obtaining probability information to make predictions, ii) it can be hard to handle many random and continuous variables and, finally, iii) it may not contain the appropriate variables if the prior explicit knowledge about the domain is wrong. Nevertheless, BNs are excellent conceptual models for forecasting because they can incorporate simulation techniques and fuzzy inference engines to evaluate imprecise causal relationships.

2.1. The design of the BN: underlying assumptions

A BN can be considered as a conceptual model for identifying the key variables and their causal relationships in R&D and innovation processes. Thus, the basic principles of the Endogenous Growth Theory, where economic growth and development are driven by technological change were used to design our preliminary BN. According to this theoretical framework, economic growth results from the R&D efforts of public (government and higher education) and private agents (business sector). This BN also integrates the principles of the Linear Model of Innovation, the Schumpeterian tradition and the Quintuple Helix Model. According to that, the structure of the BN (variables and relationships) are based on four assumptions:

Assumption 1. The BN was based on the Endogenous Growth Theory. It incorporates technological change as an endogenous factor to the economic growth process. It initially emphasizes the role of knowledge as an input of production (Romer, 1990). Since then, many empirical studies have explored how and to what extent innovation (as an output of R&D) might contribute to economic growth (Brinkman and Brinkman, 2001).

Assumption 2. The Linear Model of Innovation was selected in our research because policy makers understand that R&D investments generate innovations in a linear way and, consequently, official databases are designed following this model (Godin, 2006; Balconi et al., 2010). There are different and more complex models for describing innovation process like the chain-linked model (Kline and Rosenberg, 1986) and the cycle model (Padmore, 1998) that establish that the key element of the innovation process was the fact that research, development and innovation are linked in a non-linear and interactive way.

Assumption 3. Schumpeter's work (1934) established five types of innovations: a new thing; a new process for making a familiar thing; a new market to which to sell; a new source of supply or raw material; and industrial reorganization. Moreover, he analysed the influence that firms' size and market structure had on innovation.

Assumption 4. The Triple Helix Model (Etzkowitz and Leydesdorff, 1998) was considered to be a model that engages synergies from three main R&D stakeholders: universities, which develop academic-based research and development activities; industries, which develop commercial activities as well as R&D activities; and, finally, governments as the policy makers. The Quadruple Helix adds media-based and culture-based public and civil society as the fourth helix and, finally, the Quintuple Helix innovation model incorporates the "natural environments of society" helix as well (Carayannis and Campbell, 2010). The latter represents socio-ecological transition from the knowledge economy.

2.2. Main variables (nodes) and structure (relationships) of the BN

Relevant variables and relationships for R&D and innovation process were identified and defined (Appendixes A and B) as well as represented in the BN (Figure 1). The BN shown in Figure 1 is the last version after various improvements that incorporate the elicited implicit knowledge obtained from experts once 2LS-DEA results (efficiency scores) were statistically analyzed.

In the literature on the measurement of the innovation, certain authors (Chiesa et al., 1996; Hansen, 2001) propose some audit tools based on observable indicators to enable decision makers to collect data on innovation. The Organization for Economic Cooperation and Development (OECD) has elaborated some guidelines for the collection of data on R&D and innovation in the Frascati Manual (OECD, 2002), the Oslo Manual (OECD and Eurostat, 2005) and Main Science and Technology Indicators (OECD, 2010), which are widely used. The construction of our model was based on these manuals and the related literature² from which we have selected the variables.

(Figure 1: Bayesian Network conceptual model, about here)

The inputs and outputs to evaluate the efficiency of the RTDI EU policies in various socioeconomic scenarios were selected from the BN and algebraically interpreted according to their relationships. The BN understands a time lag of two years; that is that the effects of investments in R&D carried out in the year $t-2$ will be produced in the year t (Greif, 1985).

R&D expenditures and R&D personnel are the main indicators for the inputs of R&D and innovation activities. The former reflects the degree of relevance that agents place on R&D and their capability for generating innovation and the latter assesses the number of persons that are engaged in fundamental research, application research and experimental development (OECD, 2010). R&D expenditures and R&D personnel include both the public sector ($ERDFUN_{t-2}$ for R&D expenditures and $PUBPER_{t-2}$ for R&D personnel as a percentage of total employment) and the private sector (R&D business sector expenditure, $ERDPRI_{t-2}$, and R&D personnel, $PERPRI_{t-2}$). Only the public sector was considered to assess EU RTDI policies ($ERDFUN_{t-2}$ and $PUBPER_{t-2}$).

The number of scientific publications ($SPUBLI_t$) and patent applications ($TPATRE_t$) (OECD, 2010) are the most common variables for the outputs of R&D and innovation

² The works of Godin (2002, 2006, 2007) deserve special mention in this research.

activities. They are influenced by $ERDFUN_{t-2}$, $PUBPER_{t-2}$ and by the infrastructures of the country ($INFRAE_t$ machinery and equipment, communication infrastructures, vehicles, dwellings and other buildings, including laboratories and other scientific infrastructures). Some studies highlight that $TPATRE_t$ can be highly influenced by $SPUBLI_t$ (Meyer, 2000).

Following the structure of the BN, $ERDFUN_{t-2}$ is influenced by the existing EU strategic guidelines ($EUSTRAT_{t-2}$) and $PUBPER_{t-2}$ depends on: i) the level of mobility of researchers and PhD students ($MOBILI_{t-2}$) and ii) the attractiveness of the market for researchers due to the wage level of the countries ($MARKRE_{t-2}$).

For the private sector, $ERDPRI_{t-2}$ (R&D private expenditure) is influenced by public support ($ERDFUN_{t-2}$) in the form of subsidies, tax incentives, etc., the size of the firms ($SIZENT_{t-2}$) and the structure of the private sector for the corresponding country ($SECEST_{t-2}$) (Yu, 2005). The personnel involved in R&D activities in the private sector mainly come from $PUBPER_{t-2}$ and depend on $ERDPRI_{t-2}$.

According to Schumpeter's types of innovation (Schumpeter, 1934), the BN embodies the difference between technological innovations measured by patents ($TPATRE_t$) and non-technological innovations represented by proxy indicators like community trademarks ($TRADEM_t$), small and medium enterprises introducing marketing/organizational innovations ($INNSME_t$) and sales of new to markets and new to firm innovations as a percentage of the turnover ($INNEW_t$). Technological advance would not be possible without the development of new ways of organization to guide and support R&D and enable firms to obtain profits from this kind of investment (Nelson, 1996).

$TPATRE_t$ and indicators for non-technological innovations ($TRADEM_t$, $INNEW_t$ and $INNSME_t$) depend on a variety of variables like $INTECAP_t$ (intellectual capital: the difference between tangible physical and financial assets and human, structural and relational capital market values; Bontis et al., 1999), $SPILLOV_t$ (spillovers: some of the economic benefits of R&D activities go to economic agents other than the ones undertaking the research; Block, 2012) and social and business networks $SBNETW_t$ (connections through individuals and organizations; Pavitt, 1998; Parker, 2008). These variables emphasize that innovation can also take place without a systematic research and development process. Other influencing variables were: learning-by-doing ($LEARDO_t$) [Von Hippel and Tyre, 1995], benchmarking ($BENCM_t$) or "the continuous process of identifying, understanding and adapting practice and processes that will lead to better performance" (Auluck, 2002; p. 111) and sociocultural aspects like the existence (or not) of an entrepreneurship culture in the country, $ENTREP_t$ (Braunerhjelm, 2008).

As mentioned in section 1, the main reason for assessing the impacts of RDTI policy is its contribution to public objectives as economic growth, health outcomes or education (Boulanger and Bréchet, 2005), to justify not investing in other alternatives while creating a better-informed society to satisfy the increasing demand for public accountability.

Following this underlying idea, the resource allocation in R&D and innovation ($ERDFUN_{t-2}$, $PUBPER_{t-2}$, $ERDPRI_{t-2}$ and $PERPRI_{t-2}$), which leads to patents and scientific publications, would be desirable to positively affect global society welfare ($WELFAR_t$).

Since productivity growth plays a main role in increasing the Gross Domestic Product (GDP), it is important to point out the main factors underlying the growth of such productivity. The infrastructures of the country ($INFRAE_t$), employees' motivation ($MOTIVA_t$), the labour climate ($LABCLI_t$) and technological change ($TECCHA_t$) play a very important role in our BN to explain the Real Labour Productivity Growth per hour worked ($RLPGH_t$). Furthermore, the level of technological change ($TECCHA_t$) depends on technological innovation ($TPATRE_t$), the total R&D expenditure as % of GDP ($GERDTO_t$, which comprises the expenditure from public, private, foreign and non-profit sectors in the year t); and the technological balance of payments ($TECACQ_t$) that includes various forms of technological acquisition, such as the purchase of capital goods or technological services, and including the licensing of intellectual property. Other variables that strongly influence $RLPGH_t$ are those concerning education: $PHD06_t$ (number of Ph.D.), $ELEAV_t$ (early leavers) and $LLEARN_t$ (life-long learning).

As the Global Competitiveness Report of the World Economic Forum (2010) set out, productivity determines the return rates of the investments (physical, human, and technological) made in an economy. The $RLPGH_t$ influences the general competitiveness of a country and, in turn, sets the sustainable level of prosperity that can be reached by an economy. In other words, economies that are more competitive tend to produce higher levels of income for their citizens ($GDPGRO_t$). Conversely, a higher level of productivity per hour worked could lead to higher levels of unemployment.

Economic impacts would measure the extent to which traditional outputs (scientific publications and patents) indirectly influence and address: Real Labor Productivity Growth per hour worked ($RLPGH_t$), the corresponding GDP growth ($GDPGRO_t$) as well as their implications for High-Technology Exports ($HITECH_t$) and the reduction in the unemployment rate ($UNEMPL_t$).

The institutional framework of a country ($INSTFR_t$) also plays a relevant role in a society's welfare (World Economic Forum, 2010). The importance of the $INSTFR_t$ has

become even more significant during the economic crisis, given the direct role played by the States in the economy. This legal framework (*INSTFR_t*) has an important influence on the variable *TPATRE_t* (intellectual rights must be protected by the law to be effective), affects the image of a country abroad (*COUIM_t*), foreign direct investment (*FORINV_t*) and market efficiency (*MARKEF_t*). *COUIM_t* can also have a relevant influence on the latter variable.

The BN shows the effects on competitiveness (*COMPET_t*) that may be achieved due to productivity growth (*RLPGH_t*), technological (*TPATRE_t*) and non-technological innovation indicators (*TRADEM_t*, *INNEW_t* and *INNSME_t*) and market efficiency (*MARKEF_t*).

Finally, all these variables and their causal relationships converge into the most popular measure for determining the levels of social welfare (*WELFAR_t*), Gross Domestic Product Growth (*GDPGRO_t*).

Due to the complexity of the BN, the rest of the variables are described in section 3.2 when describing the scenarios for 2LS-DEA.

3. Evaluation of the BN: the relative technical efficiency of EU RTDI public policies

3.1. Data Envelopment Analysis (DEA)

Data Envelopment Analysis (DEA) is a non-parametric, mathematical programming-based technique for measuring and evaluating the relative efficiency of decision-making units (DMUs), which use multiple inputs to produce multiple outputs making it appropriate for comparing the efficiency of R&D in terms of multiple innovative outputs. Charnes, Cooper and Rhodes promoted this technique but its roots may be found as early as 1957 in Farrell's seminal work (Cooper et al., 2004).

In relation to the specific model for running DEA, the variable returns-to-scale (or BCC model) was selected because R&D activities can exhibit increasing or decreasing returns to scale (Bound et al., 1984). It is also implicitly assumed that the emphasis on public policies is more on what is produced and based on given resources. Therefore, this study adopted the output-oriented BCC model. The free imputation of input-output values was selected for the identification of inefficiencies (Saen, 2010).

Considering a set of n DMUs ($DMU_j; j=1, 2, \dots, n$), that each DMU_j consumes m inputs (x_{ij}) in different amounts ($x_{ij}; i=1, 2, \dots, m$) and that produces s outputs (y_{rj}) in different quantities ($y_{rj}; r=1, 2, \dots, s$), the relative technical efficiency of a DMU (output-oriented) can be evaluated by the envelopment form using the following linear model (Cooper et al., 2004):

$$\max \phi + \varepsilon \left(\sum_{i=1}^m s_i^- + \sum_{i=1}^m s_i^+ \right)$$

Subject to:

$$\begin{aligned}
 \sum_{j=1}^m x_{ij} \lambda_j + s_i^- &= x_{i0} & i=1, 2, \dots, m \\
 \sum_{j=1}^n y_{rj} \lambda_j + s_r^+ &= \phi y_{r0} & r=1, 2, \dots, s \\
 \lambda &\geq 0 \\
 \sum_{j=1}^n \lambda &= 1
 \end{aligned} \tag{1}$$

where ϕ represents the efficiency score, x_{ij} and y_{rj} are the i th input and the r th output, respectively, of the j th DMU. Similarly, x_{i0} and y_{r0} stand for the inputs and outputs, respectively, of the DMU under evaluation (DMU₀). In addition, λ_j denotes nonnegative scalars and s_i^- and s_r^+ are input and output slacks, respectively.

To apply DEA effectively requires (Cooper et al., 2007) that:

$$n \geq \max \{m \times s, 3 \times (m + s)\} \tag{2}$$

because the whole technique is based on the comparison of each DMU with all the remaining ones, thus a considerably large set of DMUs is necessary to be meaningful.

3.2. Structure of DEA scenarios: inputs and outputs

The evaluation of the BN design was performed by assessing the relative technical efficiency RTDI public policies in EU Member States analyzing selected scenarios (input and output combinations). All inputs and outputs comes from the BN (Appendix A) but not all variables in the BN were employed as inputs and outputs of the selected socioeconomic scenarios (Appendix B). The scenarios were³:

1. Traditional scenario. The relative technical efficiency of RTDI public policies in the EU was evaluated taking into consideration two inputs (R&D public expenditure, $ERDFUN_{t-2}$,

³ All variables in the BN are circled individually (nodes) except the ones included in the above mentioned scenarios for better visualization in Figure 1.

and R&D public personnel, $PUBPER_{t-2}$) and two outputs (scientific publications, $SPUBLI_t$, and patents applications, $TPATRE_t$).

The number of scientific publications is a common indicator of academic productivity but it is not the only one. Project reports, monographs and other literature are equally important (Thomas et al., 2011). However, as data on such output are not available, we use $SPUBLI_t$. The number of patents is the most widely accepted indicator for innovation due to the lack of more consistent indicators (Griliches, 1990). This also shows some disadvantages: not all inventions are technically patentable, not all inventions are patented and, finally, firms have different propensities to patent in their domestic markets and in foreign countries.

2. Socioeconomic scenarios. These scenarios are combinations of inputs and outputs to describe different faces of a sustainable socioeconomic development. All of them have the same inputs and different outputs. The inputs are the number of scientific publications $SPUBLI_t$ and number of patents $TPATRE_t$, the outputs of the traditional scenario. The aim is to analyse the contribution of these inputs to the performance of each DMU (RTDI public policies in the EU) in Health, Environment, Information Society, Education, Economic and Innovation scenarios. In order to carry out 2LS-DEA, only a selected percentage of the original values of $TPATRE_t$ and $SPUBLI_t$ was considered. A panel of experts determined this percentage because only the public sector is analyzed. If a country is efficient in the traditional scenario, it is expected to be efficient in socioeconomic ones.

Health. Improvements in the health sector due to R&D and innovation investments are well-known and their expenditures are usually well received by society. Three outputs were used to analyse this scenario: $AGEDEP_t$ (age dependency ratio), $MORRAT_t$ (mortality rate) and $LIFEXP_t$ (life expectancy).

Environment. Obviously, without a sustainable environment, life on earth would be in danger. The variables (outputs) selected to measure the efficiency in this scenario were $ENINT_t$ (energy intensity of the economy), $CARBDI_t$ (carbon dioxide damage) and $GASEM_t$ (total greenhouse gas emissions).

Information Society. The most relevant sector in which R&D outputs had an exponential influence is Information and Communication Technologies (ICT). With the increasing relative weight of the services sector induced by the knowledge economy in developed countries, ICT are now even more incalculably important. The indicators for this scenario were $BROSUS_t$ (fixed broadband Internet subscribers), $INTUSE_t$ (internet users) and $MOBILE_t$ (cellular mobile subscriptions).

Education. The European population has to acquire a range of skills while continuously adapting new ones. Knowledge and technology diffusion relevance are included in this scenario, which plays a very important role in employment. The outputs were: $PHD06_t$ (number of doctorates), $ELEAV_t$ (early leavers from education) and $LLEARN_t$ (life-long learning).

Economic. This scenario focuses its attention on economic prosperity, a key objective reflected in EU strategies. The selected outputs were: Gross Domestic Product growth ($GDPGRO_t$), real labour productivity growth per hour worked ($RLPGH_t$), the unemployment rate ($UNEMPL_t$) and High-Technology Exports ($HITECX_t$).

Innovation. This scenario reflects the concept of innovation defined in the Oslo Manual (OECD and Eurostat, 2005). The selected outputs are community trademarks ($TRADEM_t$), small and medium enterprises introducing marketing/organizational innovations ($INNSME_t$) and sales of new to markets and new to firm innovations as a percentage of the turnover ($INNEW_t$).

3.3. Two-Linked-Steps DEA approach (2LS-DEA)

The efficiency of EU RTDI public policies was not only assessed using the traditional scenario –step 1–, but also by analyzing six socioeconomic scenarios –step 2, linked with the first. Although the BN causal model represents the R&D and innovation process as a whole, this research considers only the public sector. For this reason, in the traditional scenario, only 10% of the value of total patent applications ($TPATRE_t$) comes from $PUBPER_{t-2}$ and $ERDFUN_{t-2}$ and only 90% of the value of SPUBLIt comes from $PUBPER_{t-2}$ and $ERDFUN_{t-2}$. In Appendix A (% column) it is specified the corresponding percentage of the original output value in each socioeconomic scenario.

In step one of the 2LS-DEA the traditional scenario is evaluated and in step 2 the socioeconomic scenarios are assessed, where traditional outputs turn into the inputs in all of them in a way that both steps are “linked”.

3.4. Dataset

The 27 RTDI public policies of the corresponding EU countries were selected as DMUs. Inputs and outputs were quantitative data (continuous) according to the conceptual model (BN) discussed in section 2. Data were collected from various official databases: Eurostat’s official website for all variables, except the World International Patent Organization (WIPO, 2011) for patent applications, SCImago Journal & Country Rank in the case of the number of scientific publications per country (SCImago, 2007) and the World Bank Database (World

Bank, 2012) for age dependency ratio ($AGEDEP_t$), broadband subscribers ($BROSUS_t$), high-tech exports ($HITECEX_t$) and mobile cellular subscriptions ($MOBILE_t$). Variable values were considered per 1,000 inhabitants when necessary. Table 1 illustrates descriptive statistics for all inputs and outputs.

(Table 1 about here)

The main sources for defining qualitative variables and the intervals of quantitative ones were the Global Competitiveness Index (WEF, 2010), ERAWATCH profiles of European Countries in the European Research Area⁴ and an expert panel's opinion. They were used for interpreting the values of variables for building the rule-base in the next section 3.5.

Appendix A shows the information of quantitative variables scenario by scenario and Appendix B the qualitative variables represented in the BN model.

The reference period was 2008 due to the availability of data in the Community Innovation Survey of Eurostat (CIS) that is carried out every two years (Eurostat, 2012). Meanwhile, considering the time lag for the innovation process, data acquisition was 2006 for inputs in the traditional scenario. As mentioned in section 2.2, this means that if a country promotes R&D investments in 2006 (or year $t-2$), outputs are not expected to increase until at least two years later (2008 or year t).

3.5. Expert-based input and output transformation

In order to design BCC-DEA models, the original values of quantitative variables in the dataset were transformed by applying the expert-based rule-base that emerges from the BN. Classical DEA models interpret that the greater the outputs, the greater the efficiency of a specific DMU (if input values are constant) and the lower the inputs, the greater the efficiency (if output values are constant). Output-oriented DEA tries to maximize the output values once the input values have been set. However, in some situations some inputs need to be increased and some outputs need to be decreased to improve the efficiency of a DMU (non-standard inputs and outputs). In these cases (for example, $UNEMPL_t$ and $AGEDEP_t$), we performed a linear monotone increasing/decreasing transformation on their original values based on the expert-based rule-base.

⁴ ERAWATCH is the European Commission platform on research policies and systems in the EU and beyond. (<http://erawatch.jrc.ec.europa.eu/>)

The rule-base is a set of *if-then* rules that emerged from the causal relationships in the BN model and the characteristics of the variables (Table 2). Transformed variable values were used to design final DEA models to assess the relative technical efficiency of EU RTDI public policies incorporating expert knowledge.

(Table 2 about here)

For example, a simple linear transformation rule transforms $GDPGRO_t$ according to the influence of $RLPGH_t$. If $GDPGRO_t \leq 1$ (*low*) and $RLPGH_t > 3$ (*high*), then the original value of $GDPGRO_t$ is penalised dividing its original value by two (R15). The reason for this rule relies on the fact that high levels of productivity must lead to increasing rates of GDP growth; when $GDPGRO_t$ is *low* and $RLPGH_t$ is *high*, the country analyzed is not achieving this goal. On the other hand, if $GDPGRO_t > 3$ (*high*) and $RLPGH_t \leq 1$ (*low*) then the original value of $GDPGRO_t$ is divided by two (R14 in table 2) to make DEA understand that this situation means a very high country performance. A special rule emerges in the case of Eastern economies due to their very high GDP growth rate that mainly represents the transition economies of the New Member States. As new economies, they involve a high rate of growth in GDP and in Real Labor Productivity growth per hour worked ($RLPGH_t$) due to the abundance of natural resources, low population and a high domestic demand but a low rate of employment after a period of socialism, among other reasons. To make DEA understand this, the real value of $GDPGRO_t$ is transformed by the rule R16:

$$\text{IF } GDPGRO_t > 5 \text{ THEN } GDPGRO_t' = GDPGRO_t / 2 \quad (3)$$

It is also interesting to highlight the transformation of $ERDFUN_{t-2}$ accomplished with R1, R2 and R3. $ERDFUN_{t-2}$ depends on European Strategy ($EUSTRAT_{t-2}$ node in the BN). Due to Lisbon Strategy's targets applicable in the period under study (to reach at least 3% of GDP of R&D effort and that 60% of total R&D expenditure must be funded by the private sector), $ERDFUN_{t-2}$ was transformed according to R1, R2 and R3 rules:

$$\text{IF } ERDFUN_{t-2} / GERD_{t-2} \leq 0.4 \text{ and } GERD_{t-2} > 3 \text{ THEN } ERDFUN_{t-2}' = ERDFUN_{t-2} * 2 \quad (4)$$

$$\text{IF } ERDFUN_{t-2} / GERD_{t-2} > 0.55 \text{ AND } GERD_{t-2} < 2 \text{ THEN } ERDFUN_{t-2}' = ERDFUN_{t-2} / 2 \quad (5)$$

$$\text{IF } ERDFUN_{t-2} \leq 0.2 \text{ THEN } ERDFUN_{t-2}' = ERDFUN_{t-2} / 2 \quad (6)$$

4. Empirical Results

The 2LS-DEA approach provided a ranking of 27 EU RTDI public policies based on the relative technical efficiencies obtained in each scenario (Table 3). An efficiency score equal to 100 means that the RTDI public policy of a country is efficient (relative efficiency); values lower than 100 mean different grades of inefficiency.

(Table 3 about here).

In order to summarize and analyze the efficiency scores, a *k*-means clustering was carried out and four clusters were chosen:

Cluster 1 includes Bulgaria, Estonia, France, Italy, Latvia, Poland and Romania. Their RTDI policies are the least efficient in the Traditional scenario. They also appear as the least efficient in Innovation and Information society scenarios and they show low efficiency scores in the Education scenario.

France and Italy belong to this cluster due to the poor relationship between their public efforts and the traditional inputs (patents and scientific publications) and in the innovation and education scenarios.

Cluster 2 includes the most efficient RTDI policies. It groups Austria, Belgium, Denmark, Finland, Germany, Ireland, Lithuania, Luxembourg, Netherlands, Sweden and the United Kingdom. Sweden is efficient in all the scenarios under study followed by Belgium (six scenarios) and Finland (five scenarios). Ireland and Luxembourg rank among the five top in all but three scenarios and almost reach efficiency in six scenarios. Lithuania appears in this group mainly due to its efficiency in the Traditional and Economic scenarios. In spite of the relatively low amount of resources devoted to R&D by its public sector, it shows a high number of scientific publications.

Cluster 3 includes Cyprus, Czech Republic, Hungary, Slovakia, Slovenia and Spain. They show low efficiencies in the Traditional and Environment scenarios but they perform very well in the Innovation scenario.

Finally, **cluster 4** includes Greece, Malta and Portugal RDTI public policies. It groups the least efficient RTDI policies. They show low efficiencies in the Traditional scenario and, this group reaches the lowest efficiencies in the Economic, Education and Health scenarios. Nevertheless, they are efficient in the Information Society scenario (very high number of mobiles users, especially in Greece).

In order to discriminate between the most efficient countries, a complementary measure was used that involved the number of times that those countries were peers for others in the DEA assessment. Sweden, Belgium, Finland and Ireland RDTI public policies were the most frequently selected by inefficient ones as benchmarks.

The 2LS-DEA results were compared to the Innovation Union Scoreboard (IUS) scores per dimensions (European Commission, 2010) that were comparable with traditional, innovation and economic scenarios. The IUS report uses 24 variables and develops a composite indicator (Summary Innovation Index). Mann–Whitney U test results showed some relevant discrepancies between them that could explain the utility of the 2LS-DEA approach, which incorporates expert knowledge by means of the rules-base. For example, Lithuania's efficiency score differs from IUS classification because some variable values were transformed through the rule-base. Thus, in the Traditional scenario, $ERDFUN_{t-2}$ value was penalized dividing by two and two multiplies the original value of $PUBPER_{t-2}$ as R1, R2 and R3 rules were applied to the original dataset (see Table 2). This country is efficient in the Traditional scenario but it is not in socioeconomic ones, except in Economic. Thus, it is not verified the initial hypothesis that if a country is efficient in the Traditional scenario has to be efficient in socioeconomic ones.

Another country where our efficiency result differs from IUS scores is France. In this case, $SPUBLI_t$ was penalized by R11 rule, mainly because the number of scientific publications was *low* (<1.5 per 1,000 inhabitants) with respect to public efforts ($ERDFUN_{t-2}$ and $PUBPER_{t-2}$).

Other RTDI public policies are efficient in a large number (more or equal than three) of socioeconomic scenarios but cannot be considered efficient in the Traditional scenario. This group includes countries like France, Ireland, Italy, Luxembourg or Slovenia.

To conclude, our proposal is not only to classify the countries by means of the resulting scores but also to evaluate the technical efficiency of RTDI public policies. IUS includes public and private innovation variables so it would not be possible to distinguish the effect of specific public strategies.

5. Conclusions

Evaluation of the relative technical efficiency of RTDI public policies is an important issue for policy-makers. The knowledge-based economy and sustainable growth are concepts strongly related to R&D activities and their specific policies. Thus, the underlying idea is that the more efficient the countries are in the employment of corresponding R&D inputs, the

more efficient they are supposed to be in the socioeconomic scenarios selected in this study. If a specific country does not employ its R&D outputs efficiently as inputs for other socioeconomic scenarios (health, education, economic development, etc.), other priorities and social alternatives arise and take precedence over increasing expenditure for R&D. Improvements in such efficiency can lead to a growing competitiveness, but decision-making is not an easy task.

The proposed BN conceptual model and the Two-Linked-Steps DEA (2LS-DEA) approach address the main challenges for assessing public R&D impacts on society: i) the identification of the relationships between R&D inputs, outputs and their socioeconomic impacts; ii) the existence of direct or unidirectional relationships between inputs and outputs (our BN only includes direct causal relationships); iii) the assessment of what proportion of social benefits should be attributed to R&D inputs; iv) the effect of time when measuring those impacts (time lag) and, finally, v) the use of appropriate indicators (selection of variables).

The crucial points in the BN design were: i) interaction with experts –prior expert knowledge acquisition, ii) data pre-processing with the application of a rule-base and, finally, iii) the interpretation of the results of the 2LS-DEA. Then, the conceptual model (BN) and the rule-base were improved iteratively by eliciting implicit knowledge and until no inconsistencies were found. Finally, the 2LS-DEA assessed the (final) relative technical efficiency of RTDI public policies in the 27 European Union Member States. Thus, the 2LS-DEA approach was used both to improve the design of the BN (to provide evidence of its design as a conceptual model) and for assessing such relative technical efficiency.

Some of the abilities of this methodology are to transfer expert knowledge to data analysis and to elicit implicit knowledge, as it is an operational tool for providing fundamental information for decision-making processes. In this way, the results do not depend on subjective-weight assumptions and arbitrary selection of inputs and outputs but on the incorporation of expert knowledge into the dataset by means of the rule-base that was built based on the BN.

Other methodologies study R&D and innovation performance or, more widely, the competitiveness and the state of the knowledge-based economy of countries yielding comparable results. However, the aforementioned multidimensional approaches ignore causal relationships between variables and cannot interpret special combinations of variables (scenarios) that result from specific public policies (sometimes inefficient).

As regards the results of the 2LS-DEA approach, and as to be expected given the level of economic development in the 27 EU member States, countries like Sweden, Finland, Luxembourg and Denmark were at the top of multidimensional efficiency rankings. The least efficient RTDI public policies were in Bulgaria (it was not efficient in any scenario), Estonia, Latvia and Romania.

The diverse socioeconomic efficiency scenarios were presented to contribute to the discussion on European policies. The EU2020 Strategy, the European Research Area (ERA) or the inclusion of related indicators in the core of the Eurostat reveal the relevance of achieving sustainable development for the European Union. Since most of the scenarios examined in this study are related to public policy variables (education, health, environment, infrastructures, etc.), it can be said that our empirical findings should be of interest to policy makers for monitoring the realization of sustainable and smart European Strategies.

On the other hand, the model can provide main stakeholders with the dimensions that need improvements, which can guide a political rationale to increase or decrease the amount of inputs devoted to R&D and innovation funds. With the BN conceptual model and the 2LS-DEA approach, it is possible to identify the most efficient RDTI public policies in the socioeconomic scenarios designed, which opens a door for benchmarking.

Obviously, the resulting BN would be different if other assumptions were adopted. Nevertheless, the main purpose of this research was achieved by its evaluation and the demonstration of its potential for the analysis of RTDI public policy performances in the European Union, by using a 2LS-DEA approach in both cases.

The model can be easily adjusted in terms of the variables selected and the rule-base structure. These modifications can be based on other theoretical frameworks, as well as on other perspectives from different stakeholders. Additionally, this methodological approach can perform future inferences by incorporating a Monte Carlo simulation engine. This procedure could evaluate the effects of alternative investments in RDTI public policies and study their effects on selected scenarios.

Finally, this study could be addressed to compare both the results of the period employed (2006-2008), before the financial and economic crisis, and the next period (from 2009 to the more recent available data) to analyze the effects of the economic crisis on efficiency scores, all of these considerations being a vast field of future research.

Acknowledgements

This work was partially subsidized by the Spanish Inter-Ministerial Commission of Science and Technology under Project TIN2011-22794, the European Regional Development fund, and the “Junta de Andalucía” (Spain), under Project P2011-TIC-7508.

Among others, we would like to thank the expert opinions and comments of Prof. Gabriel Pérez-Alcalá (Rector at the Loyola-Andalusia University, Spain); Prof. Lorenzo Salas-Morera (Director of Academic Organization and Degrees, Córdoba University, Spain); Prof. Fernando Lara-De Vicente (Vice Dean of Academic Planning, Faculty of Labor Science, Córdoba University, Spain); Francisco Solís-Cabrera, Secretary of Andalusia Research Plan, Andalusia Government, Spain; Prof. M^a Teresa Sánchez-Pineda de las Infantas (Vice rector of Innovation and Quality Management, Córdoba University); Prof. Panagiotis Zervatopoulos (University of Ioannina, Department of Business Administration of Food and Agricultural Products, Greece).

References

- Aghion, P., Howitt, P., 1992. A model of growth through creative destruction. *Econometrica* 60, 323-351.
- Anderson, R.D., Vastag, G., 2004. Causal modeling alternatives in operations research: Overview and application. *European Journal of Operational Research* 156, 92-109.
- Archibald, R. B., Pereira, A. M., 2003. Effects of public and private R&D on private-sector performance in the United States. *Public Finance Review* 31, 429-451.
- Auluck, R., 2002. Benchmarking: A tool for facilitating organizational learning? *Public Administration and Development* 22, 109-122.
- Balconi, M., Brusoni, S., Orsenigo, L., 2010. In defence of the linear model: An essay. *Research Policy* 39, 1-13.
- Baumol, W. J., 2003. Innovations and growth: Two common misapprehensions. *Journal of Policy Modeling* 25, 435-444.
- Block, J.H., Thurik, R., Zhou, H., 2012. What turns knowledge into innovative products? The role of entrepreneurship and knowledge spillovers. *Journal of Evolutionary Economics*, DOI: 10.1007/s00191-012-0265-5.
- Bonaccorsi, A., Daraio, C., Simar, L., 2006. Advanced indicators of productivity of universities. An application of robust nonparametric methods to Italian data. *Scientometrics* 66, 389-410.
- Bontis, N., Dragonetti, N.C., Jacobsen, K., Roos, G., 1999. The knowledge toolbox: a review of the tools available to measure and manage intangible resources, *European Management Journal* 17, 391-402.
- Boulanger, P.M., Brochet, T., 2005. Models for policy-making in sustainable development: The state of the art and perspectives for research. *Ecological Economics* 55, 337-350.
- Bound, J., Cumins, C., Griliches, Z., Hall, H.H., Jaffe, A., 1984. *R&D, Patent, and Productivity*. University of Chicago Press, Chicago.
- Braunerhjelm, P., 2008. Entrepreneurship, knowledge and economic growth. *Foundations and Trends in Entrepreneurship* 4, 451-533.
- Brinkman, R. L., Brinkman, J.E., 2001. The new growth theories: a cultural and social addendum. *International Journal of Social Economics* 28, 506-526.
- Carayannis, E.G., Campbell, D.F.J., 2010. Triple Helix, Quadruple Helix and Quintuple Helix and how do knowledge, innovation and the environment relate to each other? A proposed framework for a trans-disciplinary analysis of sustainable development and social ecology. *International Journal of Social Ecology and Sustainable Development* 1, 41-69.

- Chiesa, V., Coughlan, P., Voss, C.A., 1996. Development of a technical innovation audit. *Journal of Product Innovation Management* 13, 105-136.
- Cohen, W.M., Nelson R. R., Walsh, J.P., 2002. Links and Impacts: The Influence of Public Research on Industrial R&D. *Management Science* 48, 1-23.
- Cooper, W.W., Seiford, L.M., Zhu, J. (2004). Data envelopment analysis: history, models and interpretations, in: Cooper, W.W., Seiford, L.M., Zhu, J. (Eds), *Handbook on Data Envelopment Analysis*. Kluwer Academic Publishers, Boston, 1-45.
- Cooper, W.W., Seiford, L.M., Tone, K., 2007. *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software* (2nd ed.). New York: Springer.
- Darwiche, A., 2010. Bayesian networks. *Commun. ACM* 53, 12, 80-90. DOI: 10.1145/1859204.1859227
- David, P.A., Hall, B. H., 2000. Heart of Darkness: Modeling Public-Private Funding Interactions Inside the R&D Black Box. *Research Policy* 29, 1165-1183.
- De la Paz-Marín, M., Campoy-Muñoz, P., Hervás-Martínez, C., 2012. Non-linear classifiers for R&D performance in the EU. *Technological Forecasting and Social Change* 79, 1731-1745.
- Dosi, G., Llerena, P., Sylos Labini, M. (2009). Does the 'European Paradox' still hold? Did it ever? in: Delanghe, H., Muldur, U., Soete, L. (Eds), *European science and technology policy: Towards integration or fragmentation?* Edward Elgar, Cheltenham, pp. 214-236.
- Etzkowitz, H., Leydesdorff, L.A., 1998. The endless transition: A "triple helix" of university–industry–government. *Minerva* 36, 203–208.
- European Commission, 2009. *Mainstreaming sustainable development into EU policies: Review of the European Union Strategy for Sustainable Development*. COM (2009) 400 final, Brussels.
- European Commission. 2010. *Europe 2020 - a strategy for smart, sustainable and inclusive growth*. COM (2010) 2020. Brussels.
- Fagerberg, J., Mowery, D.C., Nelson, R.R., 2004. *The Oxford Handbook of Innovation*. Oxford University Press, Oxford.
- Fagerberg, J., Verspagen, B., 2009. Innovation studies - the emerging structure of a new scientific field. *Research Policy*, 218–233.
- Filippettia, A., Archibugia, D., 2011. Innovation in times of crisis: National Systems of Innovation, structure, and demand. *Research Policy* 40, 179-192.
- Freeman, C., 2004. Technological infrastructure and international competitiveness. *Industrial and Corporate Change* 13, 541-569.
- Freeman, C., Soete, L., 2009. Developing science, technology and innovation indicators: What we can learn from the past. *Research Policy* 38, 583-589.
- Furman, J.L., Porter, M.E, Stern, S., 2002. The determinants of national innovative capacities. *Research Policy* 31, 899-933.
- Gilbert, K., García-Alonso, C., Salvador-Carulla, L., 2010. Integrating clinicians, knowledge and data: expert-based cooperative analysis in health care decision support. *Health Research Policy and Systems* 8, 28-44.
- Godin, B., 2002. Technological gaps: an important episode in the construction of S&T statistics. *Technology in Society* 24, 387-413.
- Godin, B., 2006. The Historical Construction of an Analytical Framework, *Science Technology and Human Values* 31, 639-667.
- Godin, B., 2007. Science, accounting and statistics: The input-output framework. *Research Policy* 36, 1388-1403.
- Greif, S., 1985. Relation between R&D expenditure and patent applications. *World Patent Information* 7, 190–195.
- Griliches, Z., 1990. Patent statistics as economic indicators: a survey. *Journal of Economic Literature*, 28, 1661-1707.
- Grossman, G., Helpman, E., 1991. *Innovation and Growth in the Global Economy*. MIT Press, Cambridge, MA.
- Grupp, H., Schubert, T., 2010. Review and new evidence on composite innovation indicators for evaluating national performance. *Research Policy* 39, pp.67-78.

- Guan, J. C., Wang, J. X., 2004. Evaluation and interpretation of knowledge production efficiency. *Scientometrics* 59, 131-155.
- Hansen, J.A. (2001). Technology innovation indicators, a survey of historical development and current practice, in: Feldman, M.P., Link, A.N. (Eds.), *Innovation Policy in the Knowledge-Based Economy*, pp. 73–103.
- Jensen, F.V., 2001. *Bayesian Networks and Decision Graphs*. Springer-Verlag, New York.
- Jones, C.I., Williams, J.C., 1998. Measuring the social return to R&D. *The Quarterly Journal of Economics* 113, 1119-1135.
- Kim, H., Park, Y., 2008. The impact of R&D collaboration on innovative performance in Korea: A Bayesian network approach. *Scientometrics* 75, 535-554.
- Kline, S.J., Rosenberg, N. (1986). An overview of innovation, in: Landau, R., Rosenberg, N. (Eds), *The Positive Sum Strategy: Harnessing Technology for Economic Growth*. National Academy Press, Washington, D.C., pp. 275-305.
- Korhonen, P., Tainio, R., Wallenius, J., 2001. Value efficiency analysis of academic research. *European Journal of Operational Research* 130, 121-132.
- Lamirel, J. C., Al Shehabi, S., Francois, C., Polanco, X., 2004. Using a compound approach based on elaborated neural network for Webometrics: An example issued from the EICSTES project. *Scientometrics* 61, 427-441.
- Martin, B.R., 2012. The evolution of science policy and innovation studies. *Research Policy* 41, 1219-1239.
- Martin, S., Scott, J.T., 2000. The nature of innovation market failure and the design of public support for private innovation.
- Meyer, M., 2000. Does science push technology? Patents citing scientific literature. *Research Policy* 3, 409-434.
- Nadkarni, S., Shenoy, P.P., 2001. A Bayesian network approach to making inferences in causal maps. *European Journal of Operational Research* 128, 479-498.
- Nelson, R. R., (1996). *National Innovation Systems: A Retrospective on a Study*, in: *The Sources of Economic Growth*, Cambridge, MA, Harvard University Press, pp. 274-302.
- OECD, 2002. *The Measurement of Scientific and Technical Activities: Proposed Standard Practice for Surveys of Research and Experimental Development*, Frascati Manual (3rd ed.). OECD, Paris.
- OECD, 2010. *Main Science and Technology Indicators (MSTI) (2nd Ed)*. OECD Publications, Paris.
- OECD, Eurostat, 2005. *Oslo Manual Guidelines For Collecting And Interpreting Innovation Data (3rd ed.)*, OEDC Publications and Eurostat, Paris.
- Padmore, T., Schuetze, H., Gibson, H., 1998. Modeling systems of innovation: An enterprise-centered view. *Research Policy* 26, 05-624.
- Parker, S.C., 2008. The economics of formal business networks. *Journal of Business Venturing* 23, 627-640.
- Pavitt, K., 1998. The inevitable limits of EU R&D funding. *Research Policy* 27, 559-568.
- Pearl, J., 2000. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge.
- European Commission, Pro Inno Europe (INNO METRICS), 2010. *European Innovation Scoreboard 2010: Comparative Analysis of Innovation Performance*, Initiative of the European Union, online: <http://www.proinno-europe.eu/inno-metrics/page/innovation-union-scoreboard-2011> (last accessed on May 25, 2012).
- Romer, P., 1990. Endogenous technological change. *Journal of Political Economy* 98, S71–S102.
- Rousseau, S., Rousseau, R., 1997. Data analysis as a tool for constructing scientometric indicators, *Scientometrics* 40, 45-46.
- Ruiz, C.F., Bonilla, R., Chavarro, D., Orozco, L. A., Zarama, R., Polanco, X., 2009. Efficiency measurement of research groups using Data Envelopment Analysis and Bayesian networks. *Scientometrics* 83, 711-721.
- Saen, R.F., 2010. Restricting weights in supplier selection decisions in the presence of dual-role factors. *Applied Mathematical Modelling* 34, 2820-2830.
- Salter A.J., Martin, B.R., 2001. The economic benefits of publicly funded basic research: a critical review. *Research Policy* 30, 509-532.
- Schumpeter, J.A., 1934. *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle*. Harvard University Press, Cambridge, Mass.

- SCImago, 2007. SJR-SCImago Journal & Country Rank (last accessed on January 15, 2012 from <http://www.scimagojr.com>).
- Solow, R.M., 1956. A Contribution to the economic theory of economic growth. *Quarterly Journal of Economics* 70 (1), 65-94.
- Thomas, V.J., Sharma, S., Jain, S.K., 2011. Using patents and publications to assess R&D efficiency in the states of the USA. *World Patent Information* 33, 4-10.
- Vedung, E., (1998). Policy Instruments: Typologies and Theories, in: Bemelmans-Videc, M. L., Rist, R.C., Vedung, E. (Eds), *Carrots, Sticks & Sermons*. Transaction Publishers, New Brunswick and London, pp. 21-58.
- Von Hippel, E., Tyre, M.J., 1995. How learning by doing is done: problem identification in novel process equipment. *Research Policy* 24, 1-12.
- Wang, E.C., Huang, W., 2007. Relative efficiency of R&D activities: a cross-country study accounting for environmental factors in the DEA approach. *Research Policy* 36, 260-273.
- World Bank, 2012. Various datasets. Official website – <http://data.worldbank.org> (last accessed on January 18, 2012).
- World Economic Forum, 2010. The Global Competitiveness Report 2010–2011. World Economic Forum, Geneva, Switzerland (http://www3.weforum.org/docs/WEF_GlobalCompetitivenessReport_2010-11.pdf).
- World International Patent Office, 2011. Official website – <http://wipo.com> (last accessed on March 20, 2012).
- Yu, Z., 2005. Trade, Market Size, and Industrial Structure: Revisiting the Home-Market Effect. *The Canadian Journal of Economics* 38, 255-272.

Evaluation of Research, Technological Development and Innovation public policies in the European Union: a Bayesian network model and the Two-Linked-Steps Data Envelopment Analysis approach.

Abstract

The purpose of this work is to design a conceptual model through a Bayesian network (BN) to provide a useful tool for decision-making support in the evaluation of the relative technical efficiency of Research, Technological Development and Innovation (RTDI) public policies within a sustainable development framework. The efficiency of such policies in the 27 European Union (EU) Member States was assessed applying the two-linked-steps Data Envelopment Analysis (2LS-DEA) approach. Providing stakeholders with this conceptual model could reduce the inherent uncertainty in both RTDI and public decision-making processes. Thus, the main contributions of this study are: i) the presentation of a Bayesian Network (BN) as a conceptual model for decision-making support; ii) the evaluation of the model through the 2LS-DEA approach, which assesses the relative technical efficiency of the RTDI public policies of the 27 EU Member State in various socioeconomic scenarios embodied in the BN (the original dataset was previously transformed by a rule-base that emerged from the BN) and iii) to obtain final efficiency scores of the RTDI public policies in the EU. The results suggest that the BN along with the 2LS-DEA approach could be a useful methodology for decision-support in monitoring the efficiency of EU RTDI public policies as well as for future better resource allocation.

Keywords: *RTDI public policies; EU Strategies; technical efficiency; Bayesian network; sustainable development; two-linked-steps DEA.*

1. Introduction

The relevance of efficiency assessment in public policies like those of Research, Technological Development and Innovation (RTDI) is highlighted in many political agendas and official reports, i.e. the Budget proposed by the U.S.A. government for 2013 or the European Union (EU) 2020 Strategy (European Commission, 2010). This evaluation and its potential socio-economic impacts are even more relevant due to the worldwide economic crisis, which implies budget cuts.

However, even in times of economic downturn, it is fundamental to continue supporting Research and Development (R&D) as well as innovation activities¹ (Filippettia and Archibugi, 2011). Numerous studies and economic growth theories state that R&D –including technological change, knowledge diffusion and innovation capacity- play a crucial role in the development, economic growth and well-being of a country (Aghion and Howitt, 1992; Furman et al., 2002; Grossman and Helpman, 1991; Romer, 1990) and have positive impacts in competitiveness (Freeman, 2004; Schumpeter, 1934; Solow, 1956; Fagerberg et al., 2004). From an empirical point of view, many authors also highlight the relevance of both the direct and socioeconomic returns of R&D expenditure (Archibal and Pereira, 2003; Baumol, 2003; Jones and Williams, 1998).

According to the above statements, the main objectives of European RTDI public policies are to increase economic growth through sustainable development and to create a knowledge-based economy in the EU by supporting R&D. In addition, the EU still faces the challenge of reducing its existing technological gap with other advanced economies (the well-known “European Paradox”). This means that, in spite of the amount of investment in R&D, Europe is weak in transforming new knowledge into innovation (Dosi et al. 2009). EU strategies like the Lisbon Strategy in 2000, re-launched in Barcelona’s European Council in 2002, and the current “European 2020” Strategy (European Commission, 2010) reflect this trend.

Under this general framework, the EU has established R&D expenditure intensity targets relying on their impact on economic performance. Sustainable development also appears as an overarching objective of the EU since 1997. The EU Strategy for Sustainable Development aims “for the continuous improvement of quality of life for current and future generations” (European Commission, 2009). R&D and innovation activities also appear as being fundamental to achieve this objective.

¹ It is important to stress that in this study the term “innovation” is understood in a narrow sense that does not include R&D activities because, while R&D plays a vital role in the innovation process, much innovation activity is not R&D-based. Innovation policies can include others like patent regulation, R&D support, etc.

R&D investments are very different from other investments. These differences are (Martin and Scott, 2000): i) indivisibility; ii) economies of scale; iii) its inherent uncertainty (it is not sure that R&D efforts will result in new knowledge and innovation); iv) asymmetric information; v) market failures; and, finally, vi) limited appropriateness of the results. This study is focused on the public side of R&D expenditure to evaluate the efficiency of such expenditure. At this point, it must be stressed that, although R&D public funds are mainly spent on basic research in universities and research institutes, this public support also has a relevant impact on R&D private expenditure (Cohen et al., 2002; Salter and Martin, 2001).

The main reason for assessing the impacts of RDTI public policies is to evaluate their contribution to general public objectives: economic growth, health-care and/or education, thus avoiding the draining of scarce resources dedicated to ineffective and inefficient programs. This evaluation allows the creation of a better-informed society that increasingly demands enhanced public accountability. This issue is still a source of great debate because there are different instruments and indicators to carry it out, and there have been as well different attempts to provide classification schemes as well as for exploring the character of innovation studies as a scientific field (Vedung, 1998; Fagerberg and Verspagen, 2009; Martin, 2012).

The methodologies for assessing public R&D impacts go from econometric studies, general Equilibrium Models (which rely on a large number of assumptions) and the use of composite indicators (Grupp and Schubert, 2010; Freeman and Soete, 2009) to survey-based studies, case studies and, more recently, non-linear classifiers for R&D performance (De la Paz-Marín et al., 2012). Nevertheless, few studies consider public R&D investment from a macroeconomic perspective (David and Hall, 2000). According to the latter, the design of a conceptual model for identifying the key variables involved in R&D and innovation processes and their relationships can be a useful decision-support tool for policy-makers and analysts.

The drawback for designing this conceptual model is the existence of multiple variables and the study of their interactions in an uncertain socio-economic environment. Moreover, classical approaches to modeling could find it quite difficult to capture the whole structure of the domain (Gibert et al., 2010), particularly when information is incomplete. Classic algebraic formalisms disregard the knowledge background necessary to design and develop a conceptual model that includes both quantitative and qualitative aspects of causal socio-economic phenomena (Pearl, 2000).

There are a number of methodologies to design conceptual models like system analysis, causal modeling, etc. In social sciences, causal modeling can be very interesting due to its

ability to combine cause-effect information with statistical data to provide a quantitative – stochastic- assessment of the phenomena under study (Anderson and Vastag, 2004).

The main approach for representing a causal model is a Bayesian Network (BN) (Pearl 2000, Darwiche 2010). Among a variety of fields and in issues related to R&D and innovation, BNs have also been used to assess innovation collaboration processes (Kim and Park, 2008), webometrics (Lamirel et al., 2004) and to map the internal configuration of the efficiency results of research groups (Ruiz et al., 2009). Nevertheless, to our knowledge they are not used as conceptual models for modeling the R&D and innovation processes and as a tool for the evaluation of RTDI public policies.

On the other hand, Data Envelopment Analysis (DEA) is a non-parametric technique for assessing the technical relative efficiency of a set of comparable Decision-Making Units (DMUs), which use multiple inputs to produce multiple outputs (Cooper et al., 2004). This analysis identifies the most efficient DMUs (Pareto-optimal) that are then used as the standard for comparison with the rest of the DMUs –which are inefficient– for analysing potential improvements for them. Rousseau and Rousseau (1997) were the first using a DEA approach to assess the relative efficiency of R&D processes in 18 developed countries. DEA has also been used to approximate the efficiency of universities (Bonaccorsi et al. 2006), research centers (Korhonen et al. 2001), research groups (Guan and Wang, 2004) and R&D activities (Wang and Huang, 2007).

In this study, the relative technical efficiency of RTDI public policies in the 27 EU Member States was assessed through a new approach called Two-Linked-Steps DEA (2LS-DEA) based on a BN that identifies the inputs and outputs. By analyzing the preliminary 2LS-DEA results, the BN was improved in an iterative way and final efficiency scores for the EU Member States were obtained. To achieve this, various scenarios (combinations of specific inputs and outputs) were designed to analyze the problem. The corresponding inputs and outputs for each scenario were chosen according to the causal relationships previously defined in the BN. Original input and output values were interpreted by a knowledge base (rule-base) that stores the explicit expert knowledge used to design input/output causal relationships. This interpretation process was carried out automatically by a linear monotone algebraic transformation.

The aim of the present research is twofold: i) to design a conceptual model (Bayesian Network) that identifies the key variables and their relationships for describing R&D and innovation processes and their socioeconomic impacts in a sustainable development framework and ii) the use of the 2LS-DEA in an iterative way to evaluate the model and to

obtain the relative technical efficiency of the 27 RTDI public policies of the 27 EU Member States.

In this way, after introductory section 1, the underlying assumptions involved in the BN, the variables and its structure are presented in section 2. In section 3, the methodology and the employed dataset are described in order to assess the technical relative efficiency of RTDI public policies in the 27 EU countries employing a Two-Linked-Steps DEA (2LS-DEA) approach. Empirical results of the application of 2LS-DEA and the evaluation of the model by these results are presented in section 4. Finally, conclusions as well as future lines of research are outlined in section 5.

2. The Bayesian Network: a conceptual model

A Bayesian Network (BN) can be used for modeling complex and stochastic domains under uncertainty (Pearl 2000, Darwiche, 2010). They are usually shown as graphs, which typically have the form of a network diagram (Jensen, 2001), and incorporate variables (nodes) and their probabilistic relationships (arcs). BNs capture uncertainties in terms of probabilities to perform reasoning and inference, motivated by the need to make rational decisions (Darwiche, 2010). In general, a BN has two parts. The first part is qualitative, where nodes represent variables, and directed arcs describe conditional causal relationships. The second part is quantitative, where causal relationships are expressed in terms of joint conditional-probability distributions (Nadkarni and Shenoy, 2001). This structure lets BNs be combined with analytic decision tools like, for example, operational models (Jensen, 2001).

The most common network diagram for a BN is the Direct Acyclic Graph (DAG) in which each arrow/arc/connection that joins two variables depicts causal dependence, and two-headed arrows and feedback loops are not permitted (Pearl, 2000). A DAG could be solved using statistical and artificial intelligence-based techniques.

Some of the disadvantages of a BN include: i) the difficulty in obtaining probability information to make predictions, ii) it can be hard to handle many random and continuous variables and, finally, iii) it may not contain the appropriate variables if the prior explicit knowledge about the domain is wrong. Nevertheless, BNs are excellent conceptual models for forecasting because they can incorporate simulation techniques and fuzzy inference engines to evaluate imprecise causal relationships.

2.1. The design of the BN: underlying assumptions

A BN can be considered as a conceptual model for identifying the key variables and their causal relationships in R&D and innovation processes. Thus, the basic principles of the Endogenous Growth Theory, where economic growth and development are driven by technological change were used to design our preliminary BN. According to this theoretical framework, economic growth results from the R&D efforts of public (government and higher education) and private agents (business sector). This BN also integrates the principles of the Linear Model of Innovation, the Schumpeterian tradition and the Quintuple Helix Model. According to that, the structure of the BN (variables and relationships) are based on four assumptions:

Assumption 1. The BN was based on the Endogenous Growth Theory. It incorporates technological change as an endogenous factor to the economic growth process. It initially emphasizes the role of knowledge as an input of production (Romer, 1990). Since then, many empirical studies have explored how and to what extent innovation (as an output of R&D) might contribute to economic growth (Brinkman and Brinkman, 2001).

Assumption 2. The Linear Model of Innovation was selected in our research because policy makers understand that R&D investments generate innovations in a linear way and, consequently, official databases are designed following this model (Godin, 2006; Balconi et al., 2010). There are different and more complex models for describing innovation process like the chain-linked model (Kline and Rosenberg, 1986) and the cycle model (Padmore, 1998) that establish that the key element of the innovation process was the fact that research, development and innovation are linked in a non-linear and interactive way.

Assumption 3. Schumpeter's work (1934) established five types of innovations: a new thing; a new process for making a familiar thing; a new market to which to sell; a new source of supply or raw material; and industrial reorganization. Moreover, he analysed the influence that firms' size and market structure had on innovation.

Assumption 4. The Triple Helix Model (Etzkowitz and Leydesdorff, 1998) was considered to be a model that engages synergies from three main R&D stakeholders: universities, which develop academic-based research and development activities; industries, which develop commercial activities as well as R&D activities; and, finally, governments as the policy makers. The Quadruple Helix adds media-based and culture-based public and civil society as the fourth helix and, finally, the Quintuple Helix innovation model incorporates the "natural environments of society" helix as well (Carayannis and Campbell, 2010). The latter represents socio-ecological transition from the knowledge economy.

2.2. Main variables (nodes) and structure (relationships) of the BN

Relevant variables and relationships for R&D and innovation process were identified and defined (Appendixes A and B) as well as represented in the BN (Figure 1). The BN shown in Figure 1 is the last version after various improvements that incorporate the elicited implicit knowledge obtained from experts once 2LS-DEA results (efficiency scores) were statistically analyzed.

In the literature on the measurement of the innovation, certain authors (Chiesa et al., 1996; Hansen, 2001) propose some audit tools based on observable indicators to enable decision makers to collect data on innovation. The Organization for Economic Cooperation and Development (OECD) has elaborated some guidelines for the collection of data on R&D and innovation in the Frascati Manual (OECD, 2002), the Oslo Manual (OECD and Eurostat, 2005) and Main Science and Technology Indicators (OECD, 2010), which are widely used. The construction of our model was based on these manuals and the related literature² from which we have selected the variables.

(Figure 1: Bayesian Network conceptual model, about here)

The inputs and outputs to evaluate the efficiency of the RTDI EU policies in various socioeconomic scenarios were selected from the BN and algebraically interpreted according to their relationships. The BN understands a time lag of two years; that is that the effects of investments in R&D carried out in the year $t-2$ will be produced in the year t (Greif, 1985).

R&D expenditures and R&D personnel are the main indicators for the inputs of R&D and innovation activities. The former reflects the degree of relevance that agents place on R&D and their capability for generating innovation and the latter assesses the number of persons that are engaged in fundamental research, application research and experimental development (OECD, 2010). R&D expenditures and R&D personnel include both the public sector ($ERDFUN_{t-2}$ for R&D expenditures and $PUBPER_{t-2}$ for R&D personnel as a percentage of total employment) and the private sector (R&D business sector expenditure, $ERDPRI_{t-2}$, and R&D personnel, $PERPRI_{t-2}$). Only the public sector was considered to assess EU RTDI policies ($ERDFUN_{t-2}$ and $PUBPER_{t-2}$).

The number of scientific publications ($SPUBLI_t$) and patent applications ($TPATRE_t$) (OECD, 2010) are the most common variables for the outputs of R&D and innovation

² The works of Godin (2002, 2006, 2007) deserve special mention in this research.

activities. They are influenced by $ERDFUN_{t-2}$, $PUBPER_{t-2}$ and by the infrastructures of the country ($INFRAE_t$ machinery and equipment, communication infrastructures, vehicles, dwellings and other buildings, including laboratories and other scientific infrastructures). Some studies highlight that $TPATRE_t$ can be highly influenced by $SPUBLI_t$ (Meyer, 2000).

Following the structure of the BN, $ERDFUN_{t-2}$ is influenced by the existing EU strategic guidelines ($EUSTRAT_{t-2}$) and $PUBPER_{t-2}$ depends on: i) the level of mobility of researchers and PhD students ($MOBILI_{t-2}$) and ii) the attractiveness of the market for researchers due to the wage level of the countries ($MARKRE_{t-2}$).

For the private sector, $ERDPRI_{t-2}$ (R&D private expenditure) is influenced by public support ($ERDFUN_{t-2}$) in the form of subsidies, tax incentives, etc., the size of the firms ($SIZENT_{t-2}$) and the structure of the private sector for the corresponding country ($SECEST_{t-2}$) (Yu, 2005). The personnel involved in R&D activities in the private sector mainly come from $PUBPER_{t-2}$ and depend on $ERDPRI_{t-2}$.

According to Schumpeter's types of innovation (Schumpeter, 1934), the BN embodies the difference between technological innovations measured by patents ($TPATRE_t$) and non-technological innovations represented by proxy indicators like community trademarks ($TRADEM_t$), small and medium enterprises introducing marketing/organizational innovations ($INNSME_t$) and sales of new to markets and new to firm innovations as a percentage of the turnover ($INNEW_t$). Technological advance would not be possible without the development of new ways of organization to guide and support R&D and enable firms to obtain profits from this kind of investment (Nelson, 1996).

$TPATRE_t$ and indicators for non-technological innovations ($TRADEM_t$, $INNEW_t$ and $INNSME_t$) depend on a variety of variables like $INTECAP_t$ (intellectual capital: the difference between tangible physical and financial assets and human, structural and relational capital market values; Bontis et al., 1999), $SPILLOV_t$ (spillovers: some of the economic benefits of R&D activities go to economic agents other than the ones undertaking the research; Block, 2012) and social and business networks $SBNETW_t$ (connections through individuals and organizations; Pavitt, 1998; Parker, 2008). These variables emphasize that innovation can also take place without a systematic research and development process. Other influencing variables were: learning-by-doing ($LEARDO_t$) [Von Hippel and Tyre, 1995], benchmarking ($BENCM_t$) or "the continuous process of identifying, understanding and adapting practice and processes that will lead to better performance" (Auluck, 2002; p. 111) and sociocultural aspects like the existence (or not) of an entrepreneurship culture in the country, $ENTREP_t$ (Braunerhjelm, 2008).

As mentioned in section 1, the main reason for assessing the impacts of RDTI policy is its contribution to public objectives as economic growth, health outcomes or education (Boulanger and Bréchet, 2005), to justify not investing in other alternatives while creating a better-informed society to satisfy the increasing demand for public accountability.

Following this underlying idea, the resource allocation in R&D and innovation ($ERDFUN_{t-2}$, $PUBPER_{t-2}$, $ERDPRI_{t-2}$ and $PERPRI_{t-2}$), which leads to patents and scientific publications, would be desirable to positively affect global society welfare ($WELFAR_t$).

Since productivity growth plays a main role in increasing the Gross Domestic Product (GDP), it is important to point out the main factors underlying the growth of such productivity. The infrastructures of the country ($INFRAE_t$), employees' motivation ($MOTIVA_t$), the labour climate ($LABCLI_t$) and technological change ($TECCHA_t$) play a very important role in our BN to explain the Real Labour Productivity Growth per hour worked ($RLPGH_t$). Furthermore, the level of technological change ($TECCHA_t$) depends on technological innovation ($TPATRE_t$), the total R&D expenditure as % of GDP ($GERDTO_t$, which comprises the expenditure from public, private, foreign and non-profit sectors in the year t); and the technological balance of payments ($TECACQ_t$) that includes various forms of technological acquisition, such as the purchase of capital goods or technological services, and including the licensing of intellectual property. Other variables that strongly influence $RLPGH_t$ are those concerning education: $PHD06_t$ (number of Ph.D.), $ELEAV_t$ (early leavers) and $LLEARN_t$ (life-long learning).

As the Global Competitiveness Report of the World Economic Forum (2010) set out, productivity determines the return rates of the investments (physical, human, and technological) made in an economy. The $RLPGH_t$ influences the general competitiveness of a country and, in turn, sets the sustainable level of prosperity that can be reached by an economy. In other words, economies that are more competitive tend to produce higher levels of income for their citizens ($GDPGRO_t$). Conversely, a higher level of productivity per hour worked could lead to higher levels of unemployment.

Economic impacts would measure the extent to which traditional outputs (scientific publications and patents) indirectly influence and address: Real Labor Productivity Growth per hour worked ($RLPGH_t$), the corresponding GDP growth ($GDPGRO_t$) as well as their implications for High-Technology Exports ($HITECH_t$) and the reduction in the unemployment rate ($UNEMPL_t$).

The institutional framework of a country ($INSTFR_t$) also plays a relevant role in a society's welfare (World Economic Forum, 2010). The importance of the $INSTFR_t$ has

become even more significant during the economic crisis, given the direct role played by the States in the economy. This legal framework (*INSTFR_t*) has an important influence on the variable *TPATRE_t* (intellectual rights must be protected by the law to be effective), affects the image of a country abroad (*COUIM_t*), foreign direct investment (*FORINV_t*) and market efficiency (*MARKEF_t*). *COUIM_t* can also have a relevant influence on the latter variable.

The BN shows the effects on competitiveness (*COMPET_t*) that may be achieved due to productivity growth (*RLPGH_t*), technological (*TPATRE_t*) and non-technological innovation indicators (*TRADEM_t*, *INNEW_t* and *INNSME_t*) and market efficiency (*MARKEF_t*).

Finally, all these variables and their causal relationships converge into the most popular measure for determining the levels of social welfare (*WELFAR_t*), Gross Domestic Product Growth (*GDPGRO_t*).

Due to the complexity of the BN, the rest of the variables are described in section 3.2 when describing the scenarios for 2LS-DEA.

3. Evaluation of the BN: the relative technical efficiency of EU RTDI public policies

3.1. Data Envelopment Analysis (DEA)

Data Envelopment Analysis (DEA) is a non-parametric, mathematical programming-based technique for measuring and evaluating the relative efficiency of decision-making units (DMUs), which use multiple inputs to produce multiple outputs making it appropriate for comparing the efficiency of R&D in terms of multiple innovative outputs. Charnes, Cooper and Rhodes promoted this technique but its roots may be found as early as 1957 in Farrell's seminal work (Cooper et al., 2004).

In relation to the specific model for running DEA, the variable returns-to-scale (or BCC model) was selected because R&D activities can exhibit increasing or decreasing returns to scale (Bound et al., 1984). It is also implicitly assumed that the emphasis on public policies is more on what is produced and based on given resources. Therefore, this study adopted the output-oriented BCC model. The free imputation of input-output values was selected for the identification of inefficiencies (Saen, 2010).

Considering a set of n DMUs ($DMU_j; j=1,2, \dots, n$), that each DMU_j consumes m inputs (x_{ij}) in different amounts ($x_{ij}; i=1,2 \dots, m$) and that produces s outputs (y_{rj}) in different quantities ($y_{rj}; r=1,2, \dots, s$), the relative technical efficiency of a DMU (output-oriented) can be evaluated by the envelopment form using the following linear model (Cooper et al., 2004):

$$\max \phi + \varepsilon \left(\sum_{i=1}^m s_i^- + \sum_{i=1}^m s_i^+ \right)$$

Subject to:

$$\begin{aligned}
\sum_{j=1}^m x_{ij} \lambda_j + s_i^- &= x_{i0} & i=1,2, \dots, m \\
\sum_{j=1}^n y_{rj} \lambda_j + s_r^+ &= \phi y_{r0} & r=1,2, \dots, s \\
\lambda &\geq 0 \\
\sum_{j=1}^n \lambda &= 1
\end{aligned} \tag{1}$$

where ϕ represents the efficiency score, x_{ij} and y_{rj} are the i th input and the r th output, respectively, of the j th DMU. Similarly, x_{i0} and y_{r0} stand for the inputs and outputs, respectively, of the DMU under evaluation (DMU₀). In addition, λ_j denotes nonnegative scalars and s_i^- and s_r^+ are input and output slacks, respectively.

To apply DEA effectively requires (Cooper et al., 2007) that:

$$n \geq \max \{m \times s, 3 \times (m + s)\} \tag{2}$$

because the whole technique is based on the comparison of each DMU with all the remaining ones, thus a considerably large set of DMUs is necessary to be meaningful.

3.2. Structure of DEA scenarios: inputs and outputs

The evaluation of the BN design was performed by assessing the relative technical efficiency RTDI public policies in EU Member States analyzing selected scenarios (input and output combinations). All inputs and outputs comes from the BN (Appendix A) but not all variables in the BN were employed as inputs and outputs of the selected socioeconomic scenarios (Appendix B). The scenarios were³:

1. Traditional scenario. The relative technical efficiency of RTDI public policies in the EU was evaluated taking into consideration two inputs (R&D public expenditure, $ERDFUN_{t-2}$,

³ All variables in the BN are circled individually (nodes) except the ones included in the above mentioned scenarios for better visualization in Figure 1.

and R&D public personnel, $PUBPER_{t-2}$) and two outputs (scientific publications, $SPUBLI_t$, and patents applications, $TPATRE_t$).

The number of scientific publications is a common indicator of academic productivity but it is not the only one. Project reports, monographs and other literature are equally important (Thomas et al., 2011). However, as data on such output are not available, we use $SPUBLI_t$. The number of patents is the most widely accepted indicator for innovation due to the lack of more consistent indicators (Griliches, 1990). This also shows some disadvantages: not all inventions are technically patentable, not all inventions are patented and, finally, firms have different propensities to patent in their domestic markets and in foreign countries.

2. Socioeconomic scenarios. These scenarios are combinations of inputs and outputs to describe different faces of a sustainable socioeconomic development. All of them have the same inputs and different outputs. The inputs are the number of scientific publications $SPUBLI_t$ and number of patents $TPATRE_t$, the outputs of the traditional scenario. The aim is to analyse the contribution of these inputs to the performance of each DMU (RTDI public policies in the EU) in Health, Environment, Information Society, Education, Economic and Innovation scenarios. In order to carry out 2LS-DEA, only a selected percentage of the original values of $TPATRE_t$ and $SPUBLI_t$ was considered. A panel of experts determined this percentage because only the public sector is analyzed. If a country is efficient in the traditional scenario, it is expected to be efficient in socioeconomic ones.

Health. Improvements in the health sector due to R&D and innovation investments are well-known and their expenditures are usually well received by society. Three outputs were used to analyse this scenario: $AGEDEP_t$ (age dependency ratio), $MORRAT_t$ (mortality rate) and $LIFEXP_t$ (life expectancy).

Environment. Obviously, without a sustainable environment, life on earth would be in danger. The variables (outputs) selected to measure the efficiency in this scenario were $ENINT_t$ (energy intensity of the economy), $CARBDI_t$ (carbon dioxide damage) and $GASEM_t$ (total greenhouse gas emissions).

Information Society. The most relevant sector in which R&D outputs had an exponential influence is Information and Communication Technologies (ICT). With the increasing relative weight of the services sector induced by the knowledge economy in developed countries, ICT are now even more incalculably important. The indicators for this scenario were $BROSUS_t$ (fixed broadband Internet subscribers), $INTUSE_t$ (internet users) and $MOBILE_t$ (cellular mobile subscriptions).

Education. The European population has to acquire a range of skills while continuously adapting new ones. Knowledge and technology diffusion relevance are included in this scenario, which plays a very important role in employment. The outputs were: $PHD06_t$ (number of doctorates), $ELEAV_t$ (early leavers from education) and $LLEARN_t$ (life-long learning).

Economic. This scenario focuses its attention on economic prosperity, a key objective reflected in EU strategies. The selected outputs were: Gross Domestic Product growth ($GDPGRO_t$), real labour productivity growth per hour worked ($RLPGH_t$), the unemployment rate ($UNEMPL_t$) and High-Technology Exports ($HITECX_t$).

Innovation. This scenario reflects the concept of innovation defined in the Oslo Manual (OECD and Eurostat, 2005). The selected outputs are community trademarks ($TRADEM_t$), small and medium enterprises introducing marketing/organizational innovations ($INNSME_t$) and sales of new to markets and new to firm innovations as a percentage of the turnover ($INNEW_t$).

3.3. Two-Linked-Steps DEA approach (2LS-DEA)

The efficiency of EU RTDI public policies was not only assessed using the traditional scenario –step 1–, but also by analyzing six socioeconomic scenarios –step 2, linked with the first. Although the BN causal model represents the R&D and innovation process as a whole, this research considers only the public sector. For this reason, in the traditional scenario, only 10% of the value of total patent applications ($TPATRE_t$) comes from $PUBPER_{t-2}$ and $ERDFUN_{t-2}$ and only 90% of the value of SPUBLIt comes from $PUBPER_{t-2}$ and $ERDFUN_{t-2}$. In Appendix A (% column) it is specified the corresponding percentage of the original output value in each socioeconomic scenario.

In step one of the 2LS-DEA the traditional scenario is evaluated and in step 2 the socioeconomic scenarios are assessed, where traditional outputs turn into the inputs in all of them in a way that both steps are “linked”.

3.4. Dataset

The 27 RTDI public policies of the corresponding EU countries were selected as DMUs. Inputs and outputs were quantitative data (continuous) according to the conceptual model (BN) discussed in section 2. Data were collected from various official databases: Eurostat’s official website for all variables, except the World International Patent Organization (WIPO, 2011) for patent applications, SCImago Journal & Country Rank in the case of the number of scientific publications per country (SCImago, 2007) and the World Bank Database (World

Bank, 2012) for age dependency ratio ($AGEDEP_t$), broadband subscribers ($BROSUS_t$), high-tech exports ($HITECEX_t$) and mobile cellular subscriptions ($MOBILE_t$). Variable values were considered per 1,000 inhabitants when necessary. Table 1 illustrates descriptive statistics for all inputs and outputs.

(Table 1 about here)

The main sources for defining qualitative variables and the intervals of quantitative ones were the Global Competitiveness Index (WEF, 2010), ERAWATCH profiles of European Countries in the European Research Area⁴ and an expert panel's opinion. They were used for interpreting the values of variables for building the rule-base in the next section 3.5.

Appendix A shows the information of quantitative variables scenario by scenario and Appendix B the qualitative variables represented in the BN model.

The reference period was 2008 due to the availability of data in the Community Innovation Survey of Eurostat (CIS) that is carried out every two years (Eurostat, 2012). Meanwhile, considering the time lag for the innovation process, data acquisition was 2006 for inputs in the traditional scenario. As mentioned in section 2.2, this means that if a country promotes R&D investments in 2006 (or year $t-2$), outputs are not expected to increase until at least two years later (2008 or year t).

3.5. Expert-based input and output transformation

In order to design BCC-DEA models, the original values of quantitative variables in the dataset were transformed by applying the expert-based rule-base that emerges from the BN. Classical DEA models interpret that the greater the outputs, the greater the efficiency of a specific DMU (if input values are constant) and the lower the inputs, the greater the efficiency (if output values are constant). Output-oriented DEA tries to maximize the output values once the input values have been set. However, in some situations some inputs need to be increased and some outputs need to be decreased to improve the efficiency of a DMU (non-standard inputs and outputs). In these cases (for example, $UNEMPL_t$ and $AGEDEP_t$), we performed a linear monotone increasing/decreasing transformation on their original values based on the expert-based rule-base.

⁴ ERAWATCH is the European Commission platform on research policies and systems in the EU and beyond. (<http://erawatch.jrc.ec.europa.eu/>)

The rule-base is a set of *if-then* rules that emerged from the causal relationships in the BN model and the characteristics of the variables (Table 2). Transformed variable values were used to design final DEA models to assess the relative technical efficiency of EU RTDI public policies incorporating expert knowledge.

(Table 2 about here)

For example, a simple linear transformation rule transforms $GDPGRO_t$ according to the influence of $RLPGH_t$. If $GDPGRO_t \leq 1$ (*low*) and $RLPGH_t > 3$ (*high*), then the original value of $GDPGRO_t$ is penalised dividing its original value by two (R15). The reason for this rule relies on the fact that high levels of productivity must lead to increasing rates of GDP growth; when $GDPGRO_t$ is *low* and $RLPGH_t$ is *high*, the country analyzed is not achieving this goal. On the other hand, if $GDPGRO_t > 3$ (*high*) and $RLPGH_t \leq 1$ (*low*) then the original value of $GDPGRO_t$ is divided by two (R14 in table 2) to make DEA understand that this situation means a very high country performance. A special rule emerges in the case of Eastern economies due to their very high GDP growth rate that mainly represents the transition economies of the New Member States. As new economies, they involve a high rate of growth in GDP and in Real Labor Productivity growth per hour worked ($RLPGH_t$) due to the abundance of natural resources, low population and a high domestic demand but a low rate of employment after a period of socialism, among other reasons. To make DEA understand this, the real value of $GDPGRO_t$ is transformed by the rule R16:

$$\text{IF } GDPGRO_t > 5 \text{ THEN } GDPGRO_t' = GDPGRO_t / 2 \quad (3)$$

It is also interesting to highlight the transformation of $ERDFUN_{t-2}$ accomplished with R1, R2 and R3. $ERDFUN_{t-2}$ depends on European Strategy ($EUSTRAT_{t-2}$ node in the BN). Due to Lisbon Strategy's targets applicable in the period under study (to reach at least 3% of GDP of R&D effort and that 60% of total R&D expenditure must be funded by the private sector), $ERDFUN_{t-2}$ was transformed according to R1, R2 and R3 rules:

$$\text{IF } ERDFUN_{t-2} / GERD_{t-2} \leq 0.4 \text{ and } GERD_{t-2} > 3 \text{ THEN } ERDFUN_{t-2}' = ERDFUN_{t-2} * 2 \quad (4)$$

$$\text{IF } ERDFUN_{t-2} / GERD_{t-2} > 0.55 \text{ AND } GERD_{t-2} < 2 \text{ THEN } ERDFUN_{t-2}' = ERDFUN_{t-2} / 2 \quad (5)$$

$$\text{IF } ERDFUN_{t-2} \leq 0.2 \text{ THEN } ERDFUN_{t-2}' = ERDFUN_{t-2} / 2 \quad (6)$$

4. Empirical Results

The 2LS-DEA approach provided a ranking of 27 EU RTDI public policies based on the relative technical efficiencies obtained in each scenario (Table 3). An efficiency score equal to 100 means that the RTDI public policy of a country is efficient (relative efficiency); values lower than 100 mean different grades of inefficiency.

(Table 3 about here).

In order to summarize and analyze the efficiency scores, a *k*-means clustering was carried out and four clusters were chosen:

Cluster 1 includes Bulgaria, Estonia, France, Italy, Latvia, Poland and Romania. Their RTDI policies are the least efficient in the Traditional scenario. They also appear as the least efficient in Innovation and Information society scenarios and they show low efficiency scores in the Education scenario.

France and Italy belong to this cluster due to the poor relationship between their public efforts and the traditional inputs (patents and scientific publications) and in the innovation and education scenarios.

Cluster 2 includes the most efficient RTDI policies. It groups Austria, Belgium, Denmark, Finland, Germany, Ireland, Lithuania, Luxembourg, Netherlands, Sweden and the United Kingdom. Sweden is efficient in all the scenarios under study followed by Belgium (six scenarios) and Finland (five scenarios). Ireland and Luxembourg rank among the five top in all but three scenarios and almost reach efficiency in six scenarios. Lithuania appears in this group mainly due to its efficiency in the Traditional and Economic scenarios. In spite of the relatively low amount of resources devoted to R&D by its public sector, it shows a high number of scientific publications.

Cluster 3 includes Cyprus, Czech Republic, Hungary, Slovakia, Slovenia and Spain. They show low efficiencies in the Traditional and Environment scenarios but they perform very well in the Innovation scenario.

Finally, **cluster 4** includes Greece, Malta and Portugal RDTI public policies. It groups the least efficient RTDI policies. They show low efficiencies in the Traditional scenario and, this group reaches the lowest efficiencies in the Economic, Education and Health scenarios. Nevertheless, they are efficient in the Information Society scenario (very high number of mobiles users, especially in Greece).

In order to discriminate between the most efficient countries, a complementary measure was used that involved the number of times that those countries were peers for others in the DEA assessment. Sweden, Belgium, Finland and Ireland RDTI public policies were the most frequently selected by inefficient ones as benchmarks.

The 2LS-DEA results were compared to the Innovation Union Scoreboard (IUS) scores per dimensions (European Commission, 2010) that were comparable with traditional, innovation and economic scenarios. The IUS report uses 24 variables and develops a composite indicator (Summary Innovation Index). Mann–Whitney U test results showed some relevant discrepancies between them that could explain the utility of the 2LS-DEA approach, which incorporates expert knowledge by means of the rules-base. For example, Lithuania's efficiency score differs from IUS classification because some variable values were transformed through the rule-base. Thus, in the Traditional scenario, $ERDFUN_{t-2}$ value was penalized dividing by two and two multiplies the original value of $PUBPER_{t-2}$ as R1, R2 and R3 rules were applied to the original dataset (see Table 2). This country is efficient in the Traditional scenario but it is not in socioeconomic ones, except in Economic. Thus, it is not verified the initial hypothesis that if a country is efficient in the Traditional scenario has to be efficient in socioeconomic ones.

Another country where our efficiency result differs from IUS scores is France. In this case, $SPUBLI_t$ was penalized by R11 rule, mainly because the number of scientific publications was *low* (<1.5 per 1,000 inhabitants) with respect to public efforts ($ERDFUN_{t-2}$ and $PUBPER_{t-2}$).

Other RTDI public policies are efficient in a large number (more or equal than three) of socioeconomic scenarios but cannot be considered efficient in the Traditional scenario. This group includes countries like France, Ireland, Italy, Luxembourg or Slovenia.

To conclude, our proposal is not only to classify the countries by means of the resulting scores but also to evaluate the technical efficiency of RTDI public policies. IUS includes public and private innovation variables so it would not be possible to distinguish the effect of specific public strategies.

5. Conclusions

Evaluation of the relative technical efficiency of RTDI public policies is an important issue for policy-makers. The knowledge-based economy and sustainable growth are concepts strongly related to R&D activities and their specific policies. Thus, the underlying idea is that the more efficient the countries are in the employment of corresponding R&D inputs, the

more efficient they are supposed to be in the socioeconomic scenarios selected in this study. If a specific country does not employ its R&D outputs efficiently as inputs for other socioeconomic scenarios (health, education, economic development, etc.), other priorities and social alternatives arise and take precedence over increasing expenditure for R&D. Improvements in such efficiency can lead to a growing competitiveness, but decision-making is not an easy task.

The proposed BN conceptual model and the Two-Linked-Steps DEA (2LS-DEA) approach address the main challenges for assessing public R&D impacts on society: i) the identification of the relationships between R&D inputs, outputs and their socioeconomic impacts; ii) the existence of direct or unidirectional relationships between inputs and outputs (our BN only includes direct causal relationships); iii) the assessment of what proportion of social benefits should be attributed to R&D inputs; iv) the effect of time when measuring those impacts (time lag) and, finally, v) the use of appropriate indicators (selection of variables).

The crucial points in the BN design were: i) interaction with experts –prior expert knowledge acquisition, ii) data pre-processing with the application of a rule-base and, finally, iii) the interpretation of the results of the 2LS-DEA. Then, the conceptual model (BN) and the rule-base were improved iteratively by eliciting implicit knowledge and until no inconsistencies were found. Finally, the 2LS-DEA assessed the (final) relative technical efficiency of RTDI public policies in the 27 European Union Member States. Thus, the 2LS-DEA approach was used both to improve the design of the BN (to provide evidence of its design as a conceptual model) and for assessing such relative technical efficiency.

Some of the abilities of this methodology are to transfer expert knowledge to data analysis and to elicit implicit knowledge, as it is an operational tool for providing fundamental information for decision-making processes. In this way, the results do not depend on subjective-weight assumptions and arbitrary selection of inputs and outputs but on the incorporation of expert knowledge into the dataset by means of the rule-base that was built based on the BN.

Other methodologies study R&D and innovation performance or, more widely, the competitiveness and the state of the knowledge-based economy of countries yielding comparable results. However, the aforementioned multidimensional approaches ignore causal relationships between variables and cannot interpret special combinations of variables (scenarios) that result from specific public policies (sometimes inefficient).

As regards the results of the 2LS-DEA approach, and as to be expected given the level of economic development in the 27 EU member States, countries like Sweden, Finland, Luxembourg and Denmark were at the top of multidimensional efficiency rankings. The least efficient RTDI public policies were in Bulgaria (it was not efficient in any scenario), Estonia, Latvia and Romania.

The diverse socioeconomic efficiency scenarios were presented to contribute to the discussion on European policies. The EU2020 Strategy, the European Research Area (ERA) or the inclusion of related indicators in the core of the Eurostat reveal the relevance of achieving sustainable development for the European Union. Since most of the scenarios examined in this study are related to public policy variables (education, health, environment, infrastructures, etc.), it can be said that our empirical findings should be of interest to policy makers for monitoring the realization of sustainable and smart European Strategies.

On the other hand, the model can provide main stakeholders with the dimensions that need improvements, which can guide a political rationale to increase or decrease the amount of inputs devoted to R&D and innovation funds. With the BN conceptual model and the 2LS-DEA approach, it is possible to identify the most efficient RDTI public policies in the socioeconomic scenarios designed, which opens a door for benchmarking.

Obviously, the resulting BN would be different if other assumptions were adopted. Nevertheless, the main purpose of this research was achieved by its evaluation and the demonstration of its potential for the analysis of RTDI public policy performances in the European Union, by using a 2LS-DEA approach in both cases.

The model can be easily adjusted in terms of the variables selected and the rule-base structure. These modifications can be based on other theoretical frameworks, as well as on other perspectives from different stakeholders. Additionally, this methodological approach can perform future inferences by incorporating a Monte Carlo simulation engine. This procedure could evaluate the effects of alternative investments in RDTI public policies and study their effects on selected scenarios.

Finally, this study could be addressed to compare both the results of the period employed (2006-2008), before the financial and economic crisis, and the next period (from 2009 to the more recent available data) to analyze the effects of the economic crisis on efficiency scores, all of these considerations being a vast field of future research.

Acknowledgements

This work was partially subsidized by the Spanish Inter-Ministerial Commission of Science and Technology under Project TIN2011-22794, the European Regional Development fund, and the “Junta de Andalucía” (Spain), under Project P2011-TIC-7508.

Among others, we would like to thank the expert opinions and comments of Prof. Gabriel Pérez-Alcalá (Rector at the Loyola-Andalusia University, Spain); Prof. Lorenzo Salas-Morera (Director of Academic Organization and Degrees, Córdoba University, Spain); Prof. Fernando Lara-De Vicente (Vice Dean of Academic Planning, Faculty of Labor Science, Córdoba University, Spain); Francisco Solís-Cabrera, Secretary of Andalusia Research Plan, Andalusia Government, Spain; Prof. M^a Teresa Sánchez-Pineda de las Infantas (Vice rector of Innovation and Quality Management, Córdoba University); Prof. Panagiotis Zervatopoulos (University of Ioannina, Department of Business Administration of Food and Agricultural Products, Greece).

References

- Aghion, P., Howitt, P., 1992. A model of growth through creative destruction. *Econometrica* 60, 323-351.
- Anderson, R.D., Vastag, G., 2004. Causal modeling alternatives in operations research: Overview and application. *European Journal of Operational Research* 156, 92-109.
- Archibald, R. B., Pereira, A. M., 2003. Effects of public and private R&D on private-sector performance in the United States. *Public Finance Review* 31, 429-451.
- Auluck, R., 2002. Benchmarking: A tool for facilitating organizational learning? *Public Administration and Development* 22, 109-122.
- Balconi, M., Brusoni, S., Orsenigo, L., 2010. In defence of the linear model: An essay. *Research Policy* 39, 1-13.
- Baumol, W. J., 2003. Innovations and growth: Two common misapprehensions. *Journal of Policy Modeling* 25, 435-444.
- Block, J.H., Thurik, R., Zhou, H., 2012. What turns knowledge into innovative products? The role of entrepreneurship and knowledge spillovers. *Journal of Evolutionary Economics*, DOI: 10.1007/s00191-012-0265-5.
- Bonaccorsi, A., Daraio, C., Simar, L., 2006. Advanced indicators of productivity of universities. An application of robust nonparametric methods to Italian data. *Scientometrics* 66, 389-410.
- Bontis, N., Dragonetti, N.C., Jacobsen, K., Roos, G., 1999. The knowledge toolbox: a review of the tools available to measure and manage intangible resources, *European Management Journal* 17, 391-402.
- Boulanger, P.M., Brochet, T., 2005. Models for policy-making in sustainable development: The state of the art and perspectives for research. *Ecological Economics* 55, 337-350.
- Bound, J., Cumins, C., Griliches, Z., Hall, H.H., Jaffe, A., 1984. *R&D, Patent, and Productivity*. University of Chicago Press, Chicago.
- Braunerhjelm, P., 2008. Entrepreneurship, knowledge and economic growth. *Foundations and Trends in Entrepreneurship* 4, 451-533.
- Brinkman, R. L., Brinkman, J.E., 2001. The new growth theories: a cultural and social addendum. *International Journal of Social Economics* 28, 506-526.
- Carayannis, E.G., Campbell, D.F.J., 2010. Triple Helix, Quadruple Helix and Quintuple Helix and how do knowledge, innovation and the environment relate to each other? A proposed framework for a trans-disciplinary analysis of sustainable development and social ecology. *International Journal of Social Ecology and Sustainable Development* 1, 41-69.

- Chiesa, V., Coughlan, P., Voss, C.A., 1996. Development of a technical innovation audit. *Journal of Product Innovation Management* 13, 105-136.
- Cohen, W.M., Nelson R. R., Walsh, J.P., 2002. Links and Impacts: The Influence of Public Research on Industrial R&D. *Management Science* 48, 1-23.
- Cooper, W.W., Seiford, L.M., Zhu, J. (2004). Data envelopment analysis: history, models and interpretations, in: Cooper, W.W., Seiford, L.M., Zhu, J. (Eds), *Handbook on Data Envelopment Analysis*. Kluwer Academic Publishers, Boston, 1-45.
- Cooper, W.W., Seiford, L.M., Tone, K., 2007. *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software* (2nd ed.). New York: Springer.
- Darwiche, A., 2010. Bayesian networks. *Commun. ACM* 53, 12, 80-90. DOI: 10.1145/1859204.1859227
- David, P.A., Hall, B. H., 2000. Heart of Darkness: Modeling Public-Private Funding Interactions Inside the R&D Black Box. *Research Policy* 29, 1165-1183.
- De la Paz-Marín, M., Campoy-Muñoz, P., Hervás-Martínez, C., 2012. Non-linear classifiers for R&D performance in the EU. *Technological Forecasting and Social Change* 79, 1731-1745.
- Dosi, G., Llerena, P., Sylos Labini, M. (2009). Does the ‘European Paradox’ still hold? Did it ever? in: Delanghe, H., Muldur, U., Soete, L. (Eds), *European science and technology policy: Towards integration or fragmentation?* Edward Elgar, Cheltenham, pp. 214-236.
- Etzkowitz, H., Leydesdorff, L.A., 1998. The endless transition: A “triple helix” of university–industry–government. *Minerva* 36, 203–208.
- European Commission, 2009. Mainstreaming sustainable development into EU policies: Review of the European Union Strategy for Sustainable Development. COM (2009) 400 final, Brussels.
- European Commission. 2010. Europe 2020 - a strategy for smart, sustainable and inclusive growth. COM (2010) 2020. Brussels.
- Fagerberg, J., Mowery, D.C., Nelson, R.R., 2004. *The Oxford Handbook of Innovation*. Oxford University Press, Oxford.
- Fagerberg, J., Verspagen, B., 2009. Innovation studies - the emerging structure of a new scientific field. *Research Policy*, 218–233.
- Filippettia, A., Archibugia, D., 2011. Innovation in times of crisis: National Systems of Innovation, structure, and demand. *Research Policy* 40, 179-192.
- Freeman, C., 2004. Technological infrastructure and international competitiveness. *Industrial and Corporate Change* 13, 541-569.
- Freeman, C., Soete, L., 2009. Developing science, technology and innovation indicators: What we can learn from the past. *Research Policy* 38, 583-589.
- Furman, J.L., Porter, M.E, Stern, S., 2002. The determinants of national innovative capacities. *Research Policy* 31, 899-933.
- Gilbert, K., García-Alonso, C., Salvador-Carulla, L., 2010. Integrating clinicians, knowledge and data: expert-based cooperative analysis in health care decision support. *Health Research Policy and Systems* 8, 28-44.
- Godin, B., 2002. Technological gaps: an important episode in the construction of S&T statistics. *Technology in Society* 24, 387-413.
- Godin, B., 2006. The Historical Construction of an Analytical Framework, *Science Technology and Human Values* 31, 639-667.
- Godin, B., 2007. Science, accounting and statistics: The input-output framework. *Research Policy* 36, 1388-1403.
- Greif, S., 1985. Relation between R&D expenditure and patent applications. *World Patent Information* 7, 190–195.
- Griliches, Z., 1990. Patent statistics as economic indicators: a survey. *Journal of Economic Literature*, 28, 1661-1707.
- Grossman, G., Helpman, E., 1991. *Innovation and Growth in the Global Economy*. MIT Press, Cambridge, MA.
- Grupp, H., Schubert, T., 2010. Review and new evidence on composite innovation indicators for evaluating national performance. *Research Policy* 39, pp.67-78.

- Guan, J. C., Wang, J. X., 2004. Evaluation and interpretation of knowledge production efficiency. *Scientometrics* 59, 131-155.
- Hansen, J.A. (2001). Technology innovation indicators, a survey of historical development and current practice, in: Feldman, M.P., Link, A.N. (Eds.), *Innovation Policy in the Knowledge-Based Economy*, pp. 73–103.
- Jensen, F.V., 2001. *Bayesian Networks and Decision Graphs*. Springer-Verlag, New York.
- Jones, C.I., Williams, J.C., 1998. Measuring the social return to R&D. *The Quarterly Journal of Economics* 113, 1119-1135.
- Kim, H., Park, Y., 2008. The impact of R&D collaboration on innovative performance in Korea: A Bayesian network approach. *Scientometrics* 75, 535-554.
- Kline, S.J., Rosenberg, N. (1986). An overview of innovation, in: Landau, R., Rosenberg, N. (Eds), *The Positive Sum Strategy: Harnessing Technology for Economic Growth*. National Academy Press, Washington, D.C., pp. 275-305.
- Korhonen, P., Tainio, R., Wallenius, J., 2001. Value efficiency analysis of academic research. *European Journal of Operational Research* 130, 121-132.
- Lamirel, J. C., Al Shehabi, S., Francois, C., Polanco, X., 2004. Using a compound approach based on elaborated neural network for Webometrics: An example issued from the EICSTES project. *Scientometrics* 61, 427-441.
- Martin, B.R., 2012. The evolution of science policy and innovation studies. *Research Policy* 41, 1219-1239.
- Martin, S., Scott, J.T., 2000. The nature of innovation market failure and the design of public support for private innovation.
- Meyer, M., 2000. Does science push technology? Patents citing scientific literature. *Research Policy* 3, 409-434.
- Nadkarni, S., Shenoy, P.P., 2001. A Bayesian network approach to making inferences in causal maps. *European Journal of Operational Research* 128, 479-498.
- Nelson, R. R., (1996). *National Innovation Systems: A Retrospective on a Study*, in: *The Sources of Economic Growth*, Cambridge, MA, Harvard University Press, pp. 274-302.
- OECD, 2002. *The Measurement of Scientific and Technical Activities: Proposed Standard Practice for Surveys of Research and Experimental Development*, Frascati Manual (3rd ed.). OECD, Paris.
- OECD, 2010. *Main Science and Technology Indicators (MSTI) (2nd Ed)*. OECD Publications, Paris.
- OECD, Eurostat, 2005. *Oslo Manual Guidelines For Collecting And Interpreting Innovation Data (3rd ed.)*, OEDC Publications and Eurostat, Paris.
- Padmore, T., Schuetze, H., Gibson, H., 1998. Modeling systems of innovation: An enterprise-centered view. *Research Policy* 26, 05-624.
- Parker, S.C., 2008. The economics of formal business networks. *Journal of Business Venturing* 23, 627-640.
- Pavitt, K., 1998. The inevitable limits of EU R&D funding. *Research Policy* 27, 559-568.
- Pearl, J., 2000. *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge.
- European Commission, Pro Inno Europe (INNO METRICS), 2010. *European Innovation Scoreboard 2010: Comparative Analysis of Innovation Performance*, Initiative of the European Union, online: <http://www.proinno-europe.eu/inno-metrics/page/innovation-union-scoreboard-2011> (last accessed on May 25, 2012).
- Romer, P., 1990. Endogenous technological change. *Journal of Political Economy* 98, S71–S102.
- Rousseau, S., Rousseau, R., 1997. Data analysis as a tool for constructing scientometric indicators, *Scientometrics* 40, 45-46.
- Ruiz, C.F., Bonilla, R., Chavarro, D., Orozco, L. A., Zarama, R., Polanco, X., 2009. Efficiency measurement of research groups using Data Envelopment Analysis and Bayesian networks. *Scientometrics* 83, 711-721.
- Saen, R.F., 2010. Restricting weights in supplier selection decisions in the presence of dual-role factors. *Applied Mathematical Modelling* 34, 2820-2830.
- Salter A.J., Martin, B.R., 2001. The economic benefits of publicly funded basic research: a critical review. *Research Policy* 30, 509-532.
- Schumpeter, J.A., 1934. *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle*. Harvard University Press, Cambridge, Mass.

SCImago, 2007. SJR-SCImago Journal & Country Rank (last accessed on January 15, 2012 from <http://www.scimagojr.com>).

Solow, R.M., 1956. A Contribution to the economic theory of economic growth. *Quarterly Journal of Economics* 70 (1), 65-94.

Thomas, V.J., Sharma, S., Jain, S.K., 2011. Using patents and publications to assess R&D efficiency in the states of the USA. *World Patent Information* 33, 4-10.

Vedung, E., (1998). Policy Instruments: Typologies and Theories, in: Bemelmans-Videc, M. L., Rist, R.C., Vedung, E. (Eds), *Carrots, Sticks & Sermons*. Transaction Publishers, New Brunswick and London, pp. 21-58.

Von Hippel, E., Tyre, M.J., 1995. How learning by doing is done: problem identification in novel process equipment. *Research Policy* 24, 1-12.

Wang, E.C., Huang, W., 2007. Relative efficiency of R&D activities: a cross-country study accounting for environmental factors in the DEA approach. *Research Policy* 36, 260-273.

World Bank, 2012. Various datasets. Official website – <http://data.worldbank.org> (last accessed on January 18, 2012).

World Economic Forum, 2010. The Global Competitiveness Report 2010–2011. World Economic Forum, Geneva, Switzerland (http://www3.weforum.org/docs/WEF_GlobalCompetitivenessReport_2010-11.pdf).

World International Patent Office, 2011. Official website – <http://wipo.com> (last accessed on March 20, 2012).

Yu, Z., 2005. Trade, Market Size, and Industrial Structure: Revisiting the Home-Market Effect. *The Canadian Journal of Economics* 38, 255-272.

Table 1. Descriptive statistics of the inputs and outputs employed in the 2LS-DEA (n=27).

VARIABLES	Min.	Max.	Mean	Std. Deviation
<i>AGEDEP</i> ₂₀₀₈	38.15	53.87	46.88	4.15
<i>BROSUS</i> ₂₀₀₈	88.26	411.19	217.01	85.03
<i>CARBDI</i> ₂₀₀₈	0.08	0.87	0.27	0.19
<i>ELEAV</i> ₂₀₀₈	5.00	39.00	14.37	8.46
<i>ENINT</i> ₂₀₀₈	103.13	944.16	285.07	201.53
<i>ERDFUN</i> ₂₀₀₆	0.16	0.88	0.50	0.21
<i>GASEMI</i> ₂₀₀₈	5.25	25.57	10.70	3.96
<i>GDPGRO</i> ₂₀₀₈	-5.20	7.50	.88	3.18
<i>HITECEX</i> ₂₀₀₈	5.03	50.19	14.42	10.28
<i>INNEWP</i> ₂₀₀₈	5.88	18.67	12.89	3.58
<i>INNSME</i> ₂₀₀₈	16.82	53.61	32.39	10.48
<i>INTUSE</i> ₂₀₀₈	28.79	87.70	60.82	16.49
<i>LIFEXP</i> ₂₀₀₈	71.96	84.51	78.33	3.32
<i>LLEARN</i> ₂₀₀₈	1.40	30.00	9.57	7.18
<i>MOBILE</i> ₂₀₀₈	2.63	44.88	9.33	9.01
<i>MORRAT</i> ₂₀₀₈	1.80	26.40	5.29	4.71
<i>PHD06</i> ₂₀₀₈	0.02	0.31	0.17	0.09
<i>PUBPER</i> ₂₀₀₆	0.18	1.01	0.47	0.16
<i>RLPGH</i> ₂₀₀₈	39.30	177.70	93.50	30.30
<i>SPUBLI</i> ₂₀₀₈	0.28	2.54	1.27	0.68
<i>TPATRE</i> ₂₀₀₈	0.04	1.92	0.61	0.67
<i>TRADEM</i> ₂₀₀₈	1.03	17.56	5.09	3.52
<i>UNEMPL</i> ₂₀₀₈	2.80	11.30	6.21	1.87

Table 2. Rule-Base for the transformation of inputs and outputs values according to their parental variables in the Bayesian Network

DEPENDENCE RELATIONSHIPS IN THE BAYESIAN NETWORK (<i>t</i> year of reference)	SET OF RULES (R_i) ^(a)
$ERDFUN_{t,2} = f(EUSTRAT_{t,2})$	<p>R1: IF $ERDFUN_{t,2} < 0.4$ AND $GERDIO_{t,2} < 0.55$ AND $GERDIO_{t,2} < 2$ THEN $ERDFUN_{t,2} = ERDFUN_{t,2} * 2$</p> <p>R2: IF $ERDFUN_{t,2} > 0.55$ AND $GERDIO_{t,2} < 2$ THEN $ERDFUN_{t,2} = GERDIO_{t,2} / 2$</p> <p>R3: IF $ERDFUN_{t,2} < 0.2$ THEN $ERDFUN_{t,2} = ERDFUN_{t,2} / 2$</p> <p>R4: IF $PUBPER_{t,2}$ = LOW AND $MOBILI_{t,2}$ = HIGH AND $MARKRE_{t,2}$ = ATTRACTIVE THEN $PUBPER_{t,2} = PUBPER_{t,2} * 2$</p> <p>R5: IF $PUBPER_{t,2}$ = HIGH AND $MOBILI_{t,2}$ = LOW AND $MARKRE_{t,2}$ = NOT ATTRACTIVE THEN $PUBPER_{t,2} = PUBPER_{t,2} / 2$</p>
$TPATRE_t = f(ERDFUN_{t,2}, PUBPER_{t,2}, ERDPRI_{t,2}, PERPRI_{t,2}, SPUBLI_t, AGECOL_t, INTECAP_t, ENTREP_t, SBNETW_t, SPILLOV_t, INFRAE_t)$	<p>R6: IF $TPATRE_t$ = LOW AND $ERDFUN_{t,2}$ = HIGH AND $PUBPER_{t,2}$ = HIGH AND $ERDPRI_{t,2}$ = HIGH AND $PERPRI_{t,2}$ = HIGH AND $SPUBLI_t$ = HIGH AND $AGECOL_t$ = HIGH AND $INTECAP_t$ = HIGH AND $ENTREP_t$ = HIGH AND $SBNETW_t$ = HIGH AND $SPILLOV_t$ = HIGH AND $INFRAE_t$ = GOOD AND $INSTFR_t$ = ESTABLE THEN $TPATRE_t = TPATRE_t / 2$</p> <p>R7: IF $TPATRE_t$ = HIGH AND $ERDFUN_{t,2}$ = LOW AND $PUBPER_{t,2}$ = LOW AND $ERDPRI_{t,2}$ = LOW AND $PERPRI_{t,2}$ = LOW AND $SPUBLI_t$ = LOW AND $AGECOL_t$ = LOW AND $INTECAP_t$ = LOW AND $ENTREP_t$ = LOW AND $SBNETW_t$ = LOW AND $SPILLOV_t$ = LOW AND $INFRAE_t$ = NOT GOOD AND $INSTFR_t$ = UNSTABLE THEN $TPATRE_t = TPATRE_t * 2$</p> <p>R8: IF $TPATRE_t$ = HIGH AND $SPUBLI_t$ = LOW AND $PUBPER_{t,2}$ = LOW AND $ERDFUN_{t,2}$ = LOW AND $ERDPRI_{t,2}$ = LOW AND $PERPRI_{t,2}$ = LOW THEN $TPATRE_t = TPATRE_t * 2$</p> <p>R9: IF $TPATRE_t$ = HIGH AND $ERDPRI_{t,2}$ = LOW AND $PERPRI_{t,2}$ = LOW AND $TPATRE_t = TPATRE_t * 2$</p> <p>R10: IF $TPATRE_t$ = LOW AND $SPUBLI_t$ = HIGH AND $PUBPER_{t,2}$ = HIGH AND $ERDFUN_{t,2}$ = HIGH AND $ERDPRI_{t,2}$ = HIGH AND $TPATRE_t = TPATRE_t * 2$</p> <p>R11: IF $SPUBLI_t$ = HIGH AND $ERDFUN_{t,2}$ = HIGH AND $PUBPER_{t,2}$ = HIGH AND $ERDPRI_{t,2}$ = HIGH AND $PERPRI_{t,2}$ = HIGH AND $INFRAE_t$ = GOOD AND $AGECOL_t$ = HIGH THEN $SPUBLI_t = SPUBLI_t / 2$</p> <p>R12: IF $SPUBLI_t$ = HIGH AND $ERDFUN_{t,2}$ = LOW AND $PUBPER_{t,2}$ = LOW AND $ERDPRI_{t,2}$ = LOW AND $PERPRI_{t,2}$ = LOW AND $INFRAE_t$ = NOT GOOD AND $AGECOL_t$ = LOW THEN $SPUBLI_t = SPUBLI_t * 2$</p> <p>R13: IF $SPUBLI_t$ = LOW AND $AGECOL_t = 1$ AND $ERDFUN_{t,2}$ = HIGH AND $INFRAE_t = SI$ AND $ERDPRI_{t,2}$ = HIGH AND $SBNETW_t = 1$ THEN $SPUBLI_t = SPUBLI_t / 2$</p>
$GDPGRO_t = f(RLPGH_t)$	<p>R14: IF $GDPGRO_t$ = HIGH AND $RLPGH_t$ = LOW THEN $GDPGRO_t = GDPGRO_t * 2$</p> <p>R15: IF $GDPGRO_t$ = LOW AND $RLPGH_t$ = HIGH THEN $GDPGRO_t = GDPGRO_t / 2$</p> <p>R16: IF $GDPGRO_t > 5$ THEN $GDPGRO_t = GDPGRO_t / 2$</p>
$RLPGH_t = f(UNEMPL_t, LABCLI_t, MOTIVA_t, LLEARN_t, PHD06_t, ELEAV_t, INFRAE_t, TECCHA_t)$	<p>R17: IF $RLPGH_t$ = HIGH AND $UNEMPL_t$ = HIGH AND $LABCLI_t$ = BAD AND $MOTIVA_t$ = UNMOTIVATED AND $LLEARN_t$ = LOW AND $PHD06_t$ = LOW AND $ELEAV_t$ = HIGH AND $INFRAE_t$ = NOT GOOD AND $TECCHA_t$ = WEAK THEN $RLPGH_t = RLPGH_t * 2$</p> <p>R18: IF $RLPGH_t$ = LOW AND $UNEMPL_t$ = LOW AND $LABCLI_t$ = GOOD AND $MOTIVA_t$ = MOTIVATED AND $LLEARN_t$ = HIGH AND $PHD06_t$ = HIGH AND $ELEAV_t$ = LOW AND $INFRAE_t$ = GOOD AND $TECCHA_t$ = STRONG THEN $RLPGH_t = RLPGH_t / 2$</p> <p>R19: IF $RLPGH_t$ = HIGH AND $TECCHA_t$ = WEAK AND $LLEARN_t$ = LOW THEN $RLPGH_t = RLPGH_t / 2$</p> <p>R20: IF $RLPGH_t$ = LOW AND $MOTIVA_t = 1$ AND $LABOCLIM_t = 1$ AND $TECCHA_t$ = STRONG AND $PHD06_t$ = HIGH AND $ELEAV_t$ = HIGH AND $LLEARN_t$ = HIGH THEN $RLPGH_t = RLPGH_t / 2$</p> <p>R21: IF $RLPGH_t$ = LOW AND $PHD06_t$ = HIGH AND $ELEAV_t$ = HIGH THEN $RLPGH_t = RLPGH_t / 2$</p>
$UNEMPL_t = f(PHD06_t, LLEARN_t, ELEAV_t, GDPGRO_t)$	<p>R22: IF $UNEMPL_t$ = LOW AND $PHD06_t$ = LOW AND $LLEARN_t$ = LOW AND $ELEAV_t$ = LOW AND $GDPGRO_t$ = LOW THEN $UNEMPL_t = UNEMPL_t * 2$</p> <p>R23: IF $UNEMPL_t$ = HIGH AND $PHD06_t$ = HIGH AND $LLEARN_t$ = HIGH AND $ELEAV_t$ = HIGH AND $GDPGRO_t$ = HIGH THEN $UNEMPL_t = UNEMPL_t / 2$</p>
$HITECEX_t = f(compet)$	<p>R24: IF $HITECEX_t$ = HIGH AND $compet$ = LOW THEN $HITECEX_t = HITECEX_t * 2$</p> <p>R25: IF $HITECEX_t$ = LOW AND $compet$ = HIGH THEN $HITECEX_t = HITECEX_t / 2$</p> <p>R26: IF $HITECEX_t > 30$ AND $compet$ = LOW THEN $HITECEX_t = HITECEX_t / 3$</p>
$ELEAV_t = f(EDUEXP_t, SPUBLI_t)$	<p>R27: IF $ELEAV_t$ = HIGH AND $EDUEXP_t$ = HIGH AND $SPUBLI_t$ = HIGH THEN $ELEAV_t = ELEAV_t / 2$</p> <p>R28: IF $ELEAV_t$ = LOW AND $EDUEXP_t$ = LOW AND $SPUBLI_t$ = LOW THEN $ELEAV_t = ELEAV_t * 2$</p>
$PHD06_t = f(EDUEXP_t, SPUBLI_t)$	<p>R29: IF $PHD06_t$ = HIGH AND $EDUEXP_t$ = LOW AND $SPUBLI_t$ = LOW THEN $PHD06_t = PHD06_t * 2$</p> <p>R30: IF $PHD06_t$ = LOW AND $EDUEXP_t$ = HIGH AND $SPUBLI_t$ = HIGH THEN $PHD06_t = PHD06_t / 2$</p>
$LLEARN_t = f(EDUEXP_t, SPUBLI_t)$	<p>R31: IF $LLEARN_t$ = HIGH AND $EDUEXP_t$ = LOW AND $SPUBLI_t$ = LOW THEN $LLEARN_t = LLEARN_t * 2$</p> <p>R32: IF $LLEARN_t$ = LOW AND $EDUEXP_t$ = HIGH AND $SPUBLI_t$ = HIGH THEN $LLEARN_t = LLEARN_t / 2$</p>

$INNSME_t = f(INTECAP_t, ENTREP_t, SBNETW_t, SPILLOV_t, INSTFR_t, LEARDO_t, BENCHM_t)$	R33: IF $INNSME_t = HIGH$ AND $INTECAP_t = LOW$ AND $ENTREP_t = NO$ AND $SBNETW_t = LOW$ AND $SPILLOV_t = LOW$ AND $INSTFR_t = NOT GOOD$ AND $LEARDO_t = LOW$ AND $BENCHM_t = LOW$ THEN $INNSME_t = INNSME_t * 2$ R34: IF $INNSME_t = LOW$ AND $INTECAP_t = HIGH$ AND $ENTREP_t = YES$ AND $SBNETW_t = HIGH$ AND $SPILLOV_t = HIGH$ AND $INSTFR_t = GOOD$ AND $LEARDO_t = HIGH$ AND $BENCHM_t = HIGH$ THEN $INNSME_t = INNSME_t / 2$
$INNEWP_t = f(INTECAP_t, ENTREP_t, SBNETW_t, SPILLOV_t, INSTFR_t, LEARDO_t, BENCHM_t)$	R35: IF $INNEWP_t = HIGH$ AND $INTECAP_t = LOW$ AND $ENTREP_t = NO$ AND $SBNETW_t = LOW$ AND $SPILLOV_t = LOW$ AND $INSTFR_t = NOT GOOD$ AND $LEARDO_t = LOW$ AND $BENCHM_t = LOW$ THEN $INNEWP_t = INNEWP_t * 2$ R36: IF $INNEWP_t = LOW$ AND $INTECAP_t = HIGH$ AND $ENTREP_t = YES$ AND $SBNETW_t = HIGH$ AND $SPILLOV_t = HIGH$ AND $INSTFR_t = GOOD$ AND $LEARDO_t = HIGH$ AND $BENCHM_t = HIGH$ THEN $INNEWP_t = INNEWP_t / 2$
$TRADEM_t = f(INTECAP_t, ENTREP_t, SBNETW_t, SPILLOV_t, INSTFR_t, LEARDO_t, BENCHM_t)$	R37: IF $TRADEM_t = HIGH$ AND $INTECAP_t = LOW$ AND $ENTREP_t = NO$ AND $SBNETW_t = LOW$ AND $SPILLOV_t = LOW$ AND $INSTFR_t = NOT GOOD$ AND $LEARDO_t = LOW$ AND $BENCHM_t = LOW$ THEN $TRADEM_t = TRADEM_t * 2$ R38: IF $TRADEM_t = LOW$ AND $INTECAP_t = HIGH$ AND $ENTREP_t = YES$ AND $SBNETW_t = HIGH$ AND $SPILLOV_t = HIGH$ AND $INSTFR_t = GOOD$ AND $LEARDO_t = HIGH$ AND $BENCHM_t = HIGH$ THEN $TRADEM_t = TRADEM_t / 2$
$INTUSE_t = f(SPUBLI_t, TPATRE_t)$	R39: IF $INTUSE_t = HIGH$ AND $SPUBLI_t = LOW$ AND $TPATRE_t = LOW$ THEN $INTUSE_t = INTUSE_t * 2$ R40: IF $INTUSE_t = LOW$ AND $SPUBLI_t = HIGH$ AND $TPATRE_t = HIGH$ THEN $INTUSE_t = INTUSE_t / 2$
$BROSUS_t = f(SPUBLI_t, TPATRE_t)$	R41: IF $BROSUS_t = HIGH$ AND $SPUBLI_t = LOW$ AND $TPATRE_t = LOW$ THEN $BROSUS_t = BROSUS_t * 2$ R39: IF $BROSUS_t = LOW$ AND $SPUBLI_t = HIGH$ AND $TPATRE_t = HIGH$ THEN $BROSUS_t = BROSUS_t / 2$
$MOBILE_t = f(SPUBLI_t, TPATRE_t)$	R40: IF $MOBILE_t = HIGH$ AND $SPUBLI_t = LOW$ AND $TPATRE_t = LOW$ THEN $MOBILE_t = MOBILE_t * 2$ R42: IF $MOBILE_t = LOW$ AND $SPUBLI_t = HIGH$ AND $TPATRE_t = HIGH$ THEN $MOBILE_t = MOBILE_t / 2$
$AGEDEP_t = f(SPUBLI_t, TPATRE_t, HEAEXP_t)$	R43: IF $AGEDEP_t = HIGH$ AND $SPUBLI_t = HIGH$ AND $TPATRE_t = HIGH$ AND $HEAEXP_t = HIGH$ THEN $AGEDEP_t = AGEDEP_t * 2$ R44: IF $AGEDEP_t = LOW$ AND $SPUBLI_t = LOW$ AND $TPATRE_t = LOW$ AND $HEAEXP_t = LOW$ THEN $AGEDEP_t = AGEDEP_t / 2$
$LIFEXP_t = f(SPUBLI_t, TPATRE_t, HEAEXP_t)$	R45: IF $LIFEXP_t = HIGH$ AND $SPUBLI_t = LOW$ AND $TPATRE_t = LOW$ AND $HEAEXP_t = LOW$ THEN $LIFEXP_t = LIFEXP_t * 2$ R46: IF $LIFEXP_t = LOW$ AND $SPUBLI_t = HIGH$ AND $TPATRE_t = HIGH$ AND $HEAEXP_t = HIGH$ THEN $LIFEXP_t = LIFEXP_t / 2$
$MORRAT_t = f(SPUBLI_t, TPATRE_t, HEAEXP_t)$	R47: IF $MORRAT_t = HIGH$ AND $SPUBLI_t = HIGH$ AND $TPATRE_t = HIGH$ AND $HEAEXP_t = HIGH$ THEN $MORRAT_t = MORRAT_t * 2$ R48: IF $MORRAT_t = LOW$ AND $SPUBLI_t = LOW$ AND $TPATRE_t = LOW$ AND $HEAEXP_t = LOW$ THEN $MORRAT_t = MORRAT_t / 2$
$ENINT_t = f(SPUBLI_t, TPATRE_t)$	R49: IF $ENINT_t = HIGH$ AND $SPUBLI_t = HIGH$ AND $TPATRE_t = HIGH$ THEN $ENINT_t = ENINT_t * 2$ R50: IF $ENINT_t = LOW$ AND $SPUBLI_t = LOW$ AND $TPATRE_t = LOW$ THEN $ENINT_t = ENINT_t / 2$
$GASEMI_t = f(SPUBLI_t, TPATRE_t)$	R51: IF $GASEMI_t = HIGH$ AND $SPUBLI_t = HIGH$ AND $TPATRE_t = HIGH$ THEN $GASEMI_t = GASEMI_t * 2$ R52: IF $GASEMI_t = LOW$ AND $SPUBLI_t = LOW$ AND $TPATRE_t = LOW$ THEN $GASEMI_t = GASEMI_t / 2$
$CARBBDI_t = f(SPUBLI_t, TPATRE_t)$	R53: IF $CARBBDI_t = HIGH$ AND $SPUBLI_t = HIGH$ AND $TPATRE_t = HIGH$ THEN $CARBBDI_t = CARBBDI_t * 2$ R54: IF $CARBBDI_t = LOW$ AND $SPUBLI_t = LOW$ AND $TPATRE_t = LOW$ THEN $CARBBDI_t = CARBBDI_t / 2$

^(a) See Appendix A and Appendix B for detailed information about variables and interpretation of the values.

Table 3. Relative technical efficiency of DMUs by scenario (BCC output-oriented model; 100 for fully efficient RDTI public policies and values lower than 100 mean some kind of inefficiencies). Period 2006-2008.

DMUs (RTDI public policy of each country)	Traditional	Economic	Innovation	Education	Information Society	Health	Environment
Austria	64.01	97.09	70.39	99.55	100	84.1	94.83
Belgium	100	100	100	100	80.9	100	100
Bulgaria	11.4	98.77	43.66	70.18	82.05	98.27	92.69
Cyprus	26.39	100	100	73.49	78.23	98.85	83.72
Czech Republic	31.65	97.96	100	93.44	85.25	100	56.96
Denmark	85.79	100	92.04	100	100	74.83	100
Estonia	30.77	87.69	40.91	93.01	100	91.44	50.63
Finland	100	100	100	100	100	98.3	85.36
France	38.55	96.66	45.53	86.47	100	100	100
Germany	84.04	99.47	100	86.83	100	88.04	97.13
Greece	38.52	83.92	100	70.18	100	88.92	98.73
Hungary	20.61	100	90.16	79.52	89.37	86.08	98.38
Ireland	96.16	100	100	100	55.5	98.29	100
Italy	32.57	82.75	38.9	58.26	100	100	100
Latvia	9.01	78.87	38.57	68.07	71.42	77.83	100
Lithuania	100	100	76.94	70.18	93.25	79.14	98
Luxembourg	98.68	100	100	74.4	100	100	99.3
Malta	18.07	100	61.03	29.79	95.53	100	99.46
Netherlands	86.81	100	73.24	80.65	99.24	100	93.93
Poland	16.48	96.78	43.7	99.7	70.82	100	71.02
Portugal	28.08	84.61	96.64	33.92	100	83.84	100
Romania	8.57	100	45.81	66.87	71.8	97.4	96.38
Slovakia	21.11	96.58	96.84	100	62.72	76.1	94.74
Slovenia	55.52	100	100	99.4	43.72	100	80.98
Spain	27.87	81.83	100	71.54	82.92	85.2	97.18
Sweden	100	100	100	100	100	100	100
United Kingdom	63.02	90.12	79.7	89.95	100	84.74	97.06

Appendix A. Definition of Bayesian Network (BN) variables used as inputs and outputs in 2LS-DEA approach.

NODE LABEL	VARIABLE ^(a)	DEFINITION	IO	Scenario	% ^(b)	INTERPRETATION OF IO VALUES ^(c)
ERDFUN	(1) R&D public expenditure	Research and development expenditure, by source of funds: Government and Higher Education sectors as% of GDP.	I	Traditional	100	Low: [0-0.55]; High: (0.55-1)
PUBPER	(2) R&D public personnel	Research and development personnel, by sectors of performance: Government and Higher education sectors are considered as a percentage of total employment (Full time equivalent in denominator).	I	Traditional	100	Low: [0-0.7]; High: (0.7-1.5)
TPATRE	(3) Total patents applications ^(***)	Worldwide patent applications filed through the Patent Cooperation Treaty procedure or with a national patent office (residents) per 1000 population.	IO	Input in the traditional scenario becomes an output in the rest	10%	Low: [0-1]; High: (1-4)
SPUBLI	(4) Total Scientific publications ^(***)	Number of documents in the Scientific Journal Ranking (1000 per population).	IO	Input in traditional scenario becomes an output in the rest	90%	Low: [0-1.5]; High: (1.5-3)
GDPGRO	(5) Real GDP per capita, growth rate	Real GDP per capita, growth rate; Percentage change on previous period, GDP includes goods and services that have markets (or which could have markets) and products which are produced by general government and non-profit institutions. Real GDP per capita is calculated as the ratio of real GDP to the average population of a specific year.	O	Economic	10%	Low: <=1; Medium: (1-3); High>3
UNEMPL	(6) Unemployment rate	Percentage of the labor force based on International Labor Office (ILO) definition. The labor force is the total number of people employed and unemployed.	O	Economic	10%	Low: [0-5]; Medium: (5-8); High>8
HITCEX	(7) High-tech exports ^(***)	Share of exports of all high technology products of total exports. The total exports for the EU do not include the intra-EU trade.	O	Economic	10%	Low: [0-10]; Medium: (10-15); High: (15-100)
RLPGH	(8) Real labor productivity growth per hour worked	Real labor productivity per hour worked is calculated as real output (deflated GDP measured in chain-linked volumes, reference year 2000) per unit of labor input (measured by the total number of hours worked). Measuring labor productivity per hour worked provides a better picture of productivity developments in the economy than labor productivity per person employed, as it eliminates differences in the full time/part time composition of the workforce across countries and years.% Change over previous year, index 2000 = 100.	O	Economic	10%	Low: <=1; Medium: (1-3); High: >3
ELEAV	(9) Early Leavers from education and training	Early leavers from education and training refers to persons aged 18 to 74 fulfilling the following two conditions: first, the highest level of education or training attained is ISCED 0, 1, 2 or 3c short, second, respondents declared not having received any education or training 5in the four weeks preceding the survey (numerator). Both the numerators and the denominators come from the EU Labor Force Survey.	O	Education	5%	Low: [0-8]; High>8
PHD06	(10) PhD graduates	Number of PhD graduates per 1000 population.	O	Education	90%	Low: [0-0.15]; High: >0.15
LLEARN	(11) Lifelong Learning	Lifelong learning (% of persons aged 18 to 64).	O	Education	50%	Low: [0-15]; High: >15
INNSME	(12) SMEs introducing marketing/organizational innovations	SMEs introducing marketing or organizational innovations as % of SMEs.	O	Innovation	40%	Low: [0-30]; Medium: (30-50); High: >50
INNEWP	(13) Sales of new innovations	Sales of new to market and new to firm innovations as % of turnover.	O	Innovation	40%	Low: [0-10]; Medium: (10-15); High:>15
TRADEM	(14) Community trademarks ^(***)	Community trademarks per billion GDP	O	Innovation	20%	Low: [0-5]; Medium: (5-10); High:>10
INTUSE	(15) Internet users ^(***)	People with access to the worldwide network (per 1,000 population).	O	Information Society	80%	Low: [0-70]; High: >70
BROSUS	(16) Fixed broadband Internet subscribers	The broadband penetration rate describes the number of dedicated, high-speed connections per 1,000 inhabitants. This indicator shows how widely broadband access to the internet has spread in the countries on the general level, not specifying by user group. Broadband lines are defined as those with a capacity equal or higher than 144 Kbits/s. Various technologies are covered; ADSL, cable modem as well as other types of access lines.	O	Information Society	80%	Low: [0-275]; High: >275
MOBILE	(17) Mobile cellular subscriptions	Number of mobiles cellular subscriptions per 1,000 population.	O	Information Society	80%	Low: [0-1000]; High: >1000

NODE LABEL	VARIABLE ^(a)	DEFINITION	IO	Scenario	% ^(b)	INTERPRETATION OF IO VALUES ^(c)
LIFEXP	(18) Life Expectancy	The mean number of years that a new-born child can expect to live if subjected throughout his life to the current mortality conditions (age specific probabilities of dying).	O	Health	70%	High: [60-80]; Very high: >80
AGEDEP	(19) Age dependency ratio (% of working-age population)	This indicator is the ratio between the total number of elderly persons of an age when they are generally economically inactive (aged 65 and over) and the number of persons of working age (from 15 to 64).	O	Health	70%	Low: [0-50]; High:>50
MORRAT	(20) Mortality rate, infant (per 1,000 live births)	The ratio of the number of deaths of children under one year of age during the year to the number of live births in that year. The value is expressed per 1000 live births.	O	Health	70%	Low: [0-4]; High:>4
ENINT	(21) Energy Intensity of the Economy	Ratio between the gross inland consumption of energy and the gross domestic product (GDP) for a given calendar year. It measures the energy consumption of an economy and its overall energy efficiency. The gross inland consumption of energy is calculated as the sum of the gross inland consumption of five energy types: coal, electricity, oil, natural gas and renewable energy sources. The GDP figures are taken at chain-linked volumes with reference year 2000. Since gross inland consumption is measured in kgoe (kilogram of oil equivalent) and GDP in 1,000 eur.	O	Environment	50%	Low: [0-150]; Medium: (150-300); High:>300
GASEMI	(22) Total greenhouse gas emissions	Greenhouse gas emissions by sector (source EEA); Total emissions per 1000 population. The annual greenhouse gas (GHG) emissions are estimated and reported under the United Nations Framework Convention on Climate Change (UNFCCC), the Kyoto Protocol and the Decision 280/2004/EC.	O	Environment	50%	Low: [0-10]; Medium: (10-15); High:>20
CARBIDI	(23) Carbon dioxide damage ^(**)	Adjusted savings: carbon dioxide damage (% of GNI).	O	Environment	50%	Low: [0-0.2]; Medium: (0.2-0.4); High:>0.4

^(a) Source: Eurostat, except: ^(*) Scientific publications (SCImago), ^(**) World Bank Database, ^(***) WIPO and ^(****) Global Competitiveness Report by the World Economic Forum (WEF).

^(b) Percentage of the value of the variable used in the 2LS-DEA approach for the corresponding scenario.

^(c) Interpretation of the values by experts that will be employed for the formulation of the rule-base.

Appendix B. Definition of the rest of the Bayesian Network (BN) variables not included in any 2LS-DEA scenario (alphabetic order).

NODE LABEL	VARIABLE ^(a)	DEFINITION	STATES AND INTERVALS ^(b)
AGECOL	(24) Collaboration of agents	Collaboration of agents of the innovation system.	Low=0; High=1
BENCHM	(25) Benchmarking	The continuous process of measuring products, services and practices against the toughest competitors or those companies recognized as industry leaders or the search for industry best practices that lead to superior performance.	Low=0; High=1
COMPET	(26) Competitiveness ^(****)	The set of institutions, policies, and factors that determine the level of productivity of a country (ranking from the Global Competitiveness Report, WEF).	High=1-25; Low=27-133
COJIMA	(27) Country image abroad	Trust and reputation of the country abroad	Good=1; Bad=0
EDUEXP	(28) Expenditure on education as % of GDP	Total public expenditure on education as % of GDP, for all levels of education combined.	Low: [0-4.75]; High=>4.75
ENTREP	(29) Entrepreneurship culture	Culture is used to refer to the enduring set of values of the country. Entrepreneurship is defined as the act and process by which society or individuals identify and pursue business opportunities to create wealth.	Yes=1; No=0
ERDPRIV	(30) R&D private expenditure	Research and development expenditure, by source of funds: business sectors as% of GDP	Low: [0-1]; High: [1-3]
EUSTRAT	(31) European Strategy for R&D public policy	Main objectives of EU Strategies for reaching a knowledge-based society, a smart economic growth, more competitiveness, etc., like Lisbon Strategy (at least R&D total expenditure have to reach 3% of GDP and private sector must represent at least 60% of total GERD.	Very good: ERDFUN/GERDT<=0.4 AND GERDT>=3% Very bad: ERDFUN/GERDT>0.55 AND GERDT<1.5%
FORINV	(32) Foreign investment	Presence of multinational enterprises in the country.	Low=0; High=1
GERDIO	(33) R&D total expenditure	Research and development expenditure (total) as% of GDP.	Low:[0-1.5]; Medium(1.5-3) ; High;>=3
HEAEXP	(34) Health Expenditure	Total health expenditure as % of GDP.	Low: [0-8]; High;>8
INFRAE	(35) Infrastructures	Gross fixed capital formation consists of resident producers' acquisitions, less disposals, of fixed tangible or intangible assets. This covers in particular machinery and equipment, vehicles, dwellings and other buildings, including laboratories and other scientific infrastructures and investment.	Good=1 Not good=0
INSTFR	(36) Institutional and political Framework	The legal and administrative framework within which individuals, firms, and governments interact to generate income and wealth in an economy determines the institutional environment.	Stable=1; Unstable=0
INTECAP	(37) Intellectual capital	Refers to the difference in value between tangible assets (physical and financial) and market value (including human, structural and relational capital).	Low=0; High=1
LABCLI	(38) Labour Climate	Labour-management climate in the organizations.	Bad=0; Good=1
LEARDO	(39) Learning by doing	The earlier the innovation, the sooner a firm's employees can learn by doing, which leads to more innovation and lower costs.	Low=0; High=1
MARKEF	(40) Markets efficiency	Ranking from Global Competitiveness report.	Low=0; High=1
MARKRE	(41) Market for researchers ^(****)	Attractiveness of market for researchers (wages).	Attractive=1; Not attractive=0
MOBILI	(42) Mobility of researchers and students	Level of mobility of researchers and student from other European countries.	High, Low
MOTIVA	(43) Motivation	Skills, motivation and commitment as vital to how productive people are at work.	Motivated=1; Unmotivated=2
PERPRI	(44) R&D private personnel	Research and development personnel, by sectors of performance: business sector, considered as a percentage of total employment (Full time equivalent in denominator).	Low: (0-1]; High: (1-3]
PURDSY	(45) Public R&D system	The set of institutions involved in R&D activities.	Appropriate=1; Non appropriate=0
SBNETW	(46) Social and business networks	Social networking is the practice of expanding the number of one's business and/or social contacts by making connections through individuals. Potential of the Internet to promote such connections is only now being fully recognized and exploited.	Low=0; High=1
SFCEST	(47) Sector structure	Categorization of the countries' economy according to the prevalence or not of the industrial sector.	Industrial=1; Other sectors=0
SIZENT	(46) Size of enterprises	Size of enterprises that is predominant in the country.	SME=0; Large Enterprises=1
SPILLO	(48) Spillovers	Presence of strong or weak Spillovers.	Low=0; High=1
TECACQ	(49) Other ways of technology acquisition	Different ways of technological acquisition measured through the technological balance of payments, etc.	Positive=1; Negative=0
TECCHA	(50) Technological change	Technological change is a term that is used in this study to describe a change in the set of feasible production possibilities.	Strong=1; Weak=0
WELFAR	(51) Welfare of European societies	It is concerned with the improvement of quality of life of European societies.	High, Medium, Low

^(a) Source: Eurostat, except: ^(*) Scientific publications (SCImago), ^(**) WIPO and ^(***) Global Competitiveness Report by the World Economic Forum.

^(b) Interpretation of the values by experts that will be employed for the formulation of the rule-base.

