

Economía de la innovación: Modelos empíricos de I+D+i empresarial

Elena Huergo

GRIPICO – Universidad Complutense de Madrid

Barge-Gil, A., Huergo, E., López, A. y Moreno, L. (2018): “Empirical models of firms’ R&D”, en L. C. Corchón y M. A. Marini (eds.): *Handbook of Game Theory and Industrial Organization*, Edward Elgar. Cheltenham, UK and Northampton, MA. In press.

- La economía de la innovación incluye un conjunto muy amplio de temas, de carácter micro y macroeconómico, que se pueden abordar tanto desde un punto de vista teórico como empírico:
 - Teoría del cambio tecnológico y su repercusión económica
 - Funcionamiento de los sistemas nacionales de innovación
 - Relaciones sector privado-sector público
 - Medición del conocimiento
 - Transferencia del conocimiento
 - Análisis empírico de los determinantes y efectos de la innovación
 -
- Los estudios empíricos sobre el nuevo conocimiento generado en las empresas son solo una pequeña parte.

- Metodología habitual para abordar los trabajos empíricos en el área:
 - ¿Cuál es la pregunta de investigación? ¿Qué motiva dicha pregunta? ¿Por qué es relevante?
 - ¿Hay alguna hipótesis concreta sobre la pregunta de investigación? ¿Qué dice la teoría?
 - ¿Hay trabajos previos sobre el mismo tema? Si es así, ¿a qué conclusiones llegaron esos trabajos? ¿Qué se aporta de nuevo con respecto a ellos en la investigación?
 - ¿De qué datos se dispone para hacer la investigación? ¿Son suficientes para contestar a la pregunta de forma precisa? ¿Qué dicen los descriptivos básicos sobre el tema de estudio?

- ¿Cuáles son las técnicas (estadística, econométrica...) que mejor se adaptan a los datos disponibles que permite contestar a la pregunta? ¿Contestar a la pregunta requiere realizar previamente una mejora en las técnicas existentes?
- ¿Cuáles son los principales resultados que se derivan del análisis empírico realizado? ¿Son los resultados sensibles a un cambio de metodología? ¿Cuál es la respuesta a la pregunta inicial?
- ¿Tienen los resultados alguna implicación de política económica?
- ¿Qué limitaciones tiene el análisis realizado? ¿Cómo se podría mejorar?
- ¿Cuál es la línea de extensión futura de la investigación?

- 1. Introducción**
- 2. Determinantes de la I+D**
 - 1. Determinantes clásicos**
 - 2. El papel de la financiación pública**
- 3. I+D, innovación y productividad**
 - 1. El modelo de capital tecnológico**
 - 2. El modelo CDM**
 - 3. Modelos estructurales recientes**
- 4. Dentro de la caja negra de la I+D**
 - 1. La composición de la I+D**
 - 2. Complementariedades en las actividades de I+D**
- 5. Conclusiones**

1. Introducción

2. Determinantes de la I+D

1. Determinantes clásicos
2. El papel de la financiación pública

3. I+D, innovación y productividad

1. El modelo de capital tecnológico
2. El modelo CDM
3. Modelos estructurales recientes

4. Dentro de la caja negra de la I+D

1. La composición de la I+D
2. Complementariedades en las actividades de I+D

5. Conclusiones

- Durante las últimas décadas, los economistas han prestado una atención creciente al papel de la I+D+I en las economías modernas.
- Orígenes de la agenda de investigación clásica:
 - Solow (1957): sobre la importancia del crecimiento de la productividad total de los factores como medida del cambio tecnológico
 - Nelson (1959) and Arrow (1962): sobre la economía de la creación de conocimiento
- Desde un **punto de vista teórico**, las herramientas de la nueva teoría de juegos se aplican para analizar el comportamiento y las interacciones de las empresas que llevan a cabo actividades de I+D.
- Desde un **punto de vista empírico**, numerosos autores comienzan la tarea de medir y entender los determinantes y los resultados de la I+D.

- Hoy en día, la literatura empírica sobre I+D es abundante y sigue creciendo debido especialmente al incremento de fuentes de micro-datos disponibles.
- Tres tópicos principales cubiertos por la literatura:
 - Los determinantes de la inversión en I+D de las empresas
 - La relación entre I+D, innovación y productividad
 - El contenido de la caja negra de la I+D
- Otros tópicos también muy estudiados:
 - La relación entre innovación y empleo
 - Las causas y efectos del *outsourcing* u *offshoring* de I+D
 - La medición de las externalidades de la I+D (especialmente en el caso de la cooperación tecnológica)
 - El papel de la innovación en la creación de empresas

1. Introducción

2. Determinantes de la I+D

1. Determinantes clásicos
2. El papel de la financiación pública

3. I+D, innovación y productividad

1. El modelo de capital tecnológico
2. El modelo CDM
3. Modelos estructurales recientes

4. Dentro de la caja negra de la I+D

1. La composición de la I+D
2. Complementariedades en las actividades de I+D

5. Conclusiones

Desde un punto de vista teórico, los estudios se suelen apoyar en modelos a la Dorfman-Steiner (1954), donde empresas maximizadoras de beneficios eligen el nivel de inversión en I+D que iguala el ingreso marginal procedente de la I+D, *MRR*, con su coste marginal, *MCR*.

La relación específica entre las variables de interés depende de si se considera que la inversión en I+D, *R*, es una estrategia para incrementar la demanda (innovación de producto) y/o para reducir costes (innovación de proceso).

Siguiendo a David et al. (2000), para cada periodo de planificación y proyecto de I+D, este marco se puede representar a través de las ecuaciones siguientes:

$$MRR=f(R,X)$$

$$MCR=g(R,Z)$$

X: variables que pueden afectar la distribución de tasas de retorno de los proyectos, como las oportunidades tecnológicas y las condiciones de apropiabilidad

Z: variables que afectan al coste marginal de I+D, como los relacionados con acceso a financiación bancaria o capital de riesgo, las condiciones macroeconómicas o políticas públicas tecnológicas.

Donde se igualan MRR y MCR encontraríamos el nivel óptimo de inversión en I+D, R^* :

$$R^*=h(X,Z)$$

Esta sería la ecuación a estimar desde un punto de vista empírico. La forma de la relación esperada con las variables explicativas dependerá de las funciones utilizadas para obtener el ingreso marginal y el coste marginal.

Partiendo de este tipo de modelo, nuevas cuestiones a analizar:

- Importancia de tener en cuenta los determinantes de la I+D tanto en el **margen intensivo** (cuantía/intensidad de la inversión) como en el **extensivo** (decisión de realizar actividades de I+D).
- Relación de **complementariedad** o sustituibilidad entre R&D pública y privada R&D.
- **Persistencia** en las actividades de I+D.

Problemas para la estimación de la ecuación de R&D:

- 1. Características no observables de las empresas**, tales como la cultura o habilidad del equipo director. Esta cuestión se aborda mediante el uso de algún tipo de transformación (intragrupos, diferencias) que elimina los efectos individuales invariantes en el tiempo.
- 2. La inversión en I+D debe analizarse en un marco dinámico ya que tiene elevados costes de ajuste.** En consecuencia, las empresas tienden a suavizar su inversión en I+D en el tiempo (Lach y Schankerman, 1989).
 - El principal enfoque para incorporar la dinámica es el modelo neoclásico del acelerador con dinámica *ad hoc* (Mairesse et al., 1999), introduciendo la variable dependiente retardada.
 - Este modelo se expresa a veces en forma de corrección de error para tener en cuenta de forma explícita los efectos de corto plazo versus los de largo plazo (Bond et al, 2005).
 - Estudios recientes distinguen entre los costes de ajuste de las empresas que se inician en las actividades de I+D de los costes de ajuste de las empresas que ya invierten en I+D (Peters et al., 2013).

3. Endogeneidad de la inversión en I+D:

- Puede estar correlacionada con efectos inobservables invariantes en el tiempo
- Puede responder a expectativas sobre futuros shocks tecnológicos
- Se puede determinar simultáneamente con otras características de la empresa incluidas en X_{it} .
- Shocks aleatorios a la I+D actual pueden afectar a valores futuros de las variables explicativas

Estos problemas generalmente han sido tratados utilizando los métodos *GMM-IV* o *system-GMM*, que coinciden en utilizar los retardos de las variables explicativas como instrumentos (Bond et al., 2005).

- ### 4. El efecto de los diferentes determinantes puede no ser homogéneo entre empresas.
- Si uno cree que son heterogéneos, los coeficientes estimados reflejan efectos promedio dentro de la muestra. Algunos autores han explorado si el efecto de los determinantes de la I+D varían en diferentes tipos de empresas (p. ej., pequeñas vs grandes, de baja tecnología vs alta tecnología, o jóvenes vs maduras).

- Los principales determinantes clásicos de la inversión en I+D son el tamaño de la empresa, el poder de mercado y determinantes sectoriales, como el grado de apropiabilidad, el tirón de la demanda o la oportunidad tecnológica.
- El papel del tamaño y el poder de mercado se ha analizado desde Schumpeter (1939, 1942) mientras que el enfoque de determinantes sectoriales se origina en una serie de trabajos en la década de 1980 (Levin et al., 1985; Cohen et al., 1987; Cohen y Levin, 1989).

Tamaño de la empresa:

- Argumentos a favor de una relación positiva:
 - Las grandes empresas son generalmente capaces de generar mayores flujos de efectivo para financiar internamente sus inversiones y tienen mejor acceso a los mercados de capital.
 - Además, las inversiones en I+D presentan generalmente economías de escala y alcance.

- Hechos estilizados del análisis empírico (Cohen y Klepper, 1996; Cohen, 2010):
 - I. la probabilidad de realizar actividades de I+D aumenta con el tamaño
 - II. La inversión en I+D aumenta monótonamente y, por lo general, proporcionalmente con el tamaño (entre industrias e inversores en I+D)
 - III. Las pequeñas empresas muestran una proporción de innovaciones y patentes mayor de lo esperado según su inversión R&D.

- Estudios más recientes se centran en analizar directamente los motivos de la importancia del tamaño.
 - El efecto está relacionado positivamente con la disponibilidad de fondos (Bloch, 2005; Brown et al., 2009; Czarnitzki y Hottenrott, 2011; Borisova y Brown, 2013).
 - La relación es más fuerte en los Estados Unidos y el Reino Unido que en Europa continental (Hall, 2002; Bond et al., 2005), debido a una diferente estructura de los mercados de capital y las diferentes actitudes empresariales frente a la incertidumbre.

Poder de mercado:

- La relación entre competencia e inversión en I+D es menos clara.
- La teoría postula dos efectos contrapuestos:
 - Una disminución del poder de mercado puede reducir el incentivo para invertir porque la empresa sería menos capaz extraer la rentas resultantes la innovación (Grossman y Helpman, 1991; Aghion y Howitt, 1992).
 - La innovación podría desplazar parcialmente las rentas oligopólicas, reduciendo los incentivos a la I+D de las empresas con alto poder de mercado (Arrow, 1962), o podría utilizarse como una variable estratégica para afrontar la creciente competencia (Spencer y Brander, 1983).
- La evidencia empírica tampoco es concluyente, obteniéndose resultados a favor tanto de una relación positiva (Crèpon et al., 1998; Blundell et al., 1999), como negativa (Geroski, 1990; Harris et al., 2003) o en forma de U invertida (Aghion et al., 2005)
 - Cohen (2010) concluye que lo que emerge de la evidencia empírica anterior es que el poder de mercado no parece desempeñar un papel clave de forma independiente.

Determinantes sectoriales:

- Desde la década de 1980, gran atención a los determinantes de carácter sectorial (Levin et al., 1985; Cohen et al., 1987; Cohen y Levin, 1989; Geroski, 1990)
- **Apropiabilidad:** Aunque la mayoría de los estudios se ha centrado en el papel desempeñado por el sistema de patentes, los métodos más empleados para apropiarse de los resultados de la innovación son el tiempo, el secreto y los activos complementarios. La evidencia empírica no es concluyente:
 - Las empresas deben ser capaces de apropiarse de suficientes resultados para que su inversión valga la pena (Levin et al., 1987).
 - Sin embargo, una excesiva apropiabilidad puede desincentivar el esfuerzo innovador: Más apropiabilidad implica menos spillovers de I+D, que suelen ser complementarios de la I+D interna.
- **Oportunidad tecnológica y tirón de la demanda:** La evidencia empírica sugiere que ambos son importantes, aunque la fracción de varianza explicada por la oportunidad tecnológica es mayor.

La literatura empírica es abundante (**este tema se profundizará en la siguiente sesión**). Sin embargo, sólo unos pocos trabajos se apoyan en modelos estructurales que describen las decisiones de las empresas:

González, Jaumandreu y Pazó (2005)

Describen tanto la decisión de hacer I+D (margen extensivo) como el esfuerzo inversor (margen intensivo). En su modelo, cada competidor produce un producto diferenciado y es capaz de incrementar la demanda de ese producto aumentando su calidad mediante la I+D.

Cada empresa maximiza sus beneficios, donde sus ingresos netos son función de los gastos en I+D, R , siendo s la fracción de esos gastos que está subvencionada. El esfuerzo óptimo (no nulo) viene dado por la expresión:

$$e^* = \frac{R^*}{P^* Y^*} = \frac{R}{Y} \frac{\partial Y}{\partial R} \left/ \left(-\frac{P}{Y} \frac{\partial Y}{\partial P} E[(1-s)^\beta] \right) \right.$$

donde E representa la expectativa sobre los valores de s y β el nivel de eficiencia en el gasto de los fondos públicos.

Esta expresión refleja el esfuerzo óptimo tanto para las empresas que realizan actividades de I+D como para las que no las realizan. Pero sólo se observa si sobrepasa el esfuerzo umbral asociado al nivel de gasto que hace indiferente a la empresa entre invertir y no invertir en I+D. Por debajo de ese umbral, los costes de la I+D no son recuperados completamente mediante el incremento en las ventas.

Este marco conduce de forma natural a una modelización de tipo Tobit para una variable censurada. El modelo empírico vendría dado por las ecuaciones:

$$e^* = -\beta \ln(1 - s^e) + \mathbf{z}_1 \boldsymbol{\beta}_1 + u_1 \quad [2.5]$$

$$\bar{e} = \mathbf{z} \boldsymbol{\beta}_2 + u_2 \quad [2.6]$$

donde e^* se observa sólo si $e^* - \bar{e} > 0$. \mathbf{z} representa el vector de determinantes tradicionales. s^e es la expectativa de s , que es inobservable y debe ser estimada previamente. Para ello, esta expectativa se descompone en el producto de la esperanza condicional de recibir la subvención y el valor esperado de la subvención condicional en \mathbf{z} y en su concesión. Estos elementos se estiman respectivamente utilizando especificaciones Probit y OLS.

Takalo et al. (2013b)

Complementa el marco de González et al. (2005) con un modelo estructural de la decisión de la empresa de solicitar la subvención. El programa de ayudas se describe mediante un juego de 4 etapas de información incompleta entre los jugadores:

Etapa 0: Se determinan los tipos de jugadores (empresa/proyecto y agencia)

Etapa 1: La empresa decide solicitar o no la ayuda

Etapa 2: La agencia califica la propuesta y aprende su tipo

Etapa 3: La empresa elige la inversión en I+D con o sin la subvención.

También en este caso la inversión óptima en I+D (margen intensivo) se obtiene a través de una condición de primer orden de una empresa maximizadora de beneficios.

Dado que el objetivo de su enfoque es derivar ecuaciones que puedan estimarse empíricamente, (they) *“model the players’ payoffs by more specific functional forms that would be necessary from a purely theoretical point of view”*.

En particular, expresan la función objetivo de la empresa y la utilidad esperada que obtiene la agencia de cada proyecto solicitado de la forma siguiente:

$$\Pi(R(s), s, \mathbf{x}, \varepsilon) = \exp(\mathbf{x}\boldsymbol{\beta} + \varepsilon) \ln R(s) - (1-s)R(s)$$

$$U(R(s), s, \mathbf{x}, \mathbf{z}, \varepsilon, \eta) = V(R(s), \mathbf{z}, \eta) + \Pi(R(s), s, \mathbf{x}, \varepsilon) - gsR(s) - F$$

donde s es la tasa de subvención.

En la segunda ecuación, el primer término, $V(\cdot)$, refleja los spillovers del proyecto sin tener en cuenta los beneficios de las empresas, los costes directos de la subvención ($gsR(s)$) y los costes fijos de solicitar y procesar la solicitud (F). g representa el coste de oportunidad de los recursos de la agencia, mientras que \mathbf{x} y \mathbf{z} son vectores de características observables de las empresas.

En este marco, Takalo et al. (2013b) prueban que existe un único equilibrio Bayesiano perfecto del juego.

Modelo empírico: consiste en la estimación de cuatro tipos de ecuaciones que caracterizan:

1. La decisión de la empresa de solicitar la subvención
2. La decisión de la empresa sobre la inversión en I+D
3. La decisión de la agencia sobre la valoración (calificación) de los proyectos
4. La decisión de la agencia sobre la tasa de subvención.

Dos supuestos fuertes de este modelo es que no hay costes fijos de la I+D ni racionamiento de crédito.

En un trabajo paralelo Takalo et al. (2013a) amplían el modelo para incluir la existencia de un coste de la financiación externa y costes fijos de los proyectos de I+D.

Arqué-Castells y Mohnen (2015)

Consideran una dimensión adicional: La existencia de costes fijos irrecuperables al iniciar las actividades de I+D por primera vez. En consecuencia, diferencian entre **subvenciones extensivas** (pensadas para incrementar el número de empresas inversoras en I+D) y **subvenciones intensivas** (pensadas para incentivar el aumento del esfuerzo de las ya inversoras).

Consideran un modelo dinámico donde las empresas deciden empezar, continuar o dejar de hacer I+D dependiendo de las subvenciones. Con este modelo caracterizan las estrategias de las empresas en términos de dos umbrales de subvención, que determinan la entrada y la continuación de las actividades de I+D.

Modelo empírico: consiste en la estimación de un modelo Tobit tipo-2 dinámico con datos de panel, con dos ecuaciones:

1. Ecuación de participación en las actividades de I+D
2. Ecuación de inversión en I+D

1. Introducción

2. Determinantes de la I+D

1. Determinantes clásicos
2. El papel de la financiación pública

3. I+D, innovación y productividad

1. El modelo de capital tecnológico
2. El modelo CDM
3. Modelos estructurales recientes

4. Dentro de la caja negra de la I+D

1. La composición de la I+D
2. Complementariedades en las actividades de I+D

5. Conclusiones

- Muchos autores han analizado la relación entre las actividades de I+D y la productividad, encontrando, en general, un efecto positivo, aunque con distintas magnitudes dependiendo de la metodología empleada y el nivel de análisis.
- En este sentido, como señalan Mairesse y Sassenou (1991), **la cuestión no es tanto si existe o no esa relación, sino si los estudios econométricos pueden caracterizar dicha relación de manera satisfactoria y útil.**
- La medición del impacto es difícil porque:
 - la I+D afecta tanto a la oferta como a la demanda de productos.
 - La I+D de otras empresas en el propio sector (o diferentes sectores y países) puede generar externalidades positivas en otras empresas del mismo sector (o diferentes sectores o países).
- Podemos distinguir tres aproximaciones fundamentales:
 - el modelo de capital de conocimiento
 - el modelo CDM
 - los recientes modelos estructurales

- A nivel de empresa, la mayoría de los estudios tempranos siguen a Griliches (1979), utilizando una función de producción (usualmente Cobb-Douglas) aumentada con capital tecnológico (o gasto en I+D) para estimar los rendimientos de la I+D.

$$Y_{it} = A_{it} F(L_{it}, K_{it}, M_{it}, G_{it})$$

donde Y representa la cantidad de producto, L , K y M son, respectivamente, las cantidades de trabajo, capital físico y consumos intermedios, G es la cantidad de capital de conocimiento y A representa el nivel de eficiencia alcanzado por la empresa.

- Por tanto, este modelo se centra en el lado de la oferta. El impacto del capital de I+D se obtendría estimando esa función de producción de Cobb-Douglas en niveles o en diferencias:

$$y_{it} = a_{it} + \alpha_l l_{it} + \alpha_k k_{it} + \alpha_m m_{it} + \gamma g_{it} + u_{it}$$

$$\Delta y_{it} = \Delta a_{it} + \alpha_l \Delta l_{it} + \alpha_k \Delta k_{it} + \alpha_m \Delta m_{it} + \gamma \Delta g_{it} + \Delta u_{it}$$

- Griliches (1979) propone calcular el capital de conocimiento, G , como la acumulación de gastos en I+D en el tiempo, es decir, construir una variable que mida el **stock de capital en I+D** de la empresa. Específicamente utiliza el **método del inventario perpetuo**, que es el método que más se sigue utilizando actualmente.
- La estimación de esta función de producción conlleva una serie de **problemas econométricos**, que incluyen la simultaneidad (endogeneidad de las decisiones de inputs de la empresa) y el sesgo de selección (salida endógena).
- A ello se suman las **limitaciones de datos** que dan lugar a errores de medida :
 - La tasa de depreciación de la I+D se suele suponer constante en el tiempo.
 - Dificultad para encontrar condiciones iniciales para el stock de capital en I+D.
 - Existen cambios de calidad en factores y producto a lo largo del tiempo. Si los precios utilizados para deflactar no incluyen un ajuste de calidad, las cantidades en términos reales estarán infravaloradas.
 - El problema de medición del producto se acrecienta si la empresa tiene poder de mercado (quizá generado por sus innovaciones) y se utiliza un deflactor sectorial.

- Bajo los supuestos de rendimientos constantes a escala y competencia perfecta, algunos trabajos utilizan el **marco de la contabilidad de crecimiento**, que permite relacionar el crecimiento de la **productividad total de factores (PTF)** con la I+D de una forma no paramétrica.
- En este contexto las elasticidades de los factores se igualan a sus participaciones en el ingreso de la empresa, lo que permite aproximar la PTF a partir del denominado **residuo de Solow** como:

$$\tilde{\theta}_{it} = y_{it} - (s_l l_{it} + s_k k_{it} + s_m m_{it})$$

- Principales problemas de esta metodología? El incumplimiento de los supuestos (rendimientos constantes a escala, ajuste instantáneo de los inputs, mercados competitivos) y la interpretación de la PTF. Nótese que recoge todo lo que no se asocia al trabajo, la intensidad de capital y los consumos intermedios, como son los cambios en la eficiencia de las empresas o la utilización de la capacidad y los errores de medida en las variables.

- Alternativa? Incluir en la ecuación variables que reflejen el incumplimiento de los supuestos y otras variables de control (como el ciclo económico), \mathbf{x} . En este contexto, el impacto del capital de conocimiento sobre la PTF se calcularía estimando la ecuación:

$$\tilde{\theta}_{it} = \lambda g_{it} + \mathbf{x}_{it} \boldsymbol{\beta} + v_{it}$$

- Las tasas de rendimiento de la I+D en las economías desarrolladas durante el pasado medio siglo han sido fuertemente positivas y se han situado mayoritariamente en el rango del 20 al 30%.
- Wieser (2005) revisa la literatura empírica que utiliza este enfoque, también teniendo en cuenta el impacto de los *spillovers* de I+D:
- Existe en promedio un impacto significativo de la I+D sobre los resultados empresariales.
 - El meta-análisis de los estudios revisados muestra que las tasas estimadas de rendimiento no difieren significativamente entre países, mientras que sí lo hacen las elasticidades.
 - Las elasticidades son significativamente mayores en la década de 1980 y más altas en la década de 1990 comparadas con la década de 1970.

La interacción con el lado de la demanda

- Cuando se utilizan deflatores sectoriales con datos de empresa, las medidas de producción (ventas nominales deflactadas) también reflejan el impacto de la innovación de productos.
- Klette y Griliches (1996) tratan el problema de los precios de producto no observables mediante una ecuación de demanda. Consideran que, con una adecuada especificación del sistema de demanda, hay identificación:
 - La variable precio omitida puede expresarse en términos del crecimiento de la producción de la empresa respecto a la producción de la industria y, eventualmente en términos de las variables observables y parámetros ya presentes en la función de producción (o de costes).
- Solo unos pocos trabajos incorporan el lado de la demanda. En ellos, el impacto de la I+D sobre los ingresos suele ser mayor que el impacto sobre la producción (véase Hall et al. (2010) para una revisión de esta literatura).

- En el enfoque anterior, la variable que mide la actividad tecnológica y que se incorpora en las ecuaciones de oferta y demanda es el capital en I+D+i (o los gastos en I+D), que es una medida de input tecnológico.
- **Crèpon, Duguet y Mairesse (1998)** desarrollan un modelo multiecuacional que explica la productividad a partir de los resultados tecnológicos (innovaciones, patentes, etc.) y éstos a partir del esfuerzo tecnológico (input tecnológico). Su planteamiento se basa en la idea de **son los resultados innovadores (y no los factores empleados en su obtención) los que incrementan la productividad** de las empresas. Este modelo recibe desde entonces la denominación de modelo CDM.
- Desde la aparición de este trabajo seminal, muchos investigadores han aplicado la misma metodología a diferentes países europeos, usando datos de corte transversal esencialmente de las Encuestas Comunitarias de Innovación (CIS).
- La mayoría de ellos estima las ecuaciones de este modelo de forma recursiva, aunque algunos trabajos recientes estiman las ecuaciones simultáneamente.

- En este modelo, la secuencia de decisiones sería la siguiente:

1ª etapa: Las empresas deciden si llevar a cabo actividades de I+D. Esta decisión se representa mediante la ecuación:

$$dR_{it} = \begin{cases} 1 & \text{if } dR_{it}^* = \mathbf{x}_{1it}\boldsymbol{\beta}_1 + \varepsilon_{1it} > 0 \\ 0 & \text{if } dR_{it}^* = \mathbf{x}_{1it}\boldsymbol{\beta}_1 + \varepsilon_{1it} \leq 0 \end{cases}$$

donde dR es una variable binaria que toma valor 1 si la empresa invierte en I+D y 0 en caso contrario. Si la variable latente dR^* es mayor que un umbral constante (que puede ser 0) observaríamos que la empresa hace I+D. \mathbf{x}_1 es el vector de variables explicativas observables.

Condiciona a la realización (y declaración) de actividades de I+D, la empresa decide la intensidad del gasto. Esta decisión se representa como sigue:

$$R_{it} = \begin{cases} R_{it}^* = \mathbf{x}_{2it}\boldsymbol{\beta}_2 + \varepsilon_{2it} & \text{if } dR_{it} = 1 \\ 0 & \text{if } dR_{it} = 0 \end{cases}$$

donde \mathbf{x}_2 representa los determinantes del esfuerzo innovador, que pueden diferir de los que explican la decisión de realizar actividades.

La mayoría de los trabajos presuponen que los términos de error de las ecuaciones previas siguen una distribución normal bivalente y suelen estimarlas conjuntamente por máxima verosimilitud como un modelo Tobit generalizado (Heckman 1976, 1979).

2ª etapa: Se obtienen resultados innovadores a partir del esfuerzo tecnológico.

Para representar esta fase se utiliza una función de producción de nuevo conocimiento, g , definida como:

$$g_{it} = \lambda R_{it}^* + \mathbf{x}_{3it} \boldsymbol{\beta}_3 + \varepsilon_{3it}$$

donde \mathbf{x}_3 incluye otros determinantes de la producción de conocimiento además del propio esfuerzo inversor.

En el trabajo original, el nuevo conocimiento se mide por dos variables: el número de patentes y el porcentaje de ventas innovadoras (por intervalos, ventas de productos nuevos para el mercado en los últimos 5 años). El método econométrico de estimación se adapta al tipo de variable (*count data, ordered probit*).

Estudios posteriores complementan estos indicadores con variables dicotómicas que reflejan el logro de innovaciones de producto, proceso y, recientemente, de innovación organizacional.

3ª etapa: Los resultados innovadores incrementan la productividad de la empresa.

La mayoría de estudios consideran una función aumentada de producción tipo Cobb-Douglas con capital físico, empleo, resultados innovadores y otras variables de control. Además, generalmente se suponen rendimientos constantes a escala (en los inputs tradicionales). Siguiendo el modelo CDM original, la forma funcional mas utilizada es la siguiente:

$$y_{it} - l_{it} = \delta g_{it} + \mathbf{x}_{4it} \boldsymbol{\beta}_4 + \varepsilon_{4it}$$

donde $y - l$ es la productividad del trabajo (normalmente valor añadido o producción por empleado, en logs.) y \mathbf{x}_4 incluye otros determinantes como el capital físico por empleado (en logs.) o la composición de cualificaciones.

En el modelo CDM original, todas las ecuaciones se estiman conjuntamente por mínimos cuadrados asintóticos (ALS) o un estimador de mínima distancia.

Sin embargo, trabajos posteriores utilizan un enfoque secuencial: el valor predicho de la variable dependiente de cada sistema entra como factor determinante en la ecuación siguiente.

- Revisiones de esta literatura han sido proporcionadas por Hall (2011) y Mohnen y Hall (2013):
 - La productividad se relaciona positivamente tanto con las ventas innovadoras como con el indicador binario de innovaciones de producto.
 - La asociación es mayor en sectores de alta tecnología.
 - El efecto de las innovaciones de proceso presenta una gran variabilidad (incluyendo resultados negativos). Posibles explicaciones (Hall, 2011):
 - Las empresas operan en la parte inelástica de su curva de oferta, de forma que la productividad de sus ingresos no se ve afectada por mejoras en la eficiencia del proceso de producción
 - hay tanto error en la medición de las variables de innovación que sólo una de las dos (innovaciones de producto o de proceso) es positiva cuando se añade a la ecuación de productividad.

La persistencia en los inputs y outputs tecnológicos:

- Aunque existe amplia evidencia empírica que contrasta el modelo CDM, en su mayoría se basa en datos de corte transversal. Este hecho no permite tener en cuenta en las estimaciones la persistencia de las actividades de I+D, los vínculos dinámicos entre innovación y resultados económicos o la heterogeneidad de empresa inobservable.
- Recientemente, algunos estudios han comenzado a considerar el ritmo de la innovación y determinados aspectos dinámicos.
- La persistencia en las actividades de innovación podría explicar la persistencia en los resultados económicos.
- Dos explicaciones fundamentales del comportamiento persistente:
 - a) Verdadera dependencia del estado (*true state dependence*)
 - b) Dependencia espuria

- a) Dependencia espuria: Algunas características de la empresa pueden afectar positivamente a su decisión de hacer I+D o a la generación de innovaciones y, si están correlacionadas en el tiempo, podrían crear una relación espuria entre el estatus actual y el futuro. Estas variables pueden ser observables (tamaño) o inobservables (habilidad empresarial, oportunidades tecnológicas, actitudes frente al riesgo).
- b) Verdadera dependencia del estado: Implica efecto causal real, es decir, la probabilidad de invertir en $t-1$ incrementa la probabilidad de invertir en t . Explicaciones (Peters, 2009; Mañez-Castillejo et al., 2009; Raymond et al., 2010):
- **Costes irrecuperables** (de entrada). Barreras a la salida.
 - **Hipótesis “success breeds success”**. Oportunidades tecnológicas. Beneficios de innovaciones pasadas incrementan los fondos internos para proyectos futuros.
 - **Rendimientos a escala dinámicos**. El conocimiento se acumula en el tiempo (Nelson y Winter, 1982). Efectos de aprendizaje (*learning by doing*) Trayectoria tecnológica (a innovaciones radicales le sigue una sucesión de innovaciones incrementales).

- Desde un punto de vista empírico, número menor de trabajos sobre la persistencia en inputs que en outputs tecnológicos.
 - Sobre inputs (gasto en I+D o gasto en innovación): Mañez Castillejo et al. (2009) y Peters (2009) obtienen evidencia de verdadera persistencia del estado independientemente de la metodología.
 - Sobre innovaciones de producto, de proceso y organizativas: también evidencia de alta persistencia (Duguet y Monjon, 2004; Rogers, 2004; Flaig and Stadler, 1994).
 - Sobre patentes: la mayoría de trabajos encuentran un bajo nivel de persistencia (Geroski et al. 1997; Crèpon y Duguet, 1997; Malerba y Orsenigo, 1999; Cefis y Orsenigo, 2001; Cefis, 2003). Razones:
 - No todas las invenciones se patentan.
 - Patentar implica tanto innovar como ser el primero en innovar (Duguet y Monjon, 2004).

- En el contexto del modelo CDM, se obtiene evidencia de verdadera dependencia del estado tanto en la decisión de hacer I+D como en la generación de innovaciones (Huelgo y Moreno, 2011). La omisión de este tipo de persistencia conlleva una sobreestimación del impacto de las innovaciones en el crecimiento de la productividad.

Los modelos estructurales recientes, suponen que la productividad es inobservable y la representan como un proceso de Markov que depende de los gastos en I+D. Es decir, la productividad viene afectada por la decisión endógena de invertir en I+D.

Doraszelski y Jaumandreu (2013):

A partir de la función de producción Cobb-Douglas estándar en logaritmos,

$$y_{it} = a_0 + \alpha_t t + \alpha_l l_{it} + \alpha_k k_{it} + \alpha_m m_{it} + w_{it} + u_{it}$$

suponen que la productividad en el periodo t , w_{it} , se puede descomponer en la productividad esperada y un shock aleatorio:

$$w_{it} = E \left[w_{it} \mid w_{it-1}, R_{it-1} \right] + \xi_{it} = G(w_{it-1}, R_{it-1}) + \xi_{it}$$

donde R_{it-1} son los gastos en I+D en $t-1$. Es decir, la empresa anticipa el efecto en la productividad de la I+D en el período t al tomar la decisión sobre la inversión en conocimiento en el período $t-1$.

A partir de un panel de empresas manufactureras españolas, encuentran que los gastos en I+D determinan las diferencias en la productividad entre empresas y la evolución de la productividad de la empresa en el tiempo. También obtienen evidencia de las no linealidades y la incertidumbre presentes en el proceso de I+D.

Aw et al. (2011):

Utilizan un modelo de ventas procedentes del mercado interior y de exportación, y en lugar de utilizar una función de producción estándar, consideran una función de coste marginal de corto plazo y empresas que operan en mercados de competencia.

Los ingresos de la empresa en cada mercado dependen de las condiciones del mercado global, el capital, el vector de precios de los factores variables y la productividad de la empresa que expresan como:

$$w_{it} = G(w_{it-1}, dR_{it-1}, dX_{it-1}) + \xi_{it}$$

donde dR_{it-1} y dX_{it-1} representan, respectivamente, la participación de la empresa en el mercado de I+D y de exportación en t-1. Ello implica que también consideran la posibilidad de aprendizaje mediante la exportación (*learning-by-exporting*)

Utilizando datos de empresas de la industria electrónica taiwanesa, obtienen que ambas actividades afectan positivamente la futura productividad. Sin embargo, el efecto marginal de la I+D varía mucho con la productividad; es mucho mayor en empresas de alta productividad, y es mayor para las empresas no-exportadoras (aunque en magnitudes bastante pequeñas en este segundo caso).

Peters et al. (2013) y Roberts y Vuong (2013)

De forma complementaria, tratan el problema de la simultaneidad mediante la modelización y estimación de la decisión dinámica de la empresa de invertir en I+D en lugar de estimar solo una función de producción. Desarrollan un **modelo dinámico estructural de la demanda de I+D** de forma que el beneficio esperado de la inversión en I+D se infiere a partir de la decisión racional de esa inversión.

Una ventaja importante de estos modelos es que permiten el cambio de parámetros en el entorno de la empresa y la cuantificación del efecto de estos cambios sobre la decisión de la empresa de invertir en I+D, la productividad y el impacto de largo plazo en la rentabilidad.

Resultados (Peters et al., 2013):

- En proporción sobre el valor de la empresa, el beneficio neto para la empresa mediana con experiencia previa en I+D varía del 2.4 al 3.2% en 5 industrias de alta tecnología, y del -4.6 al 0.6% en empresas sin experiencia previa. Para la empresa inexperta mediana no resultaría rentable invertir en I+D. En industrias de baja tecnología, los beneficios netos son bastante menores.
- En industrias de alta tecnología, una reducción del 20% del coste fijo de la I+D da lugar a los 5 años a un aumento promedio de 7 p.p. en la probabilidad de invertir en I+D y del 4% en la productividad.

1. Introducción

2. Determinantes de la I+D

1. Determinantes clásicos
2. El papel de la financiación pública

3. I+D, innovación y productividad

1. El modelo de capital tecnológico
2. El modelo CDM
3. Modelos estructurales recientes

4. Dentro de la caja negra de la I+D

1. La composición de la I+D
2. Complementariedades en las actividades de I+D

5. Conclusiones

4. Dentro de la caja negra de la I+D

- La literatura revisada hasta ahora generalmente está limitada por la falta de información y de análisis de las especificidades de la I+D.
- Trabajos recientes tratan explícitamente de abrir la caja negra de la I+D teniendo en cuenta que investigación y desarrollo son dos actividades diferentes y, por tanto, pueden diferir en términos de sus determinantes y efectos.
- Por otra parte, las empresas a menudo adoptan a varias estrategias de I+D+i al mismo tiempo, y esta coexistencia de estrategias sugiere la existencia de complementariedades.

4.1. La composición de la I+D

- Las actividades de investigación (*I*) y desarrollo (*D*) difieren en propósitos, bases de conocimiento, personas involucradas y estilos de gestión (Barge-Gil y López, 2015):
 - El objetivo principal de *I* es adquirir nuevos conocimientos, mientras que el principal propósito de *D* es la introducción de productos o procesos nuevos o mejorados (OCDE, 2005).
 - *I* es de naturaleza más teórica (aunque con frecuencia orientada a un objetivo práctico) y se basa en el conocimiento analítico. *D* es esencialmente aplicado y basado en el conocimiento sintético (Asheim y Coenen, 2005).
 - *I* necesita capital humano especializado que trabaje sin mucha jerarquía y de forma relativamente independiente del resto de la organización, mientras que *D* muestra una jerarquía clara y necesita a generalistas capaces de coordinarse con otras funciones de la organización (Leifer y Triscari, 1987).
 - En consecuencia, *I* y *D* cada vez más se llevan a cabo en diferentes departamentos, e incluso en lugares distantes (Chiesa, 2001).

➤ Temas de estudio analizados en este ámbito:

- **Determinantes de la composición**

Mansfield (1981): las grandes empresas están más orientadas a *I* básica. Además, hay cierta correlación entre el número de innovaciones y la proporción de *I* básica sobre los gastos totales en I+D

Link (1982, 1985): la orientación a *D* es mayor en empresas que operan en mercados más concentrados y reciben financiación pública, mientras que las empresas con un mayor nivel de ganancias están más orientadas a *I* aplicada. Las cuestiones administrativas son importantes dado que las empresas con una estrategia más ofensiva y con laboratorios de I+D centrales son también las que cada vez están más orientadas hacia *I* básica y de largo plazo.

- **Efectos de la composición**

Mansfield (1980): existe una relación positiva entre la cantidad de *I* básica y la productividad, dados otros gastos en I+D (que no tienen un efecto en la productividad).

Griliches (1986): *I* básica parece ser más importante como determinante de la productividad que otros tipos de I+D.

Lichtenberg y Siegel (1991): sólo la inversión en *I* básica muestra un efecto positivo sobre la productividad.

- **Relación entre financiación pública y composición**

Aerts and Thorwarth (2009): hay adicionalidad de la financiación pública en D pero no en I .

Clausen (2009): hay adicionalidad en las subvenciones para I , pero no en las subvenciones para D .

Czarnitzki et al. (2011): efecto mayor de las restricciones financieras sobre I .

- **Determinantes de cada uno de los componentes**

Barge-Gil y López (2014): el tirón de la demanda y la apropiabilidad tienen un mayor efecto sobre D , mientras que la oportunidad tecnológica sobre I . Para las empresas grandes, el efecto del tamaño es generalmente mayor sobre D , y no se capta efecto del poder de mercado.

- **Efectos de cada uno de los componentes**

Lim (2008): I aplicada tiene un efecto mucho mayor sobre el número de patentes que I básica (industria farmacéutica y de semiconductores)

Czarnitzki y Thorwall (2012): Existe un premio para la inversión en I básica en las industrias de alta tecnología, que se incrementa con el tamaño de la empresa.

Barge-Gil y López (2015): I contribuye más para obtener innovaciones de proceso, mientras que D contribuye más para lograr innovaciones de producto (y ventas asociadas).

- En su definición más general, se entiende que existe **complementariedad** entre prácticas si los resultados/beneficios de adoptar una práctica son mayores cuando están presentes las otras prácticas.
- Temas de estudio en este ámbito se refieren a la complementariedad entre:
 - **I+D interna y fuentes externas de tecnología** (subcontratación de I+D, licencias de I+D, contratación de personal...): Hou y Mohnen (2003), Hu et al. (2005), Cassiman y Veugelers (2006), Schmiedeberg (2008), Lokshin et al. (2008), Hagedoorn y Wang (2012), Ceccagnoli et al. (2014).
 - **Cooperación en I+D con socios diferentes**: Belderbos et al. (2006).
 - **Cooperación en I+D y otras prácticas** (habilidades técnicas, I+D externa, stock de I+D): Leiponen (2005), Schmiedeberg (2008), Harhoff et al. (2014), Arvanities et al. (2015)
 - **I+D e inversión en TICs**: Hall et al. (2013)
 - **I+D nacional e internacional**: Belderbos et al. (2015)
 - **I+D y capital humano**: Ballot et al. (2001)
 - **Componentes de la I+D**: Barge-Gil y López (2013)

- Enfoques principales para contrastar la existencia de complementariedades:

Enfoque-adopción:

- Se basa en el principio de las preferencias reveladas: bajo el supuesto de comportamiento optimizador de la empresa, la adopción conjunta de dos prácticas es potencialmente informativa acerca de los rendimientos conjuntos generados por ellas.
- Este enfoque ha sido popular entre los investigadores debido a su simplicidad. No requiere datos sobre la función objetivo, sólo la disponibilidad de datos sobre las prácticas. En este sentido, se considera un enfoque "indirecto".
- Dos tipos de metodologías usadas para implementar este enfoque:
 1. Medir el coeficiente de correlación, k_C , entre prácticas: se basa en la idea de que esta correlación puede interpretarse como la primera evidencia de complementariedad.
 2. Medir las correlaciones entre los términos de error de ecuaciones que representan la demanda de esas prácticas:

$$y_1^* = \mathbf{z}\alpha_1 + \varepsilon_1, \quad y_1 = 1 \text{ si } y_1^* > 0 \text{ y } 0 \text{ en caso contrario}$$

$$y_2^* = \mathbf{z}\alpha_2 + \varepsilon_2, \quad y_2 = 1 \text{ if } y_2^* > 0 \text{ y } 0 \text{ en caso contrario}$$

Enfoque-productividad:

- Se apoya en la estimación de una ecuación de resultados, en la que se incluyen interacciones (combinaciones) de las prácticas.
- **Si las prácticas tienen carácter discreto**, el análisis se basa en el concepto de supermodularidad de Topkis (1979). Para una función bidimensional $f(y_1, y_2)$, donde $y_1 = \{0, 1\}$ e $y_2 = \{0, 1\}$, las prácticas y_1 e y_2 son (estrictamente) complementarias si:

$$f(1,1) - f(0,1) > f(1,0) - f(0,0)$$

Para implementar este procedimiento empíricamente, típicamente se estima la ecuación:

$$f(y_1, y_2, \mathbf{z}) = \theta_{00}(1 - y_1)(1 - y_2) + \theta_{10}y_1(1 - y_2) + \theta_{01}(1 - y_1)y_2 + \theta_{11}y_1y_2 + \mathbf{z}\boldsymbol{\theta}_z + \varepsilon$$

donde \mathbf{z} es un vector de variables exógenas y $\varepsilon \sim N(0, \sigma_\varepsilon^2)$.

El test estadístico de complementariedad entre y_1 e y_2 se corresponde con :

$$\kappa_p \equiv \theta_{11} - \theta_{01} - \theta_{10} + \theta_{00}$$

Un valor de κ_p significativamente mayor que cero indica que podemos rechazar la hipótesis nula de ausencia de complementariedad.

- **Si las prácticas tienen carácter continuo**, la complementariedad entre las dos variables continuas significa que el efecto incremental de una variable en la función de resultados se incrementa de forma condicional al aumento de la otra variable

En este caso, el contraste de complementariedad se basa en la estimación de una función de resultados en un marco de regresión multivariante en función de las versiones continuas de las prácticas de interés junto con un término de interacción entre estas prácticas y otros factores exógenos:

$$f(y_1, y_2, \mathbf{z}) = \alpha_1 y_1 + \alpha_2 y_2 + \alpha_{12} y_1 y_2 + \mathbf{z} \alpha_{\mathbf{z}} + \varepsilon$$

donde \mathbf{z} es un vector de variables exógenas y $\varepsilon \sim N(0, \sigma_\varepsilon^2)$. Existe evidencia de complementariedad si el coeficiente α_{12} es significativamente mayor que 0.

- Tanto si las prácticas tienen carácter continuo como discreto, las metodologías se pueden ampliar al caso de tres o más prácticas (Caroli y Van Reenen, 2001; Mohnen y Röller, 2005).
- Es importante señalar que la **presencia de heterogeneidad inobservable** puede sesgar los contraste de complementariedades tanto en el enfoque-adopción como en el enfoque-productividad. Trabajos recientes proporcionan alternativas a este problema (Miravete y Pernías, 2006; Kretschmer et al., 2012; Brynjolfsson y Milgrom, 2013).

1. Introducción

2. Determinantes de la I+D

1. Determinantes clásicos
2. El papel de la financiación pública

3. I+D, innovación y productividad

1. El modelo de capital tecnológico
2. El modelo CDM
3. Modelos estructurales recientes

4. Dentro de la caja negra de la I+D

1. La composición de la I+D
2. Complementariedades en las actividades de I+D

5. Conclusiones

Sobre las preguntas de investigación:

Temas clásicos como el papel del tamaño o el poder de mercado como determinantes de la I+D o el efecto de la I+D en la productividad siguen recibiendo hoy día una atención considerable.

Sin embargo, existe una tendencia a centrarse en otras cuestiones:

- el papel desempeñado por la financiación pública como factor determinante,
- los canales (por ejemplo, nuevos productos y procesos) a través de los que la I+D influye en la productividad,
- la importancia de la composición de la I+D,
- la complementariedad de la I+D interna con otras actividades, como la I+D externa, la cooperación, las inversiones en TICs o el capital humano.

Otros temas que caen fuera de la revisión realizada y que también están recibiendo una atención creciente son: la relación entre I+D y empleo, las causas y efectos del abastecimiento internacional de I+D o, en asociación con la reciente crisis económica, la respuesta de la inversión I+D a ciclos.

Sobre la vinculación con la teoría:

Hay una tendencia cada vez mayor a usar modelos teóricos estructurales como base del análisis empírico. Estos modelos han hecho posible:

- I. distinguir entre los determinantes del margen extensivo de la I+D (la decisión de llevar a cabo actividades de I+D+i) y los del margen intensivo de la I+D (la magnitud o intensidad de los gastos en I+D+i);
- II. revelar los problemas de incertidumbre e información incompleta que suelen estar presentes en los mercados de I+D;
- III. vincular la adicionalidad del apoyo público a la I+D privada (cuando existe) con el bienestar social;
- IV. tener en cuenta la persistencia y la naturaleza dinámica de las decisiones de I+D y sus efectos;
- V. cuantificar el efecto de cambios en el entorno de la empresa sobre su decisión de invertir en I+D y la rentabilidad de largo plazo.

¿Hacia donde podría dirigirse la investigación empírica sobre la I+D empresarial?

Hay al menos dos áreas relacionadas en que sería interesante extender la labor investigadora: la metodología y la disponibilidad de datos.

- Aunque ha habido avances significativos en el uso de modelos estructurales, ésta sigue siendo una línea emergente de trabajo. Una segunda cuestión metodológica se refiere a la explotación y el diseño de experimentos controlados. El enfoque experimental puede complementar lo que aprendamos de modelos más sofisticados.
- En el lado de los datos, una oportunidad interesante es el uso de datos de panel que permitirían un análisis más adecuado de la relación entre I+D y la dinámica de entrada y salida de las empresas y también de la persistencia en las diferencias de comportamiento de las empresas. La disponibilidad de más datos cuantitativos también mejoraría los análisis cualitativos que con frecuencia están condicionados por el carácter categórico de las medidas de resultados de innovación en las bases de datos públicas.