

La asignación de ayudas públicas a la I+D empresarial: escogen los evaluadores los mejores proyectos? *

Versión provisional

Por favor no citar sin el consentimiento de los autores

Néstor Duch^{1,2}, Jose García-Quevedo^{1,2} y Daniel Montolio^{2,3}

Resumen. Uno de los problemas principales con que normalmente se enfrentan las agencias públicas locales dedicadas a promover actividades de I+D, es que reciban más solicitudes para proyectos de I+D que recursos públicos disponibles para financiarlos. Esta restricción exige el establecimiento de un proceso de selección entre proyectos. Esta parte del proceso de otorgamiento de subvenciones pública es crucial (y costoso) para la agencia si quiere dar una señal inequívoca de imparcialidad y de no corrupción al asignar los fondos disponibles para proyectos de I+D. Además, la selección de proyectos se enfrenta a problemas relevantes. Dos son los problemas principales. El primero hace referencia a cómo medir el posible rendimiento de los proyectos de I+D (potencial de proyecto). El segundo se refiere a cómo optimizar el proceso de selección entre proyectos con múltiples y (a veces) incomparables medidas de resultados. En este trabajo se dispone de una base de datos y se propone una metodología para solventar ambos problemas. Utilizando la técnica de la envolvente de datos (DEA) evalúa la eficiencia de los evaluadores al seleccionar entre proyectos de I+D presentados a una agencia de desarrollo regional para obtener financiación pública. La base de datos disponible nos permite definir de forma precisa para cada proyecto *inputs* y *outputs* que se pueden utilizar como sistema de selección para filtrar aquellos proyectos que se podrían considerar relativamente más (o menos) deseables en comparación con el resto. Además, y dado que se dispone de las puntuaciones que otorgan los evaluadores (contratados por la agencia pública para seleccionar proyectos) a cada proyecto podemos evaluar la actuación de los propios evaluadores. En otras palabras, se evalúa la actuación de los evaluadores mediante la comparación con una metodología alternativa y objetiva (DEA) ampliamente utilizada en el sector privado para la selección de proyectos de I+D.

Keywords: subvenciones, I+D, DEA, *Multicriteria decision analysis*, evaluadores

JEL: O32, C61, H25

* Queremos manifestar nuestro agradecimiento al *Centre d'Innovació i Desenvolupament Empresarial* (CIDEM) y, en particular a su Servicio de estudios, por la información facilitada y por su colaboración en el tratamiento y revisión de la misma. José García Quevedo agradece el apoyo del MEC (SEJ2007-65806) y de la Generalitat de Catalunya (2005SGR00285).

¹ Departament d'Econometria i Economia Espanyola. Universitat de Barcelona. Av Diagonal 690. 08034 Barcelona. Tel.: +34-93-4034729 Fax: +34-93-4037242.

² Institut d'Economia de Barcelona (IEB)

³ Departament d'Hisenda Pública. Universitat de Barcelona .Av Diagonal 690. 08034 Barcelona. Tel.: +34-93-4021812 Fax: +34-93-4021813

“I think that the World market for computers is for no more than five computers”, Chief Executive Officer of IBM, 1958.

1. Introducción

Las aportaciones de la teorías del crecimiento endógeno (Romer, 1990) y la evidencia empírica (Griliches, 1992) han destacado la relación existente entre la innovación y el crecimiento económico. Actualmente, el fomento de la I+D y la innovación ocupa un lugar de especial relevancia en las políticas públicas de los países desarrollados. En particular, la Unión Europea se propuso, en la cumbre de Lisboa, avanzar en la sociedad del conocimiento y alcanzar un nivel de gasto en I+D sobre el PIB del 3% en 2010. En España, también destacan, en los últimos años, los esfuerzos en el fomento de la innovación. En los países de la OCDE y de la UE-27, los gobiernos financian directamente en torno al 7% de los gastos privados en I+D (OCDE, 2007). En España este porcentaje se sitúa, con datos de 2005, en el 13,6%, cuatro puntos porcentuales más que en 2001 (OCDE, 2007).

El apoyo público a las actividades privadas de I+D se apoya en argumentos sobre los que existe un amplio consenso. En concreto, la razón fundamental para la intervención pública es la existencia de fallos de mercado, que comportan que la inversión privada en I+D se situaría, sin la intervención pública, por debajo de lo que sería socialmente óptimo (Arrow, 1962; Klette *et al.*, 2000). Con el objetivo de estimular la inversión privada en I+D, los gobiernos y las agencias públicas utilizan diversos instrumentos como son los incentivos fiscales a la I+D o el apoyo a la creación de empresas de base tecnológica. En particular, un instrumento de uso muy extendido son las subvenciones a los proyectos empresariales de I+D. Aunque las razones para el apoyo público están bien establecidas y existe un cierto consenso sobre la conveniencia de aumentar los recursos públicos, es necesario examinar el grado de eficiencia de la intervención pública.

En los últimos años, la evaluación de las políticas públicas ha adquirido una importancia creciente, con el objetivo de determinar su impacto y examinar si generan un efecto adicional. En el ámbito concreto de las subvenciones a la I+D, y aunque los

resultados nos son concluyentes, los estudios existentes (David *et al.*, 2000; García Quevedo, 2004) apuntan a la existencia de un efecto adicional. Estas son, sin embargo, evaluaciones de carácter *ex – post*, de los proyectos seleccionados y subvencionados por la agencia pública. En este sentido, el grado de adicionalidad e, incluso su posible existencia, está estrechamente relacionado con la evaluación *ex – ante* y posterior selección de proyectos para ser subvencionados, aspecto escasamente estudiado.

Las agencias públicas, tanto en Estados Unidos como en Europa y en España, utilizan como método de selección y asignación de subvenciones a los proyectos empresariales de I+D, mayoritariamente el sistema “*peer review*” o de revisión por pares o expertos, método que, aunque presenta ventajas, no está exento de críticas, tal como se presenta posteriormente. En cambio, las empresas privadas para seleccionar y establecer un ranking de sus propios proyectos de I+D, con el propósito de optimizar su inversión en I+D, utilizan preferentemente métodos más cuantitativos como es el caso del *Multiple Criteria Decision Making* (MCDM) o del *Data Envelopment Analysis* (DEA).¹

El objetivo de este trabajo es examinar diferentes alternativas o métodos para la evaluación y asignación de subvenciones a los proyectos empresariales de I+D. Aunque existen diferencias entre los objetivos de la agencia pública y de las empresas privadas, en ambos casos, la evaluación de proyectos trata de garantizar que se seleccionen proyectos con un alto potencial para generar beneficios económicos. En consecuencia, conviene contrastar si existen metodologías preferibles en la selección de proyectos de I+D. Este aspecto, como se ha señalado, ha sido muy poco tratado en la literatura debido, en gran medida, a las limitaciones de información. Los estudios existentes sobre la participación de las empresas en los programas públicos de I+D se han centrado fundamentalmente sobre las características de las empresas que inciden en la probabilidad de obtener una subvención (Blanes y Busom, 2004, Herrera y Heijis, 2007) al no disponer de información de los proyectos individuales.

La disponibilidad, en nuestro caso, de una base de datos de proyectos de I+D, que han recibido tanto subvención como no, con información individual de los mismos y con la valoración que le otorgan los expertos externos permite examinar si los evaluadores está

¹ Ver Linton *et al.* (2002) para un breve resumen de las principales metodologías utilizadas en el sector privado para escoger proyectos de I+D en grandes *portfolios*.

escogiendo los proyectos más eficientes o, en otras palabras, permite evaluar a los evaluadores. Para lograr este objetivo se llevan a cabo dos etapas. En una primera se comparan las puntuaciones que los evaluadores otorgan a cada proyecto con los que les otorga el método del DEA. En una segunda fase se analiza qué distribución de proyectos consigue alcanzar los objetivos que la agencia establece para dichas ayudas a la I+D.

Al margen de esta introducción, el resto del artículo se organiza de la siguiente manera. En la sección 2 se examinan los distintos métodos de selección de proyectos de I+D y se discuten sus ventajas e inconvenientes. La sección 3 presenta la base de datos disponible y se lleva a cabo un análisis aplicado, con el método DEA, lo que permite contrastar los distintos métodos de selección de proyectos y, en particular, en relación al método del “*peer review*” utilizado por la agencia pública. La sección 4 presenta una segunda fase de estudio que analiza si existe un impacto diferencial de los proyectos seleccionados por la agencia de aquellos que se seleccionarían con el método alternativo (DEA) propuesto en este trabajo. Finalmente en la sección 5 se presentan las principales conclusiones a la que llega el estudio.

2. Selección de proyectos de I+D

La definición de un programa público de subvenciones a proyectos empresariales de I+D comporta definir un sistema de selección y ordenación de proyectos para decidir cuáles han de ser subvencionados y el importe de la subvención. Los criterios de decisión de la agencia deben formar parte de la evaluación de un programa público de subvenciones, tal como se pone de manifiesto en los modelos estructurales en este ámbito (David *et al.*, 2000), ya que las decisiones de la agencia en la selección de proyectos tienen una influencia determinante en los resultados del programa y en los efectos de las subvenciones.

Las agencias públicas establecen, en sus convocatorias de subvenciones, con un grado mayor o menor de precisión, los criterios en los que se basa la selección de proyectos que presentan las empresas para una posible subvención. En principio, estos criterios deben estar en relación con el objetivo de superar los fallos de mercado de modo que las subvenciones a las empresas generen, en presencia de “*spillovers*”, los incentivos

suficientes para lograr una asignación eficiente de recursos. Sin embargo, de modo frecuente, otros objetivos también están presentes en los criterios de decisión de la agencia (Blanes y Busom, 2004). Estos objetivos pueden ser, a modo de ejemplo, el apoyo a tecnologías o sectores específicos, el fomento de proyectos con mayor capacidad de arrastre e impacto económico, o el apoyo prioritario a proyectos que supongan cambios en el comportamiento de las empresas y que generen “*behavioural additionality*” (OCDE, 2006), mediante el fomento, por ejemplo, a la cooperación entre empresas y universidades. Además, habitualmente, diferentes niveles de administraciones como, por ejemplo, la Unión Europea, las Administraciones centrales y las agencias regionales, disponen de programas de fomento de la I+D. Ello exige un cierto grado de coordinación y evitar solapamientos con lo que, a menudo, las agencias regionales, que disponen de un volumen menor de recursos, se orientan más al fomento de proyectos más próximos a la innovación que a la investigación básica y con más posibilidades de aplicación comercial.

La selección y ordenación de proyectos para la posterior asignación de subvenciones se enfrenta a dificultades sustanciales debido a los problemas de información asimétrica y, en general, a la limitada información de la dispone la agencia, particularmente en el caso de los proyectos de I+D caracterizados por un grado elevado de incertidumbre. En términos generales, el proceso de selección comporta valorar el grado de contribución del proyecto al avance tecnológico, su potencial económico y la conveniencia de un apoyo público para el desarrollo del proyecto (Feldman y Kelley, 2003). Para ello, el método mayoritariamente utilizado tanto en Estados Unidos, la Unión Europea, como en la mayoría de agencias nacionales y regionales es el uso de expertos externos o el “*peer review*”. En particular, la Unión Europea, en el VII Programa Marco de I+D, destaca que la evaluación por expertos es la base del sistema de selección y de decisión de concesión de apoyo financiero. También la Academia Nacional de las Ciencias de Estados Unidos considera que el modo más efectivo para evaluar proyectos de I+D es el “*peer review*”. El uso del “*peer review*” tiene distintas modalidades pero, en general, el método más común es que los expertos externos valoren las propuestas de las empresas, otorgándole una puntuación y que un comité organizado por la agencia, a partir de estas valoraciones, seleccione y establezca una ordenación de los proyectos (Jaffe, 2002).

El “*peer review*” presenta ventajas sustanciales, particularmente en términos de independencia, imparcialidad y transparencia en el proceso de evaluación y selección de proyectos. Además permite disponer del conocimiento de expertos en la valoración de proyectos en un ámbito, como es el caso de la tecnología, con un grado elevado de complejidad e incertidumbre, en el que difícilmente la agencia puede contar con expertos propios. Sin embargo, también presenta limitaciones y ha sido objeto de críticas.

En el ámbito específico de la subvención a los proyectos empresariales de I+D, las principales críticas son los denominados sesgo conservador y sesgo institucional (Brezis, 2007). El sesgo conservador se refiere a la tendencia a rechazar proyectos que puedan suponer innovaciones sustanciales pero que están lejos de ser aplicables, y que presentan, en consecuencia, dificultades sustanciales de evaluación. Esta tendencia que favorece la aprobación de proyectos convencionales puede tener, como destaca Brezis (2007), efectos negativos sobre el impacto de las subvenciones en la productividad. Por su parte, el sesgo institucional, en este ámbito, comporta que las empresas de cierto tamaño y con experiencia en actividades de I+D se vean favorecidas en la concesión de subvenciones. A estas limitaciones, se añade la subjetividad de los expertos, la dificultad a la que se enfrenta la agencia para identificar expertos adecuados y los posibles problemas de conflicto de intereses sobre todo en ámbitos de dimensión territorial reducida como es el caso de las regiones (Rigby, 2002). Además, el uso de expertos externos puede comportar costes sustanciales, tanto por el tiempo requerido para el proceso de evaluación, como derivados de la remuneración a expertos. En el caso de la Unión Europea, que dispone de una base de datos de 50.000 expertos para sus programas de I+D, se solicitan en torno a 5.000 evaluaciones anuales, con un coste estimado próximo al 2% del presupuesto total del programa.

Por su parte, las empresas privadas también se enfrentan a la necesidad de valorar y establecer un ranking de su cartera de proyectos de I+D. Tal como se pone de manifiesto a partir de un modelo de inversión empresarial en I+D, el nivel de gasto óptimo en I+D tiene lugar cuando el rendimiento marginal de la inversión en I+D coincide con el coste marginal. La curva del rendimiento marginal de la inversión en I+D tiene una pendiente decreciente dado que a medida que se invierte más en I+D, los proyectos tienen una rentabilidad menor. Ello comporta para las empresas una

valoración y ordenación de sus proyectos en función de la tasa esperada de retorno de cada proyecto. Para ello se han propuesto numerosos métodos (Linton *et al.*, 2002) sin que exista una solución definitiva. Como destacan Linton *et al.* (2002) existen dos problemas sustanciales como son la medición de los rendimientos y potencial de la cartera de proyectos de los que dispone una empresa y optimizar la selección entre un amplio conjunto de proyectos que presentan, en general, dificultades de comparación. La opción más adecuada es disponer de medidas cuantitativas para lo que se han propuesto diversas posibilidades como es el valor actual neto (VAN) u otras medidas financieras. Sin embargo, todas estas aproximaciones son insuficientes dado que no capturan la complejidad de los proyectos de I+D. Por ello se han propuesto métodos cuantitativos alternativos entre los que destacan el uso del DEA que puede ser particularmente adecuado para seleccionar y clasificar proyectos como los de I+D caracterizados por un alto grado de incertidumbre (Linton *et al.*, 2002; Linton *et al.*, 2007).

El método DEA presenta como principales ventajas su capacidad de analizar simultáneamente múltiples *inputs* y múltiples *outputs*; no requiere además suponer formas funcionales específicas que relacionen *inputs* con *outputs*, permite comparar las unidades de análisis entre ellas o con un subconjunto y finalmente, los *inputs* y *outputs* pueden estar medidos en unidades diferentes como, por ejemplo, cantidades físicas y monetarias. Frente a estas ventajas también presenta limitaciones como son que el DEA es una técnica de punto extremo con lo que los errores de medida pueden causar problemas significativos y que permite obtener medidas de la “eficiencia relativa” pero no de la “eficiencia absoluta”, con lo que no es posible comparar los resultados con un “máximo teórico”. Además, el DEA es una técnica no paramétrica con lo que las pruebas de hipótesis estadísticas son difíciles de implementar. Vale la pena destacar que la característica del DEA de ofrecer resultados de eficiencia relativa se ajusta a los objetivos de este trabajo, pues el objetivo de utilizar el DEA es escoger de entre todos los proyectos presentados aquellos que declaran conseguir un máximo *output* dado un nivel de *input*. En este caso los resultados del DEA no deben interpretarse en valores absolutos de qué proyecto tienen un mejor resultado por lo que respecta, por ejemplo, a las variables objetivo de la convocatoria, sino que los resultados del DEA deben interpretarse como una manera de ordenar los proyectos presentados y de poder

compararlos entre ellos dados unos *inputs* y unos (múltiples) *outputs* determinados por la agencia.

3. Base de datos disponible y análisis DEA

Uno de los atractivos del presente trabajo es explotar la base de datos detallada de las puntuaciones otorgadas por los evaluadores a los proyectos de I+D presentados a la Agencia catalana de innovación y desarrollo empresarial (CIDEM) que cada año abre una convocatoria de ayudas para promover la realización de proyectos en I+D por parte de las empresas establecidas en el territorio catalán. La Agencia, en general, persigue varios objetivos que se definen en diversas líneas de actuación o de subvención. Se utiliza como caso de estudio la convocatoria de proyectos estratégicos de I+D para el año 2005. Estas ayudas subvencionan proyectos que cumplan una serie de requisitos que satisfagan los objetivos concretos que persigue la Agencia con esta línea de ayuda. Los requisitos que se valoran por parte de la Agencia en el caso de los proyectos estratégicos son establecidos en la convocatoria pública de subvención (tabla 1).

En la tabla 1 se presentan los criterios de selección que establece la Agencia, así como las variables específicas que la propia Agencia determina para la evaluación de los proyectos. Todas las empresas que desean presentar un proyecto de I+D para su posible financiación por parte de la Agencia han de rellenar un impreso de solicitud.² En este impreso de solicitud se pide a la empresa una serie de información tanto de la propia empresa como del proyecto que presenta. Esta información (cualitativa y cuantitativa) es la que posteriormente se utiliza por los evaluadores para valorar los proyectos y que nosotros utilizamos en nuestras estimaciones. En nuestro caso hemos escogido aquellos criterios que son evaluables con variables cuantitativas de la solicitud, eliminando del estudio las valoraciones más cualitativas del evaluador. Así, las puntuaciones de los proyectos que se obtienen utilizando la metodología DEA sólo hacen referencia a objetivos cuantificables y se contraponen con las puntuaciones de los evaluadores para esos objetivos concretos.

² Vale la pena destacar el importante esfuerzo realizado por la Agencia por establecer un sistema de recogida de información cuantitativa (y cualitativa) de las empresas (y los proyectos) que se presentan para permitir mejorar tanto la asignación como la evaluación del impacto de las ayudas públicas que otorgan.

Tabla 1. Criterios de selección de proyectos estratégicos establecidos por la Agencia.

Criterios generales de la Agencia	Variable específica
Criterio 1: Viabilidad técnica del proyecto.	<ul style="list-style-type: none"> • Contribución tecnológica del proyecto.
Criterio 2: Impacto socioeconómico.	<ul style="list-style-type: none"> • Impacto sobre el VAB. • Impacto sobre el empleo.
Criterio 3: Contribución a la internacionalización de la economía.	<ul style="list-style-type: none"> • Contribución a la internacionalización. • Proyectos de crecimiento industrial con vocación innovadora.
Criterio 4: Incidencia de los resultados del proyecto en el propio sector de actividad.	<ul style="list-style-type: none"> • Participación empresas cadena de valor. • Incidencia del proyecto sobre resultados del sector.
Criterio 5: Aumento de la capacidad de I+D del solicitante.	<ul style="list-style-type: none"> • Impacto del proyecto sobre el incremento del I+D de la empresa.
Criterio 6: consolidación de la empresa en el mercado.	<ul style="list-style-type: none"> • Inversiones futuras derivadas del proyecto.

Fuente: Elaboración propia a partir de la orden TRI/163/2005 de 13 de abril, DOGC 4369.

3.1. Definición de *inputs* y *outputs*

La base de datos consta de información completa para 148 proyectos (después de eliminar todos aquellos proyectos para los cuales falta información relevante) presentados en el 2005. Vale la pena destacar que los datos procedentes de las solicitudes de los proyectos incluyen información no sólo para el año en que se pide la ayuda (2005) sino que para muchas variables se piden datos de la situación anterior al proyecto (2004) como de previsiones para el año posterior (2006). Esto permite llevar a cabo ejercicios de robustez de las estimaciones.

Una de las claves de la aplicación de la metodología del DEA es la definición de los *inputs* y de los *outputs*. En el caso de la selección de proyectos de I+D, y siguiendo Linton *et al.* (2002, 2007) se considera el *input* del proyecto el coste del mismo. La base

de datos disponible permite identificar 11 *outputs* que cuantifican los objetivos de la Agencia:

- *Output* 1: % de actividad subcontratada en investigación 2005-2006.
- *Output* 2: % de actividad subcontratada en desarrollo tecnológico 2005-2006.
- *Output* 3: % incremento gastos en I+D en relación al año anterior 2005-2006.
- *Output* 4: Previsión de nuevas actuaciones I+D inducidas por los resultados del proyecto 2005-2006 (presupuesto estimado en €).
- *Output* 5: Inversiones industriales inducidas como resultado del proyecto 2005-2006 (presupuesto estimado en €).
- *Output* 6: Creación de nueva ocupación directa en I+D derivada del proyecto 2005-2006 (número de personas).
- *Output* 7: Creación de nueva ocupación total derivada del proyecto 2005-2006 (número de personas).
- *Output* 8: Previsión de incremento de la cuota de mercado global derivada del proyecto 2005-2006 (%).
- *Output* 9: Previsión de contribución del proyecto al margen global de explotación 2005-2006 (%).
- *Output* 10: Previsión de contribución del proyecto a las exportaciones 2005-2006 (%).
- *Output* 11: Número de patentes propias que se prevé generar como resultado del proyecto 2005-2006.

Estos *outputs* son utilizados por los evaluadores de la Agencia para puntuar los proyectos recibidos. La tabla 2 presenta las principales estadísticas descriptivas de los 148 proyectos analizados.

El método DEA se usa normalmente para evaluar la eficiencia de las denominadas unidades de toma de decisiones (DMU),³ como pueden ser empresas, hospitales, escuelas, universidades, agencias no gubernamentales o cualquier otro tipo de organización. El DEA es un método de punto extremo y compara cada DMU sólo con la mejor DMU. El supuesto fundamental detrás del método DEA es que si, por ejemplo,

³ De la siglas en inglés *Decision Making Unit*.

una DMU dada, A, es capaz de obtener un *output* Y(A) con X(A) *inputs*, entonces otras DMU también deben poder hacerlo si operasen eficientemente. De igual manera, si una DMU B es capaz de obtener Y(B) unidades de *output* con X(B) *inputs*, entonces otros productores también deben ser capaces de obtener el mismo *output*. Las DMU A, B, y otras pueden combinarse para formar a una DMU “compuesta” con bs *inputs* y *outputs* compuestos. Dado que esta DMU compuesta no existe, se le denomina una DMU virtual. La base del método DEA reside en encontrar a la mejor DMU virtual para cada DMU real. Si la DMU virtual es mejor que la DMU real ya sea porque obtiene más *output* con los mismos *inputs* o bien porque obtiene el mismo *output* con menos *inputs*, entonces la DMU real es ineficiente.

Taula 2. Estadística descriptiva de los *inputs* y *outputs* utilizados

	Media	Error típico	Mediana	Desviación estándar	Mínimo	Máximo
<i>Input</i>	848,555.5	147,277.7	398,107.6	1,791,711.0	60,269.8	15,806,918
<i>Output 1</i>	0.14	0.01	0.08	0.17	0.00	0.85
<i>Output 2</i>	0.13	0.02	0.05	0.18	0.00	0.80
<i>Output 3</i>	0.20	0.02	0.10	0.27	0.00	1.00
<i>Output 4</i>	355,163.0	112,971.5	84,326	1,374,357.9	0.00	16,000,000
<i>Output 5</i>	301,183.1	63,696.5	69,620	774,902.1	0.00	6,000,000
<i>Output 6</i>	1.52	0.15	1.00	1.79	0.00	10.00
<i>Output 7</i>	2.80	0.27	2.00	3.24	0.00	25.00
<i>Output 8</i>	0.12	0.02	0.05	0.22	0.00	1.00
<i>Output 9</i>	0.14	0.02	0.10	0.20	0.00	1.00
<i>Output 10</i>	0.17	0.02	0.05	0.27	0.00	1.00
<i>Output 11</i>	0.76	0.13	0.00	1.52	0.00	10.00

Nota: las estadísticas descriptivas para los 148 proyectos corresponden a los valores de los *outputs* para el año 2006.

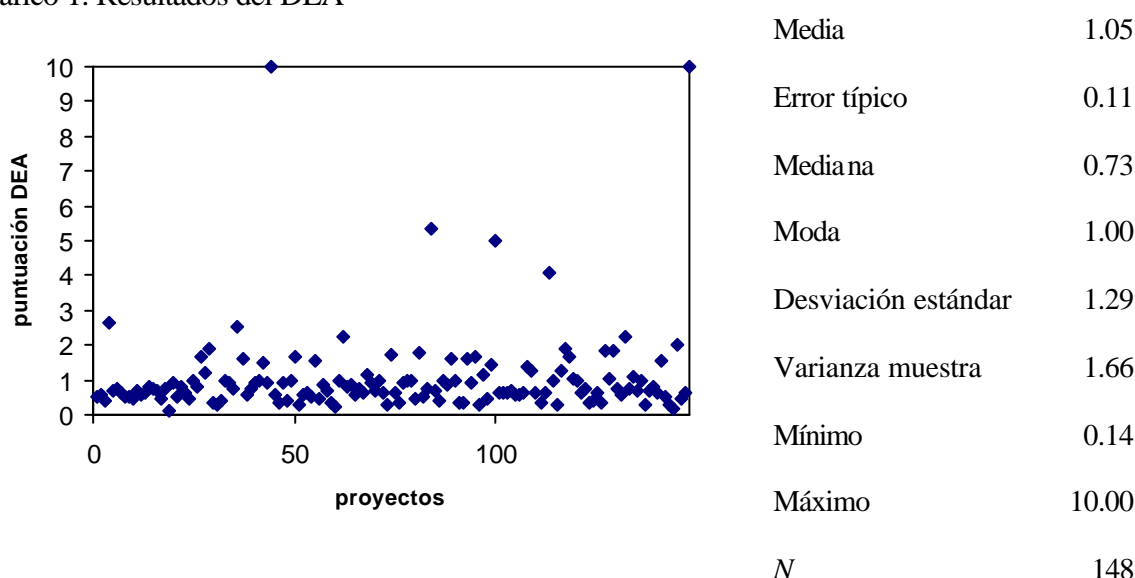
Una vez definidos *inputs* y *outputs* se estima una función de producción con rendimientos variables a escala (VRS) asumiendo que se maximiza el *output* dado un nivel determinado de *input*. La hipótesis de rendimientos constantes a escala parece del todo adecuada en nuestro caso dado que se comparan proyectos con dimensiones muy distintas entre ellos, lo que afectaría de forma clara los resultados del DEA en caso de utilizar rendimientos constantes a escala (CRS). Los resultados del DEA que a continuación se presentan (puntuaciones por proyecto o índices de eficiencia relativa) deben ser interpretados como una forma de poder ordenar los proyectos y poder seleccionar aquellos que permiten obtener un máximo *output* (objetivo de la agencia) dado un nivel dado de coste (*input*). Como es bien sabido, los índices de eficiencia están acotados entre 0 y 1, siendo el índice de eficiencia de los mejores proyectos igual a 1.

En nuestro caso esta característica de los resultados tradicionales del DEA lleva consigo un problema, no podemos ordenar entre los proyectos que el DEA considere más eficientes. Para solucionar este problema se utiliza la *superefficiency*, opción de la estimación DEA que permite asignar valores superiores a 1 a proyectos eficientes, reduciéndose así el número de proyectos con un mismo valor y permitiendo una ordenación más precisa de los proyectos. Cuando se utiliza la *superefficiency* los índices de eficiencia del DEA se distribuyen entre 0 y 10.

3.2. Principales resultados: valoración y ordenación de los proyectos según el DEA

En esta sección se presentan los principales resultados que se obtiene de utilizar la técnica DEA para los *inputs* y *outputs* de los proyectos definidos anteriormente. Por motivos de espacio no se presentan las puntuaciones (índices de eficiencia) que asigna el DEA a cada uno de los 148 proyectos. En su lugar se presenta un resumen de dichas puntuaciones en el gráfico 1 para el caso de la estimación DEA para 1 *input* (coste del proyecto) y los 11 *outputs* presentados en la tabla 2 asumiendo una tecnología con VRS y permitiendo que las puntuaciones de la eficiencia relativa de cada proyecto sean superiores a 1.

Gráfico 1. Resultados del DEA



Nota: resultados del DEA para los valores de los *outputs* en el año 2006

Puede observarse como los índices de eficiencia relativa de los proyectos que calcula el DEA nos permite hacer una ordenación precisa de los proyectos. Es precisamente esta

ordenación la que se compara con la que se obtiene del ‘*peer review*’, es decir, de las puntuaciones de los evaluadores de los proyectos.

Un primer ejercicio es comparar el grado de similitud entre los resultados del DEA y de los evaluadores. Para ello se calcula el coeficiente de correlación entre los resultados del DEA y las puntuaciones de los evaluadores de la Agencia.

Tabla 3. Coeficientes de correlación entre las puntuaciones del DEA (2 casos) y la de los evaluadores.

	DEA 1 <i>input</i> – 11 <i>outputs</i> 2006	DEA 1 <i>input</i> – 11 <i>outputs</i> 2005	Puntuaciones evaluadores
DEA 1 <i>input</i> – 11 <i>outputs</i> 2006	100%		
DEA 1 <i>input</i> – 11 <i>outputs</i> 2005	83%	100%	
Puntuaciones evaluadores	16%	14%	100%

Nota: se presentan los resultados del DEA para los 11 *outputs* en el año 2005 y 2006.

Se observa como el grado de similitud entre las puntuaciones otorgadas por el sistema del ‘*peer review*’ y el sistema del DEA es bastante bajo, y el grado de coincidencia es sólo del 14%-16% lo que nos indica que ambas metodologías de selección de proyectos dan resultados dispares, abriéndose claramente la incógnita de qué metodología es superior a la otra. Antes de abordar este importante punto con más detalle, la tabla 4 nos da también información sobre el grado de similitud entre ambas distribuciones de puntuaciones. Así, se ha establecido el criterio (totalmente *ad hoc* pero útil a efectos ilustrativos) que en ambos casos (DEA y evaluadores) se escogen los 50 mejores proyectos. La tabla 4 nos informa del grado de coincidencia entre los proyectos aprobados o no en cada caso.

La mitad de los proyectos que los evaluadores consideran como eficientes también obtienen una alta puntuación por parte del DEA. En el caso de los proyectos no aprobados por la Agencia hay una coincidencia del 73% con los resultados del DEA (72 de 98) mientras que en 26 casos se denegaría la ayuda en función de la puntuación del

evaluador mientras que se aprobaría si se utilizasen las puntuaciones del DEA. En resumen, y como medida global de similitud entre las metodologías se observa como evaluadores y DEA coinciden en un 65% en la valoración (aprobar/denegar) de los proyectos, mientras que en un 35% de los casos un criterio simple de selección como el utilizado en este ejemplo (los 50 mejores proyectos) lleva a que un 35% de los proyectos difieren en la valoración final otorgada por uno u otra metodología.

Tabla 4. Asignación de proyectos (aprobados/denegados) según metodología

	Aprobado DEA	Denegado DEA	Total
Aprobado Evaluador	24	26	50
Denegado Evaluador	26	72	98
Total	50	98	148

Nota: Se utilizan los resultados del DEA con 1 *input* y los 11 *outputs* del 2006 y con VRS.

Los resultados son poco concluyentes. El análisis muestra que la utilización del “*peer review*” o un método objetivo como la selección de proyectos a través de la técnica del DEA lleva a resultados diferentes.

4. ¿“*Peer review*” o DEA? Una segunda etapa en el proceso de selección de proyectos.

Los resultados de la sección anterior muestran un reducido grado de coincidencia entre las puntuaciones de los proyectos seleccionados por cada método de evaluación *ex ante*. Lo anterior pone de manifiesto la necesidad de comparar los resultados sociales esperados de cada conjunto de proyectos para poder determinar, de forma informada, si alguno es socialmente preferible al otro, siempre en función de los objetivos que persigue la agencia pública.

Cada método de selección de proyectos presenta fortalezas y debilidades. La determinación de la preferencia de uno respecto a otros, en cuanto a las subvenciones a la I+D se refiere, debe depender de los objetivos que persiga la agencia pública encargada de su concesión. Si lo que se pretende es generar el máximo impacto posible, entendido como la consecución de una mayor rentabilidad social derivada de la puesta en marcha de dichos proyectos, entonces es necesario comparar la adicionalidad que se

genera en cada uno de los escenarios ya que se preferirá aquel mecanismo que genere el mayor impacto.

Para ello, se dispone de una situación real, los proyectos efectivamente subvencionados por la Agencia través de las puntuaciones de los evaluadores externos y es posible simular una situación virtual, escogiendo los proyectos más eficientes identificados mediante el método DEA presentado anteriormente. Dado que las empresas indican las expectativas que tienen sobre variables referidas a las actividades de I+D que realizarán a partir de los proyectos, es posible comparar si la adicionalidad esperada es mayor en el caso de los proyectos seleccionados por el "peer review" implantado por la Agencia o bien aquellos identificados como más eficientes mediante el método DEA.

El procedimiento ideal por poder llevar a cabo una evaluación sería observar a la misma empresa en dos situaciones diferentes en el mismo momento: analizar sus resultados con ayuda y sin ayuda. Dado que esto es imposible, se debe buscar metodologías que permitan controlar el estado contrafactual, es decir, el que hubiera ocurrido en el caso contrario al que realmente se encuentra la empresa. La naturaleza de la mayoría de metodologías se basa en la utilización de dos grupos de empresas. El primero de ellos incluye las empresas que han recibido una ayuda y el segundo lo forman empresas que no la han recibido. Para que la evaluación funcione, las empresas de los dos grupos deberán ser idénticas en todas sus características, a excepción de la ayuda. Como esto tampoco es posible, se debe construir un grupo de control que sea lo más similar posible al grupo de tratamiento tanto en relación a las características no observables como a las observables.

Para tener en cuenta estos problemas, es posible aplicar un método de asociación no paramétrico denominado *Propensity Score Matching* (PSM). A partir del trabajo de Rosenbaum y Rubin (1983), el uso del PSM ha sido ampliamente utilizado en la evaluación de las políticas públicas. En el caso de los estudios de evaluación de las ayudas dirigidas a las empresas, este método se ha utilizado de forma recurrente en los últimos años (por ejemplo Duch et al, 2006; Herrera y Heijs, 2007; Almus y Czarnitski, 2003). El PSM es un método que permite estimar el efecto promedio de la ayuda sobre las unidades que la reciben (el denominado *Average Effect of Treatment of the Treated*

o ATT) observando la variable resultado expuesta al tratamiento (Y1) o estado factual frente a la no exposición (Y0) o estado contrafactual.

A partir de los datos disponibles se identifican las empresas que han solicitado y recibido una ayuda y las que la han solicitado pero no la han recibido, siendo las empresas de este último grupo las candidatas ideales para el grupo de control. Una vez realizado el ejercicio de evaluación *ex – ante* utilizando el modelo DEA, es posible saber cuales son los proyectos más eficientes y a que empresa corresponden. Así, es posible simular un escenario en el que los proyectos más eficientes fuesen subvencionados. Se supondrá por falta de datos en el tiempo de los resultados de las actividades de I+D que las empresas incorporan a sus decisiones estratégicas el hecho de conocer si se le ha otorgado una ayuda o no, aún cuando los pagos correspondientes no se hayan efectuado. Como resultado, el comportamiento de cada empresa se ve inmediatamente influenciado por la decisión de la agencia pública de concederle una ayuda.

A pesar de su creciente popularidad, relativamente poco se ha escrito sobre el problema de la selección de variables en el caso del PSM. Algunos estudios basados en simulaciones (Brookhart et al. 2006, Judkins et al. 2007) ilustran como la selección de variables que se incluyen en el cálculo del indicador de propensión (*Propensity Score* o PS) pueden afectar el sesgo, la varianza y el error cuadrático medio de un efecto estimado del ATT. Los resultados de estos estudios sugieren que las variables no relacionadas con el tratamiento (subvención) pero relacionadas con el resultado deberían incluirse siempre en el cálculo del PS. En el caso contrario, la inclusión de variables relacionadas con la exposición pero no con el resultado, tiende a aumentar la varianza del ATT sin reducir el sesgo.

Una estimación del PS no es suficiente para estimar el efecto de interés. La razón es que la probabilidad de observar dos empresas con exactamente el mismo PS es en principio cero. Se han propuesto diferentes métodos en la literatura por superar este problema y cuatro de los estimadores más ampliamente utilizados son el vecino más próximo, el radio, el kernel y la estratificación.

Una forma de emparejar unidades tratadas y de control consiste en coger cada unidad tratada y buscar la unidad de control con el PS más parecido, es decir, el vecino más próximo (*nearest neighbor matching*, NNM). Aunque no es necesario, el método se aplica normalmente con reemplazo, en el sentido que una unidad del grupo de control puede ser la mejor pareja para más de una unidad con tratamiento. Una vez cada una de las unidades tratadas se empareja con una unidad de control, se calcula la diferencia entre el resultado de las unidades tratadas y el resultado de las unidades del grupo de control. El ATT de interés se obtiene entonces como la mediana de estas diferencias.

Sin embargo, es obvio que algunos de estos emparejamientos son bastante pobres porque para algunas unidades tratadas los vecinos más próximos pueden tener un PS muy diferente, contribuyendo a la estimación del ATT con independencia de esta diferencia. Con el estimador kernel todas las unidades del grupo de tratamiento se emparejan con una media ponderada de todas las unidades del grupo de control con pesos que son inversamente proporcionales a la distancia entre los PS de las tratadas y las del grupo de control. Está claro de las consideraciones anteriores que estos cuatro métodos pueden conducir a resultados diferentes, sobre todo en cuanto al *trade-off* entre la calidad y la cantidad de los emparejamientos y ninguno de ellos es a priori superior a los otros. Su consideración conjunta, no obstante, ofrece una manera de analizar la robustez de las estimaciones.

Las tablas siguientes muestran los resultados de la evaluación *ex – post* bajo los dos escenarios descritos, la selección de proyectos a través del “*peer review*” de la Agencia (ATT-Agencia) y la elección de los proyectos más eficientes a través del método DEA (ATT-DEA) utilizando solo dos de los estimadores descritos, el NNM y el kernel. Para este ejercicio de evaluación, se ha seleccionado una variable que capture la adicionalidad financiera, como es la intensidad en el gasto en I+D (gasto en I+D sobre ventas) y otra variable que mide la adicionalidad de comportamiento, aquí aproximada por el porcentaje de actividades subcontratadas de investigación.

En el primer caso, señalado en la tabla 5, se aprecia que existe un efecto de adicionalidad financiera positivo y significativo derivado de los proyectos seleccionados a través del método DEA mientras que, a pesar de obtenerse un valor positivo en el caso

de los proyectos seleccionados por la Agencia, éste no es estadísticamente significativo⁴.

Tabla 5.

	ATT-Agencia	ATT DEA	Tratadas	Control
NNM	6,8 (1,050) [1,452]	8,1 (2,016**) [2,053**]	62	31
Kernel	6,5 (-) [1,214]	9,3 (-) [2,751***]	62	41

Notas: Errores estándar entre paréntesis. Errores estándar por bootstrap entre corchetes. *, ** y *** indican significatividad estadística al 90, 95 y 99%, respectivamente. Elaboración propia.

La tabla 6 muestra una situación diferente. En este caso, la adicionalidad de comportamiento es positiva y significativa una vez más en el caso de las empresas cuyos proyectos de I+D han sido seleccionados por el DEA como los más eficientes, mientras que en el caso de los proyectos seleccionados por los evaluadores externos de la Agencia, el efecto es negativo pero no es estadísticamente significativo. En este caso, resultados anteriores (García-Quevedo et al., 2007) indicaban la no existencia de adicionalidad de comportamiento en una muestra más amplia de proyectos.

Tabla 6.

	ATT-Agencia	ATT DEA	Tratadas	Control
NNM	-8,8 (-1,329) [-0,868]	8,6 (2,237**) [2,259**]	62	32
Kernel	-0,4 (-) [-0,053]	7,4 (-) [2,204**]	62	41

Notas: Errores estándar entre paréntesis. Errores estándar por bootstrap entre corchetes. *, ** y *** indican significatividad estadística al 90, 95 y 99%, respectivamente. Elaboración propia.

Los resultados anteriores deben interpretarse con la cautela que la elevada sensibilidad del método PSM ante la inclusión u omisión de determinadas variables impone. Asimismo, la imposibilidad de disponer de datos antes y después de recibir la subvención condiciona los efectos ATT estimados. En cualquier caso, los resultados indican que es necesario un análisis más riguroso *ex – ante* ya que, sin él, es posible

⁴ Este resultado puede deberse a las particularidades de la (sub) muestra utilizada, ya que considerados globalmente, éstos proyectos generan una adicionalidad positiva y significativa de alrededor del 6-7%. Véase García-Quevedo et al. (2007).

estar subvencionando proyectos de I+D que no contribuyen a maximizar el rendimiento social de los recursos públicos dedicados a ello. El método DEA debe ser, en este sentido, un complemento a la evaluación realizada a través del *‘peer review’*. Además, los resultados ponen de manifiesto la importancia de vincular la evaluación *ex – ante* con la evaluación *ex – post* de los proyectos de I+D.

5. Conclusiones

El fomento de la innovación empresarial constituye un elemento central en las políticas públicas, tanto en la Unión Europea, como a escala nacional o regional. Aunque existen razones consistentes que justifican la intervención pública en este ámbito, es necesario demostrar la eficiencia de los programas públicos, lo que ha conducido a un creciente uso de análisis de evaluación, particularmente de carácter *ex –post* o de impacto para determinar si la intervención pública genera un efecto adicional.

Sin embargo, un aspecto menos estudiado es la adecuación de los sistemas de evaluación *ex – ante* para la selección y ordenación de proyectos y la posterior asignación de subvenciones. El impacto de un programa público de subvenciones está en estrecha relación con el proceso de selección de proyectos, ya que los resultados del programa, en términos de adicionalidad, dependen del éxito y la precisión en la selección de los mejores proyectos, es decir, de aquellos que cumplen en mayor grado con los objetivos del programa y que generan un impacto superior.

El objetivo de este trabajo ha sido valorar la adecuación de los sistemas de evaluación *ex – ante* de los programas de subvenciones públicas a la I+D. La evaluación *ex – ante* se enfrenta a las dificultades derivadas de la elevada complejidad e incertidumbre propia de los proyectos tecnológicos. El método mayoritariamente utilizado por las agencias públicas para la valoración de proyectos es el *“peer review”* o revisión por expertos. Este método, aunque presenta ventajas sustanciales, no está exento de críticas. La principal, desde un punto de vista del impacto tecnológico y económico, es el denominado sesgo conservador, que se define como la tendencia a aprobar, en mayor medida, proyectos más convencionales en relación a proyectos de más riesgo (Brezis, 2007). Las empresas privadas, por su parte, utilizan distintos métodos para valorar y

ordenar sus proyectos de I+D en función de los rendimientos esperados. Estos métodos tienen un carácter más cuantitativo y, entre ellos, destaca el uso del DEA.

En este trabajo se ha propuesto una metodología para valorar los sistemas de selección de proyectos de I+D y de asignación de subvenciones. Asimismo se ha ilustrado la aplicabilidad de esta metodología a partir de la información del caso de un programa de subvenciones a proyectos empresariales de I+D de una agencia regional. En síntesis, la metodología propuesta consiste en comparar los resultados de la selección y ordenación de proyectos que se establece mediante el método “*peer review*” frente a otro posible método como el DEA. Posteriormente se utilizan las técnicas cuasi-experimentales de evaluación *ex -post* como el *Propensity Score Matching*, para valorar los resultados que generan en términos de adicionalidad ambos métodos de selección de proyectos.

Los resultados muestran que la selección y clasificación de proyectos que generan el “*peer review*” y el DEA presentan un escaso grado de coincidencia. Además, la comparación de los resultados del análisis de simulación del impacto que se deriva de los proyectos seleccionados por ambos métodos, muestra la dificultad de llegar a conclusiones definitivas sobre la superioridad del “*peer review*” frente a otro método. Estos resultados deben considerarse con cautela dadas las limitaciones de los métodos de evaluación *ex -post* particularmente en lo que se refiere a la definición de grupos de control adecuados.

Aunque el método “*peer review*” puede ser preferible para evaluar aspectos cualitativos de los proyectos y para solucionar los problemas de información asimétrica a los que se enfrenta la agencia pública, los resultados muestran la conveniencia de complementarlo con otros métodos y ponen de manifiesto, especialmente, la necesidad de avanzar en el conocimiento y resultados que se derivan de los sistemas utilizados de evaluación *ex - ante*. Los avances en este tema requieren mejorar la información disponible y considerar la evaluación como un proceso integral con una vinculación estrecha entre la evaluación *ex - ante* y la evaluación *ex -post*, tal como se ha ilustrado con la metodología propuesta.

En términos generales, los resultados de este análisis apuntan, en línea con otros estudios (David *et al.*, 2000; García Quevedo, 2004), en la necesidad de mejorar los

procedimientos de evaluación de las subvenciones a la I+D y en la conveniencia de examinar distintas alternativas para la selección y posterior evaluación de proyectos por parte de la agencia pública tal como sugiere este trabajo y algunos análisis recientes (Jaffe, 2002; Brezis, 2007).

Bibliografía

- Almus, M. y D. Czarnitzki (2003): "The effects of public R&D subsidies on firms' innovation activities: the case of Eastern Germany", *Journal of Business & Economic Statistics* 21-2, 226-236.
- Arrow, K. (1962): "Economic welfare and allocation of resources for invention", en R. Nelson (ed.) *The rate and direction of inventive activity*, Princeton: Princeton University Press for the NBER, 609-625.
- Blanes, J., y Busom, I. (2004), "Who participates in R&D subsidy programs? The case of Spanish manufacturing firms", *Research Policy* 33, 1459-1476.
- Brookhart, M. A., Sebastian Schneeweiss, Kenneth J. Rothman, Robert J. Glynn, Jerry Avorn y Til Sturmer, (2006): "Variable Selection for Propensity Score Models", *American Journal of Epidemiology*, 163(12), 1149-1156.
- Brezis, E. (2007), "Focal randomization: an optimal mechanism for the evaluation of R&D projects", *Science and Public Policy*, 34 (10), 691-698
- David P, Hall B y Toole A (2000), "Is public R&D a complement or substitute for private R&D? A review of the econometric evidence", *Research Policy* 29, 497-529.
- Duch, N., Montolio, D y Mediavilla, M. (2007) Evaluating the impact of public subsidies on firm's performance: a quasi-experimental approach. *Document de treball 2007/3 IEB*.
- Feldman, M., y Kelley, M. (2003), "Leveraging research and development: Assessing the impact of the U.S. Advanced Technology Program", *Small Business Economics* 20, 153-165
- García Quevedo J, (2004), "Do public subsidies complement business R&D? A meta-analysis of the econometric evidence" *Kyklos* 57, 87-102.
- García-Quevedo, J., N. Duch, D. Montolio y J. Polo-Otero (2007): "Avaluació dels ajuts atorgats per al foment de l'R+D i la Innovació empresarial per la Secretaria d'Indústria i el CIDEM en el període 2004-2006". *Col·lecció Documents de treball 15, CIDEM*. Generalitat de Catalunya.
- Griliches, Z. (1992): "The search for R&D spillovers", *Scandinavian Journal of Economics* 94, 29-47.

- Herrera, L., y Heijs, J. (2007), “Difusión y adicionalidad de las ayudas públicas a la innovación”, *Revista de Economía Aplicada* 44, 177-197.
- Jaffe, A., (2002), “Building programme evaluation into the design of public research-support programmes”, *Oxford Economic Review* 18(1), 22-34.
- Judkins, D. D. Morganstein, P. Zador§ , A. Piesse, B. Barrett y P. Mukhopadhyay (2007), “Variable selection and raking in propensity scoring”, *Statistics in Medicine* 26, 1022-1033.
- Klette, J., Moen, J. y Griliches, Z. (2000): “Do subsidies to commercial R&D reduce market failures? Microeconomic evaluation studies”, *Research Policy* 29, 471-495.
- Linton, J.D., Walsh, S., y Morabito, J. (2002), “Analysis, ranking and selection of R&D projects in a portfolio”, *R&D Management* 32, 139-148.
- Linton, J.D., Morabito, J., y Scott Yeomans, J. (2007), “An extension to DEA support system used for assessing R&D projects”, *R&D Management* 37, 29-36.
- OCDE (2006), *Government R&D funding and company behaviour: measuring behavioural additionality*, Organización para la Cooperación y el Desarrollo Económico, París
- OCDE (2007): *Main Science and Technology Indicators 2007/2*, Organización para la Cooperación y el Desarrollo Económico, París.
- Rigby, J. (2002), “Expert panels/Peer Review” en Comisión Europea (2002), *RTD Evaluation Toolbox – Assessing the socio-economic impact of RTD policies*, IPTS Technical Report series EUR 20382 EN, Institute for Prospective Technological Studies y Joanneun Research.
- Romer, P. (1990), “Endogenous technological change”, *Journal of Political Economy*, 98(5), 71-102.
- Rosenbaum, P. y Rubin, D.B., (1983): “The Central Role of the Propensity Score in Observational Studies for Causal Effects”, *Biometrika* 70, 41-55.