



Universitat Autònoma
de Barcelona

Departament d'Economia Aplicada

Trabajo 8 Créditos

La importancia económica de los procesos de innovación a nivel de empresa. Un análisis empírico del sector manufacturero español

Resumen

Este trabajo analiza los lazos existentes entre productividad, innovación e investigación a nivel de empresa utilizando una extensión del modelo estructural de función de producción de *Crépon, Duguet y Mairesse (1998)*. El modelo explica la productividad por medio de la innovación y, la innovación a través de la inversión en investigación.

El estudio se realiza para las empresas del sector manufacturero español utilizando datos provenientes de la Encuesta sobre innovación tecnológica en las empresas 2004. Se corrigen los sesgos de selectividad y simultaneidad que se presentan dadas las características estadísticas de los datos y de las relaciones económicas entre las variables: sólo una pequeña proporción de las empresas realizan actividades de investigación; y la investigación, innovación y productividad se determinan de forma endógena.

Al aplicar la forma tradicional de estimación del modelo se encuentran resultados muy similares a trabajos anteriores realizados para España. Sin embargo, se demuestra que dicha formulación no produce estimaciones consistentes, por lo que aquí se utilizan extensiones alternativas, y se exploran nuevas especificaciones para las decisiones de realizar actividades internas de investigación y desarrollo.

Autor: LEONEL MUINELO GALLO

Directora: ISABEL BUSOM PIQUER

Febrero de 2008

ÍNDICE TEMÁTICO

| | |
|---|-----------|
| I. Introducción | 3 |
| II. Marco teórico | 5 |
| II.1. Literatura e hipótesis..... | 6 |
| II.2. El modelo estructural de Función de Producción. Un sistema de ecuaciones secuenciales | 9 |
| III. El Análisis econométrico del modelo estructural..... | 13 |
| III.1. Problemas específicos | 13 |
| III.2. El modelo de Crépon Duguet y Mairesse..... | 18 |
| III.3. Otras aplicaciones del modelo estructural | 21 |
| IV. Los datos | 32 |
| IV.1. La Encuesta sobre innovación tecnológica en las empresas del INE..... | 32 |
| IV.2. Características de la muestra en el período 2002-2004..... | 35 |
| V. Réplica del modelo GHMP | 37 |
| V.1. Modelo Econométrico | 38 |
| V.2. Definición de variables | 40 |
| V.3. Resultados de la Estimación | 41 |
| VI. Contrastes de la especificación GHMP y extensiones alternativas..... | 50 |
| VI.1. Problemas de endogeneidad en la ecuación de realización de actividades de I+D | 50 |
| VI.2. Inconsistencia de los métodos de estimación en etapas en sistemas ecuacionales no lineales | 54 |
| VI.3. La relación entre I+D interna e Innovaciones. Modelos Probit Bivariantes | 57 |
| VI.4. El impacto diferenciado de las innovaciones sobre la productividad | 64 |
| VII. Nuevas extensiones del modelo estructural: la decisión de realizar actividades de I+D | 65 |
| VII.1. Un modelo multinomial de decisión de I+D..... | 66 |
| VII.2. Un modelo ordenado de decisión de I+D | 66 |
| VIII. Comentarios finales y futuras líneas de investigación | 69 |
| Referencias | 73 |
| Anexo I. Definición de variables..... | 76 |
| Anexo II. Descriptivos de variables | 80 |
| Anexo III. Los efectos marginales en los modelos probit bivariantes recursivos.. | 82 |

I. Introducción

El estudio económico del tópico de la innovación ha ganado en interés desde el trabajo seminal de *Schumpeter (1911)*. Desde entonces, una gran cantidad de estudios económicos, comenzando con los trabajos pioneros basados en una aproximación de contabilidad del crecimiento de *Solow (1957)* y *Denison (1985)*, han demostrado empíricamente el rol central de la innovación tecnológica en el crecimiento económico.¹

A pesar que *Solow (1956)* destacara el papel de los rendimientos decrecientes y el descenso de la rentabilidad del capital como frenos en última instancia del proceso de crecimiento económico. El desarrollo de modelos macroeconómicos de crecimiento endógeno de primera y segunda generación ha demostrado cómo el avance tecnológico vía el flujo de nuevas ideas tiene un papel destacado para explicar el crecimiento económico en el largo plazo.

La primera generación de este tipo de modelos hace referencia al trabajo de *Romer (1994)*, en el cual se plantean modelos en los que, postergando los rendimientos decrecientes del capital por medio de suponer la existencia de externalidades en los procesos de Investigación y Desarrollo (I+D) o de introducir el capital humano, se logran generar tasas positivas de crecimiento en el largo plazo. Nuevamente dentro del enfoque de *Romer*, en los modelos de segunda generación el motor del crecimiento es alguno de una serie de procesos "generadores de progreso tecnológico en sentido amplio". Estas actividades incluyen, entre otras, la acumulación de capital humano por parte de las unidades familiares y la realización de actividades de I+D por parte de las empresas.

De forma complementaria, muchos investigadores han preferido utilizar métodos basados en estudios a nivel de empresa para evaluar la contribución del avance tecnológico al crecimiento económico. Es así que, en la literatura de la organización industrial, se atribuye a la innovación un papel determinante en el aumento de la productividad de las empresas.²

¹ En su estudio de 1957, *Solow* encontró que la cuota de la variación del producto atribuible al capital y al trabajo era muy pequeña. Él propuso una interpretación del residuo como una medida del progreso tecnológico. A pesar que planteó ciertas precauciones en cuanto a la interpretación de este residuo, la metodología que surgió a raíz de esta contribución se encuentra muy difundida en el análisis macroeconómico.

² Un ejemplo de análisis muy detallado sobre este punto es realizado por *Olley y Pakes (1996)*.

Sin embargo, a nivel microeconómico aún existen lagunas importantes en cuanto a los determinantes de la innovación y su impacto sobre el desempeño económico de las empresas. A pesar de la demostrada importancia de la innovación para el crecimiento de las empresas, no todas innovan. La innovación no se distribuye de forma aleatoria, sino que suele ser el producto de estrategias sistemáticas e inversiones que se realizan en busca de nuevos productos y/o procesos. En este marco, es importante plantear las siguientes preguntas:

- 1) ¿Por qué algunas empresas son más innovadoras que otras?; en el mundo real, ¿qué nos puede ayudar a explicar las diferencias en la creación o acceso a las nuevas ideas?
- 2) A su vez, ¿cómo repercuten estas innovaciones en su desempeño económico?

En un intento por encontrar respuestas a estas cuestiones, la investigación a nivel microeconómico se ha visto reforzada tanto a nivel teórico, con el desarrollo de modelos estructurales de función de producción, como a nivel empírico, debido a la disponibilidad de nuevas bases de datos provenientes de encuestas en las que se pregunta directamente a las empresas acerca de características de sus procesos innovadores.

Considerando este marco de análisis, la presente investigación se basa en una aproximación de modelo estructural de función de producción para analizar los determinantes del proceso global de innovación de las empresas manufactureras españolas, y evaluar el impacto en su desempeño económico.

En una primera instancia, utilizando datos de la “Encuesta sobre innovación tecnológica en las empresas 2004”, se reproduce la metodología teórica y empírica del estudio de *Griffith et al (2006)*; uno de los pocos estudios que han estimado de forma completa el modelo estructural a datos provenientes de esta encuesta. Sin embargo, se demuestra que dicha formulación no produce estimaciones consistentes. Debido a ello, en una segunda instancia, se utilizan extensiones alternativas y se exploran nuevas especificaciones para las decisiones de realizar actividades internas de I+D.

El trabajo está organizado de la siguiente forma. En la primera parte se expone el marco teórico. Se detallan diferentes hipótesis relacionadas con los procesos de innovación a nivel de empresa y la metodología teórica básica del modelo estructural de función de producción. Seguidamente, se discute la implementación empírica del modelo, analizando los problemas econométricos específicos que conlleva y el planteamiento

realizado por *Crépon, Duguet y Mairesse (1998)*. En el siguiente apartado, se incluye una exposición de los estudios que han utilizando dicha metodología. Aquí, sin pretender realizar un análisis exhaustivo, se exponen algunos trabajos empíricos que caracterizan el comportamiento innovador de las empresas manufactureras españolas utilizando el modelo estructural. A continuación, se analiza la importancia y los problemas que surgen al trabajar con microdatos provenientes de encuestas en las que se pregunta directamente a las empresas acerca de sus procesos de innovación. En la sección siguiente, se presentan las estimaciones del modelo estructural siguiendo la metodología del estudio de *Griffith et al (2006)*. Para luego, exponer una serie de debilidades tanto teóricas como empíricas de este planteamiento y, proponer extensiones alternativas. En el último apartado, se presenta una valoración global, en la que se exponen las conclusiones del trabajo y posibles futuras líneas de investigación en este campo.

II. Marco teórico

A pesar del avance en el conocimiento de los procesos de innovación a nivel de empresa debido al surgimiento de una gran cantidad de estudios en los últimos años, aún existen lagunas importantes en cuanto a los factores determinantes de estos procesos y su impacto económico.

Una parte del grado de ignorancia se debe a la dificultad en distinguir y medir adecuadamente distintos tipos de innovación, dado que ésta puede referirse a productos o procesos nuevos (innovación tecnológica) pero también a innovaciones organizativas, de gestión y de servicios; innovaciones que además pueden ser complementarias. A su vez, los costes y beneficios de innovar dependen de diversos elementos del conjunto del sistema económico como: el funcionamiento del mercado de personal altamente cualificado, el funcionamiento de las universidades y de los mercados de productos y servicios, la facilidad para establecer nuevas empresas, y el grado de desarrollo tanto de los mercados financieros, como del capital humano de la empresa.

La metodología empírica basada en modelos estructurales de función de producción conjuntamente con las nuevas bases de datos sobre innovación a nivel de empresa, constituyen herramientas de suma utilidad para indagar y contrastar empíricamente las diversas hipótesis referidas a los procesos de innovación.

II.1. Literatura e hipótesis

Siguiendo una “tradición Schumpeteriana”, a lo largo de las últimas décadas los diversos estudios empíricos se han focalizado básicamente en el análisis del impacto de diversas variables económicas sobre las inversiones en investigación y las innovaciones realizadas por las empresas.

Un primer tipo de hipótesis se basa en la relación entre innovación y diversas características de la empresa relacionadas entre sí. Por ejemplo, la relación de los procesos de innovación de la empresa con: su tamaño, edad, cuota de mercado, sector industrial al que pertenece, su pertenencia a un grupo industrial y las restricciones financieras que enfrenta.

Estudios recientes exploran la importancia de diferentes características intrínsecas del proceso innovador, por ejemplo: las fuentes de información utilizadas para innovar, los obstáculos que enfrenta la empresa para innovar exitosamente y las condiciones de apropiabilidad de los resultados de la innovación.

Existen ciertos factores externos a la empresa que propician la generación de innovaciones. Mientras que algunos estudios destacan la importancia de la demanda y las oportunidades tecnológicas para impulsar los procesos de innovación empresarial. Otros, argumentan que el incremento de la competencia internacional que enfrentan las empresas es central para comprender sus procesos de innovación tecnológica. Las empresas compiten en mercados globales en los que se observan constantes cambios en las tecnologías. La generación de innovaciones se ha convertido en una herramienta estratégica de suma importancia tanto para el desarrollo de las empresas como para su supervivencia en el mercado.

Innovación y tamaño de la empresa:

Schumpeter (1911) concebía a la empresa innovadora como el fruto del trabajo de un individuo extraordinario. En las décadas siguientes, resultado del estudio de las actuales economías desarrolladas, y del papel de las grandes corporaciones como líderes de este

proceso, resaltó la importancia de las grandes empresas como una de las fuentes principales de generación e implementación de los nuevos conocimientos.³

Existen varios argumentos que sustentan la hipótesis que las grandes empresas son más innovadoras que el resto. Primero, no se ven afectadas por restricciones de liquidez dado su fácil acceso al financiamiento externo y mayor disponibilidad de fondos propios. Segundo, se caracterizan por un alto grado de diversificación productiva que las ayuda a tratar con la incertidumbre de la innovación. Por último, la existencia de economías de escala y las diferencias en la organización del trabajo hacen a estas más innovadoras.⁴

Por el contrario, los trabajos de *Winter (1984)* y *Acs et al. (1988)* sustentan la hipótesis que las pequeñas empresas enfrentan un entorno económico y tecnológico diferente que puede ser propicio para la innovación. En el mismo sentido, el estudio de *Cohen y Klepper (1996)*, focalizado en la relación entre el tamaño de la empresa y su actividad innovadora, estableció una serie de hechos estilizados que sugieren que no siempre las grandes empresas son más innovadoras que las pequeñas.⁵

En una industria típica, por encima de un umbral de tamaño modesto, todas las empresas realizan actividades de I+D, no incrementándose más que proporcionalmente este esfuerzo con su tamaño. A su vez, entre las empresas que realizan actividades de I+D, el número de patentes e innovaciones por dólar invertido en estas actividades decrece con su tamaño y/o el nivel de gasto y, considerando todas las empresas, las pequeñas reportan un número desproporcionadamente grande de patentes e innovaciones con relación a su tamaño. Finalmente, según estos hechos el tamaño de la empresa no explica los esfuerzos realizados en I+D pero sí su desempeño económico pues son las pequeñas empresas las que obtienen mayores tasas de crecimiento.

³ En su edición revisada de la “Teoría del Desarrollo Económico” de 1934, *Schumpeter* argumenta que mientras que se espera que las empresas pequeñas y recién establecidas sean innovadoras cuando el cambio tecnológico es considerado como un proceso de “destrucción creadora”, cuando lo que destaca es el proceso de “acumulación creativa” son las empresas grandes y ya establecidas las que lideran el rol del proceso innovador.

⁴ Para un ejemplo empírico sobre este punto se recomienda el estudio de *Mairesse y Mohen (2005)*.

⁵ La investigación se basa en la revisión de varios trabajos empíricos realizados para empresas estadounidenses durante las décadas de 1960, 1970 y 1980. A pesar de diferir en el período de tiempo analizado, la muestra de empresas e industrias utilizada y las medidas de actividad innovadora y tamaño de empresa; encuentran una serie de hallazgos comunes acerca de la relación entre el tamaño de la empresa y su esfuerzo en I+D.

Innovación y poder de mercado:

Otro elemento importante, basado nuevamente en los trabajos de *Schumpeter*, se focaliza en la relación de la innovación con el poder de mercado de la empresa.

Por un lado, la presión competitiva actual del mercado puede ser un incentivo para innovar y obtener un mayor poder de mercado en el futuro. Por otra parte, el poder ex ante genera medios financieros para innovar y reduce los niveles de riesgo que todo proceso de innovación implica.⁶

Innovación y condiciones de apropiabilidad

Los incentivos a innovar dependerán del alcance con el cual los resultados de las actividades de innovación pueden ser apropiados o difundirse fácilmente dentro o fuera de la empresa o industria. Existen además de los mecanismos legales por los cuales se pueden proteger las innovaciones, mecanismos estratégicos como la protección de la información o de la complejidad de la innovación que repercuten en el tiempo de liderazgo sobre los competidores.⁷

En la literatura se encuentran dos efectos opuestos de los niveles de apropiabilidad. Por un lado, un bajo nivel puede conducir a un efecto de desincentivo. Las empresas reducen la realización de actividades internas relacionadas con la innovación porque son incapaces de apropiarse completamente de los beneficios de su innovación.⁸

Sin embargo, bajos niveles de apropiabilidad permiten altos niveles de derrames de conocimientos. Para poder capitalizar estos derrames, las empresas necesitan desarrollar una suficiente capacidad de absorción lo cual conduce una mayor inversión en la realización de actividades de I+D internas.⁹

⁶ Para un detalle sobre este punto ver *Bozeman y Link* (1983). A su vez, durante la década de los noventas ha surgido una plorífica literatura focalizada específicamente en la relación existente entre la estructura del mercado y la innovación a nivel de empresa. Ejemplos destacados de estos últimos aportes son los trabajos de *Geroski 1990* y *1991*, y *Aghion et al. (2005)*.

⁷ Ver *Mansfield, (1985)*.

⁸ Hipótesis basada en los trabajos de *Arrow (1962)* y *Spence (1984)*.

⁹ Ejemplo destacado que sustenta este segundo tipo de hipótesis es el trabajo de *Cohen y Levinthal (1990)*.

Innovación y sector industrial.

Las características de la industria y tecnología son determinantes cruciales del comportamiento innovador de las empresas. Dos fuerzas externas a la empresa la llevan a realizar actividades de I+D y eventualmente a innovar.

Una de ellas, conocida como la hipótesis de “arrastre de la demanda”,¹⁰ explica los procesos de innovación a través de la importancia de los factores de demanda, tales como el crecimiento de mercado. Para innovar no sólo el tamaño y crecimiento del mercado importan, sino también la disponibilidad a pagar de los consumidores por productos nuevos o mejorados.

Otra fuerza externa se basa en la importancia de los avances científicos que al estimular los esfuerzos innovadores de las empresas, influencian el sendero y la tasa de avance tecnológico (“empuje” tecnológico). Los avances científicos permiten generar innovaciones a las empresas por medio del desarrollo de nuevos conceptos o a través de su incorporación a nuevas maquinarias o insumos.¹¹

Como veremos más adelante, este conjunto de hipótesis se han tenido en cuenta, al menos parcialmente, en el diseño de las encuestas sobre innovación a nivel de empresa. Esto ha enriquecido la estimación de los modelos estructurales de función de producción.

II.2. El modelo estructural de Función de Producción. Un sistema de ecuaciones secuenciales

II.2.1. El modelo uni-ecuacional de función de producción

Aunque tenían en cuenta que la I+D eran sólo uno de los muchos factores del progreso tecnológico, los primeros trabajos empíricos de “función de producción” focalizaron su atención en estas actividades.¹² La ventaja de la utilización de estas medidas aproximadas del progreso tecnológico era su fácil cuantificación y medición, lo cual permitió construir indicadores que se podían incorporar explícitamente en los modelos y estudios econométricos.

¹⁰ Usualmente atribuida a Schmookler (1966).

¹¹ Ver Scherer (1965) y Levin y Reiss (1988)

¹² Por ejemplo, Denison (1985) estimó que la contribución de estas actividades al crecimiento económico sólo representaba el 20% del mismo.

Estos estudios parten de un marco teórico con una única ecuación de función de producción ampliada, tipo *Cobb-Douglas*, que incluye como insumo adicional una medida del esfuerzo acumulado en actividades de I+D por parte de la empresa:

$$\log Y = a(t) + \beta(\log X) + \gamma_K (\log K) + u \quad (1)$$

La variable Y es una medida de producción de la empresa, X un vector de insumos factoriales tradicionales, K es una medida del esfuerzo acumulado en actividades de I+D por parte de la empresa, $a(t)$ representa otras fuerzas que afectan a su producto y que varían sistemáticamente en el tiempo, mientras que u refleja otras fluctuaciones de carácter aleatorio en el mismo.

En este caso, el foco de atención se encontraba tanto en la definición y medición de la variable K (“capital de conocimiento” de la empresa), como en la estimación del parámetro γ_K , la elasticidad del producto respecto al esfuerzo acumulado en I+D. Al considerar a K como una medida de este esfuerzo, una de las formas más usuales de medirla fue a través del gasto acumulado en actividades de I+D realizado por la empresa.¹³

A pesar que dichos estudios contaban con una serie de problemas estadísticos de consideración,¹⁴ su limitación principal era que sólo medían la relación entre insumos del proceso de innovación, como son las actividades de I+D, y una medida de desempeño económico de la empresa; no tomando en cuenta el lazo que existe entre la inversión en investigación y la generación de innovaciones.

Ya en su trabajo de 1979, *Griliches* sostenía que el planteamiento de un modelo uniecuacional era una simplificación excesiva para el análisis de proceso de generación de

¹³ Ver trabajo de 4 créditos para un detalle sobre la construcción del “capital de conocimiento” de la empresa.

¹⁴ Se podían cometer errores de especificación, debido a la omisión de variables relevantes que caracterizan a las empresas o a las industrias. A su vez, existía la posibilidad que las empresas exitosas, y con poder económico importante, fueran a la vez más productivas y pudieran invertir recursos en actividades de I+D, dando a lugar a un importante *sesgo de selección*. Una de las formas más usuales para reducir este sesgo fue trabajar con tasas de variación temporal de las variables analizadas, lo cual es equivalente a realizar un análisis “dentro” de la empresa, al controlar los efectos fijos específicos a la misma. La estimación de este segundo tipo de ecuaciones también condujo a estimaciones sesgadas y poco robustas. Una de las posibles razones era la colinealidad temporal del capital físico y la inversión en I+D. Otra, la no inclusión de una tendencia temporal para considerar el cambio tecnológico en la estimación de la función de producción. Ambos problemas, conducían a sobreestimaciones de la elasticidad del capital físico y del gasto en I+D.

innovaciones a nivel de empresa y su impacto económico. Se debía considerar que la inversión en investigación genera innovaciones, lo que conducía al concepto de “función de producción de innovaciones”. Esta función describía el proceso de transformación que va desde los insumos innovadores, por ejemplo gasto en actividades de I+D, hacia las innovaciones en sí mismas.

Los estudios subsecuentes de *Griliches* llevaron al planteamiento de la función de producción en un sistema de ecuaciones que permitía describir el proceso de transformación que va desde los insumos hasta las innovaciones en sí mismas y, a su vez, medir el impacto que tienen las innovaciones sobre el desempeño económico de las empresas.

II.2.2. El sistema multi-ecuacional de función de producción

Es en el trabajo de *Pakes y Griliches (1984)* donde se sugiere un nuevo modelo de función de producción multi-ecuacional.

La serie de procesos que resume la función de producción a nivel de empresa se puede describir adecuadamente a través de la siguiente figura:

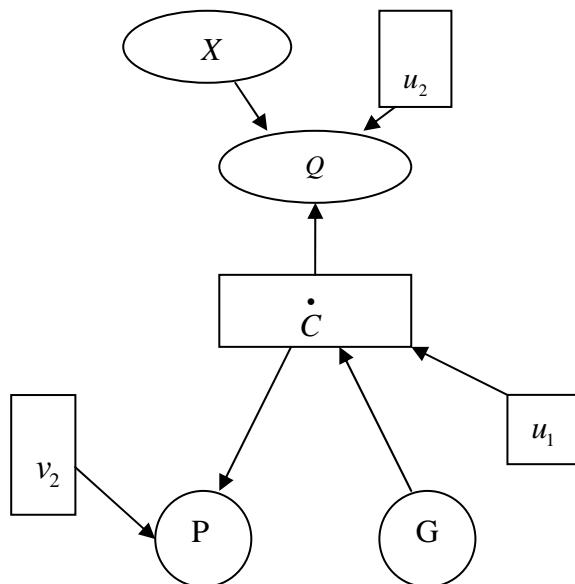


Figura 1. Análisis de sendero del modelo de función de producción.

La variable C define el nivel económicamente valioso de conocimiento tecnológico de la empresa en determinado momento del tiempo. Por tanto, $\dot{C} = dC/dt$ se puede interpretar como la creación neta de conocimiento tecnológico por unidad de tiempo. Esta variación temporal se analiza a través de una función específica, la cual combina

los insumos relacionados con los gastos realizados en investigación por parte de la empresa (variable G) y las perturbaciones u_1 , para producir conocimiento.

La variación temporal de C , que es central en el diagrama, no es observable, por lo que se hace necesario aproximarla con variables que si se puedan observar. Ante la falta de indicadores directos de las innovaciones, consideran a las patentes (P) como un indicador imperfecto de su número, siendo v_2 el ruido aleatorio en su relación con \dot{C} .

La variable P conjuntamente con los factores de producción tradicionales (resumidos en la variable X), y las perturbaciones (u_2), determinan la magnitud de varios indicadores interrelacionados de desempeño económico de la empresa, resumidos en la variable Q .

Siguiendo esta lógica de análisis es posible expresar el proceso de innovación empresarial mediante un sistema de ecuaciones. En este sistema, compuesto por tres funciones de producción secuenciales, cada ecuación pretende describir un aspecto del proceso.

Inicialmente las empresas deciden si emprender o no actividades tecnológicas; luego, eligen la intensidad de estas actividades que determinan la producción de innovaciones, para finalmente, medir su impacto, conjuntamente con el resto de los factores productivos, sobre alguna medida de desempeño económico de la empresa:

$$G = \beta_0^1 + \sum_m \beta_m^1 X_m^1 + u_0 \quad (2)$$

$$P = \beta_0^2 + \beta_G G + \sum_l \beta_l^2 X_l^2 + u_1 \quad (3)$$

$$Q = \beta_0^3 + \beta_P P + \sum_j \beta_j^3 X_j^3 + u_2 \quad (4)$$

La ecuación (2) explica el gasto en actividades de I+D por parte de la empresa, que se considera como un insumo para innovar. La ecuación (3) considera la generación de innovaciones. Y, finalmente, la ecuación (4) describe el comportamiento de alguna medida de desempeño económico de la empresas (Q).¹⁵ Los vectores X^0 , X^1 , X^2

¹⁵ Esta medida puede incluir variables como el valor de mercado de la empresa o alguna medida de productividad de los factores de producción.

resumen los insumos que explican el comportamiento de las empresas con relación al gasto en I+D, generación de innovaciones y desempeño económico, respectivamente.

La función de generación de innovaciones incluye como variable explicativa el gasto en actividades de I+D, representando el coeficiente β_G la elasticidad de las innovaciones respecto a estos gastos. En la ecuación (4) el coeficiente β_j representa la elasticidad de la medida de desempeño económico de las empresas con respecto al vector de insumos factoriales, y β_P es la elasticidad de esta variable con respecto a las innovaciones.

A diferencia del modelo uni-ecuacional de función de producción, es una medida de producto del proceso innovador y no una relacionada con sus insumos, la que conjuntamente con los factores productivos tradicionales, determinan diferencias observables en el desempeño económico de las empresas.

III. El Análisis econométrico del modelo estructural

III.1. Problemas específicos

El objetivo fundamental de los estudios basados en el modelo estructural de función de producción es obtener estimaciones consistentes del efecto causal de la inversión en I+D sobre las innovaciones y de estas sobre el desempeño económico de las empresas. Por tanto, los métodos econométricos utilizados intentan corregir los problemas inherentes a este tipo de modelos así como a la naturaleza de los datos provenientes de las encuestas.

Dado que cada una de las variables dependientes de las diferentes ecuaciones del modelo se comporta de forma diferente su tratamiento econométrico es distinto.

Muchas variables que surgen de las encuestas empresariales son cualitativas por naturaleza, en la mayoría de los casos se tienen respuestas binarias que requieren la aplicación de técnicas de variable dependiente. Estas respuestas son simples de modelizar a través de modelos de elección discreta, y su estimación se realiza generalmente a través del método de máxima verosimilitud pues, la distribución de los datos necesariamente se define como un modelo tipo Bernoulli.

En estos casos, la regresión de tipo lineal estimada por mínimos cuadrados ordinarios ignora el carácter discreto de la variable dependiente, y al no utilizar una función de

distribución acumulativa para parametrizar la probabilidad de ocurrencia del evento de interés, muchas veces no restringe las probabilidades predichas al intervalo $(0, 1)$.¹⁶

Dos parametrizaciones comúnmente utilizadas son los modelos probit o logit. Estos modelos especifican diferentes formas funcionales de la probabilidad de observar el evento de interés.

Por ejemplo, en el caso de la ecuación de investigación de cierta empresa i , aunque el interés está en explicar el comportamiento de la variable aleatoria continua no observada o variable latente g_i^* ,¹⁷ únicamente se observa la variable binaria g_i , que toma valores 0 o 1 de acuerdo a si g_i^* cruza cierto umbral, por ejemplo fijado en cero, o no:

$$g_i^* = x_{i0} b_0 + u_{i0} \quad (5)$$

No observamos g_i^* , en su lugar observamos:

$$g_i = \begin{cases} 1 & \text{si } g_i^* > 0 \\ 0 & \text{en otro caso} \end{cases}$$

Entonces:

$$\Pr(g_i = 1 / x_{i0}) = \Pr(g_i^* > 0) = \Pr(x_{i0} b_0 + u_{i0} > 0) = \Pr(-u_{i0} < x_{i0} b_0) = F(x_{i0} b_0) \quad (6)$$

La función F es la función de distribución condicional de $-u$, que iguala a la función de distribución condicional de u en el caso que la función de densidad sea simétrica en 0.

El modelo logit surge si se supone una distribución tipo logística del término de error del modelo de variable latente:

$$\Lambda(u_{i0}) = \exp(u_{i0}) / (1 + \exp(u_{i0})) , \text{ con } -\infty < u_{i0} < \infty \quad (7)$$

¹⁶ De todas formas, dicha estimación puede ser útil como herramienta exploratoria pues, aunque sea un modelo pobre para las probabilidades individuales, brinda una estimación razonablemente directa del efecto marginal del promedio muestral sobre la probabilidad de ocurrencia del evento de interés a medida que varía el valor de los regresores. En la práctica brinda una guía para observar qué variables son estadísticamente significativas en la relación estimada.

¹⁷ Por latentes se entiende a variables que se observan de forma incompleta. Estas se pueden introducir en los modelos de elección discreta a través de dos formas:

- a. Como un índice de la propensión, no observada, a que ocurra el evento de interés.
- b. Como la diferencia en la utilidad si ocurre el evento, lo cual presume que la variable binaria observada es el resultado de una elección individual.

Si se supone una distribución normal estándar del término de error, entonces se tiene un modelo probit:

$$\Pr(-u_{i0} < x_{i0}' b_0) = \Phi(x_{i0}' b_0) \quad (8)$$

Esta última característica hace que el modelo probit cuente con la atracción de extenderse naturalmente para la construcción de otros modelos, como por ejemplo los modelos Tobit.

Por otra parte, dada la metodología con la que se elaboran los cuestionarios de las encuestas, mucha de la información se encuentra disponible únicamente para las empresas innovadoras. Esto determina que exista un problema importante de *sesgo de selección* que exige trabajar con modelos econométricos de variable dependiente censurada.

Una forma de tratar con la censura consiste en aplicar una aproximación parámetrica completa que especifica la distribución condicional a los regresores de la variable latente.

Por ejemplo, si se suponen la función de densidad y de distribución acumulativa de cierta variable latente k_i^* condicional a cierto vector de variables x_{i1} : $f(k_i^* / x_{i1}), F(k_i^* / x_{i1})$; entonces al plantear el mecanismo de censura como una transformación de k_i^* : $k_i = h(k_i^*)$, es posible obtener las funciones $f(k_i / x_{i1})$ y $F(k_i / x_{i1})$ correspondientes a la variable observada k_i .

Sin embargo, la limitación principal de esta aproximación parámetrica es que se basa en fuertes supuestos distribucionales. Una segunda posibilidad, es intentar minimizar o salvar estos supuestos paramétricos a través de una estimación semiparamétrica.

Optar por una especificación parámetrica exige trabajar usualmente con modelos tipo Tobit. Al aplicar un modelo de este tipo se puede pensar, por ejemplo, que las empresas resuelven un problema de optimización al decidir el monto de recursos que van a destinar a cierta actividad y, para algunas de ellas, la elección óptima será una solución de esquina ($k_i = 0$, no gastar). En este caso el valor esperado condicional de la variable

dependiente observada, $E(k_i / x_{i1})$ cuando $k_i \geq 0$, es no lineal en x_{i1} (modelo Tobit tipo 1).¹⁸

El modelo Tobit tipo 1 es aquel con censura por debajo de cero, donde la variable latente k_i^* se supone lineal en los regresores con término de error aditivo que se distribuye normal y homoscedástico:

$$\begin{aligned} k_i^* &= x_{i1}b_1 + u_{i1} & (9) \\ u_{i1} &\sim N(0, \sigma^2) \end{aligned}$$

Dicho modelo expresa la variable observada k_i en función de una variable latente k_i^* :

$$\begin{aligned} k_i &= 0 & si & k_i^* \leq 0 \\ k_i &= k_i^* & si & k_i^* > 0 \end{aligned} \quad (10)$$

La distribución que siguen los datos de la muestra es una mezcla entre una distribución continua y otra discreta, existiendo un punto de acumulación de probabilidad en el punto de censura 0.

Sin embargo, el modelo tobit tipo 1 es adecuado cuando se supone que los ceros observados se deben a una sola causa. Estos modelos restringen el mecanismo de censura a ser del mismo modelo que genera la variable dependiente. Para obtener una aproximación más adecuada, el mecanismo de censura y de respuesta deberían modelarse utilizando procesos separados. Existen dos aproximaciones a tal generalización: el *modelo de dos partes*, que especifica un modelo para el mecanismo de censura y otro para la respuesta condicional a que se observada y,¹⁹ el *modelo de selección muestral*, que especifica una distribución conjunta para los mecanismos de censura y de respuesta, para luego especificar la distribución condicional de la respuesta observada.

¹⁸ Otra alternativa sería aplicar un modelo exponencial estimando por mínimos cuadrado no lineales (MCNL), lo cual asegura que los valores predichos sean positivos. Sin embargo, esto tiene dos limitaciones importantes. En primer lugar, si la variable observada k es una solución de esquina, su varianza condicional probablemente sea heteroscedástica y, por tanto, los estimadores MCNL podrían ser ineficientes. Aunque, este problema podría resolverse al menos parcialmente utilizando MCNL ponderados, cualquier modelo para la varianza condicional sería arbitrario. Sin embargo, una desventaja más importante, es que no seríamos capaces de medir el efecto de cada variable explicativa sobre otras características de la distribución condicional de k , como por ejemplo $\Pr(k = 0 / x_1)$ y $E(k / x_1, k > 0)$.

¹⁹ Este modelo fue presentado por Cragg (1971) como una generalización del modelo Tobit tipo 1.

En el caso del *modelo de dos partes* una posible formulación es considerar que se observa la variable latente si esta es mayor a cero, el modelo se visualiza como un “modelo de valla” dado que al cruzar este umbral se observa la participación. Se utiliza el método de máxima verosimilitud al estimar separadamente la ecuación de elección discreta utilizando todas las observaciones y, luego, los parámetros de la función de densidad de la variable de intensidad utilizando únicamente las observaciones que son mayores a cero.

En el caso del *modelo de selección muestral* se tiene una ecuación de participación y una ecuación de respuesta que se estiman conjuntamente. La primera ecuación toma en cuenta el hecho si la empresa decide participar en cierta actividad, por ejemplo realizar actividades de I+D o no, esto eventualmente se modela a través de un modelo probit o logit. En una segunda etapa, se considera la magnitud o intensidad de estas actividades, donde las empresas deciden el monto a invertir, condicional ha ya haber tomado la decisión de realizarlas. *Amemiya (1985)* denominó a este modelo Tobit tipo 2.

En el caso de la inversión en investigación por parte de cierta empresa i , el *modelo de selección muestral* incluirá:

- a. Una ecuación de participación (gasta o no en actividades de investigación):

$$g_i = \begin{cases} 1 & \text{si } g_i^* > 0 \\ 0 & \text{si } g_i^* \leq 0 \end{cases} \quad (11)$$

- b. Una ecuación de respuesta observada (intensidad del gasto en investigación):

$$k_i = \begin{cases} k_i^* & \text{si } g_i^* > 0 \\ - & \text{si } g_i^* \leq 0 \end{cases} \quad (12)$$

Este modelo especifica que la variable k_i se observa cuando $g_i^* > 0$; mientras que no tiene porque tomar ningún valor concreto cuando $g_i^* \leq 0$.

En estos casos, la estimación máxima verosímil se basa en el supuesto adicional que los errores correlacionados de ambas ecuaciones se distribuyen conjuntamente normal y son homoscedásticos:

$$\begin{pmatrix} u_{i0} \\ u_{i1} \end{pmatrix} \xrightarrow{iid} N_2 \left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_{u_{i0}}^2 & \rho\sigma_{u_{i0}}\sigma_{u_{i1}} \\ \rho\sigma_{u_{i0}}\sigma_{u_{i1}} & \sigma_{u_{i1}}^2 \end{pmatrix} \right)$$

Desde el punto de vista teórico, este *modelo de selección muestral* es identificable sin ninguna restricción sobre los regresores. Exactamente los mismos regresores pueden aparecer en ambas ecuaciones. Sin embargo, pueden surgir problemas de identificación debido a multicolinealidad. En estos casos es necesario imponer *restricciones de exclusión* (supuestos sobre la forma funcional de las dos variables latentes). A los efectos prácticos su estimación requiere que al menos un regresor en la ecuación de participación esté excluido de la ecuación de respuesta.

Otro problema econométrico importante que surge al aplicar el modelo estructural de función de producción es el surgimiento de *sesgo de simultaneidad* característico en los sistemas de ecuaciones. Una empresa que innova es probable que también incurra en gastos de I+D, con variables explicativas comunes para cada una de estas variables. A su vez, las innovaciones son endógenas en la ecuación de productividad. En ambos casos, las variables explicativas se pueden determinar conjuntamente con las variables dependientes. Además, las perturbaciones pueden reflejar en parte variables no observadas y efectos de empresa, que probablemente también estén correlacionadas.

La existencia conjunta de *sesgos de selectividad y simultaneidad* conduce a utilizar algún tipo de estimador de sistemas de ecuaciones simultáneas. En algunos de los estudios analizados se ha observado la utilización de un procedimiento de variables instrumentales. Una vez que se ha decidido el conjunto de instrumentos a utilizar, las ecuaciones del sistema no lineal se estiman a través de mínimos cuadrados en dos o tres etapas o método generalizado de los momentos aplicados directamente a las ecuaciones estructurales.

En gran parte de los estudios analizados la forma más frecuente de instrumentar las variables endógenas de I+D en la ecuación de innovación y de la innovación en la ecuación de productividad, es aplicar un procedimiento de estimación de dos o tres etapas, en el cual se sustituyen las variables endógenas por sus estimaciones de la etapa anterior, tratando de reproducir el proceso de mínimos cuadrados en dos o tres etapas.

III.2. El modelo de Crépon Duguet y Mairesse

Fue el trabajo de *Crépon, Duguet y Mairesse (1998)* (de ahora en adelante CDM), realizado para Francia, el primero capaz de unir las líneas de investigación empírica en un modelo estructural similar al planteado por *Pakes y Griliches (1984)*.

Tomando en cuenta el aporte de estos autores y considerando la estructura de los datos provenientes de las encuestas sobre innovación empresarial, el modelo CDM consiste en un sistema ecuacional no-lineal y recursivo sin efecto retroalimentación que se formaliza en cuatro ecuaciones.

La primera ecuación describe la decisión de la empresa de dedicar un esfuerzo suficiente a actividades de I+D internas, lo cual resulta en una inversión observable en estas actividades (hacer I+D interna o no). La segunda describe la intensidad con la cual la empresa realiza estas actividades (“función de inversión en I+D”). Una tercera ecuación especifica la función de producción de conocimiento o innovaciones en la cual los gastos en I+D son uno de sus determinantes. Y, finalmente, la última ecuación es la función de producción de la empresa donde el conocimiento es considerado como un insumo más conjuntamente con el resto de factores productivos.

Dos relaciones vinculan el capital de conocimiento de la empresa i con sus determinantes:

$$g_i^* = x_{i0} b_0 + u_{i0} \quad (13)$$

$$k_i^* = x_{i1} b_1 + u_{i1} \quad (14)$$

En una primera etapa, la empresa decide si realizar o no actividades de investigación (ecuación 13). Esta decisión se basa en algún criterio determinado, por ejemplo en el valor presente esperado neto de la inversión en actividades de I+D internas, recogido por la variable dependiente latente g_i^* .

Una posibilidad es observar la variable g_i^* de forma limitada, de manera que se observa únicamente si supera cierto umbral y la empresa decide declarar o no tales gastos. Por tanto, si se dispone de una muestra que contiene empresas que declaran y empresas que no declaran realizar I+D es posible plantear la siguiente ecuación:

$$g_i = \begin{cases} 1 & \text{si } g_i^* > 0 \\ 0 & \text{en otro caso} \end{cases} \quad (13')$$

En este caso, g_i es una variable binaria endógena igual a cero para las empresas que no reportan la realización de actividades de I+D internas e igual a uno para aquellas que sí lo hacen. El vector x_{i0} es de variables explicativas, b_0 es el vector de coeficientes asociado y u_{i0} el término de error.

En una segunda etapa (ecuación 14), las empresas deciden el monto a invertir en actividades de investigación internas, que representa su inversión en conocimiento científico-tecnológico. Se supone una intensidad latente de estas actividades k_i^* , siendo x_{i1} el vector de variables explicativas, b_1 el vector de coeficientes asociado y u_{i1} es el término de error.

La variable latente k_i^* es observable cuando la variable g_i^* es mayor a cierto umbral mínimo. Por ello se necesita especificar la distribución conjunta a los efectos de obtener un modelo estimable. Como ya fuera señalado en la sección III.1, una posibilidad basada en la literatura estándar relacionada con lo especificación de modelos tobit tipo 2, es suponer que los términos de error de las ecuaciones (13) y (14) siguen una distribución normal bivariada.

La siguiente ecuación del modelo es la función de producción de innovaciones:

$$I_i^* = k_i^* \gamma_{k^*} + x_{i2} b_2 + u_{i2} \quad (15)$$

Donde, I_i^* representa la variable latente referida a las innovaciones introducidas por la empresa i . Como en los casos anteriores, se supone que la introducción de innovaciones no se observa plenamente sino de forma limitada. Por ejemplo, se puede observar si la empresa patenta o no, y en caso positivo el número de patentes solicitadas u obtenidas. O bien la empresa puede declarar si ha introducido o no innovaciones de proceso o de producto a lo largo de determinado período. Por ello, el modelo empírico se puede especificar como sigue:

$$I_i = \begin{cases} 1 & \text{si } I_i^* > 0 \\ 0 & \text{en otro caso} \end{cases} \quad (15')$$

En este caso, I_i es una variable binaria endógena igual a cero para las empresas que no reportan la obtención de innovaciones e igual a uno para aquellas que sí lo hacen; k_i^* es la variable latente de investigación y x_{i2} es un vector de otros factores determinantes de la función de producción de innovaciones. El coeficiente γ_{k^*} es la elasticidad de la probabilidad de observar una innovación con relación al capital de conocimiento, una medida de los impactos o retornos de la investigación interna sobre la innovación.

Mientras que, b_2 es el vector de coeficientes asociado al resto de variables explicativas y u_{i2} el término de error.

Finalmente, en el marco de la consideración de una función de producción ampliada tipo *Cobb-Douglas*, se supone que las empresas producen a través de una tecnología con retornos constantes a escala utilizando, además de los insumos factoriales tradicionales, las innovaciones introducidas por la empresa que no están recogidas en el capital físico:

$$q_i = I_i^* \alpha_{I^*} + x_{i3} b_3 + u_{i3} \quad (16)$$

La variable q_i representa una medida de desempeño económico de la empresa i . El coeficiente α_{I^*} es la elasticidad de la medida de desempeño con relación a la innovación, mientras que b_3 está compuesto por los coeficientes que resumen la elasticidad a escala de los insumos factoriales resumidos en x_{i3} .

III.3. Otras aplicaciones del modelo estructural

A partir del modelo original CDM las investigaciones subsecuentes han avanzado tanto en su sofisticación teórica como empírica, esto ha permitido ampliar el conocimiento sobre los procesos de innovación a nivel de empresa y su impacto económico.

Si bien dichos estudios tienen como objetivo común establecer de forma empírica una caracterización de las empresas innovadoras, difieren en cuanto al planteamiento del modelo estructural, el análisis econométrico y la elección de las variables endógenas y explicativas.

En el siguiente cuadro se exponen los principales estudios analizados, destacando las variables endógenas y exógenas utilizadas, así como el método de estimación empleado.

| Estudio | Datos a nivel de empresa | Variables endógenas | Variables independientes | Método de estimación | Otros Comentarios |
|---|-------------------------------------|---|---|---|--------------------------|
| Crépon, Duguet, Mairesse (1998) y | Francia 1986-1990 | Gastos en I+D (deflactados y depreciados) Patentes y proporción de ventas de innovaciones en el total de ventas Valor agregado por trabajador | Tamaño de empresa, cuota de mercado, diversificación productiva, arrastre de demanda y condiciones de oportunidad tecnológica, sector industrial Mismas que ecuación anterior incluyendo la variable de intensidad en I+D Innovaciones, tamaño de empresa, capital físico por trabajador, personal calificado, sector industrial | Estimador de Momentos en primera etapa y Mínimos Cuadrados Asintóticos (ALS) en segunda etapa | |
| Benavente (2006) | Chile 1995-1998 | Gasto en I+D por trabajador Proporción de ventas de innovaciones en el total de ventas Productividad del trabajo | Tamaño de empresa, cuota de mercado, diversificación productiva Mismas que ecuación anterior incluyendo la variable de intensidad en I+D Tamaño de empresa, innovación, intensidad de capital | Estimador de Momentos en primera etapa y Mínimos Cuadrados Asintóticos (ALS) en segunda etapa | |
| Jefferson, Huambo, Xiaojing Xiaoayun (2006) y | China 1995-1999 (Panel de datos) | Gasto en I+D por trabajador Proporción de ventas de innovaciones en el total de ventas. Productividad Total de los Factores Rentabilidad | Tamaño de empresa, diversificación productiva, beneficios rezagados, intensidad en I+D rezagada, tipo de propiedad de la empresa, factores que dificultan la innovación Intensidad en I+D, interacción intensidad en I+D-tamaño, edad de la empresa, sector industrial, tipo de propiedad de la empresa Tamaño de empresa, proporción de ventas de innovaciones, capital físico y materiales, sector industrial, tipo de propiedad de la empresa Tamaño de la empresa, participación de ventas de innovaciones, capital físico, sector industrial, tipo de propiedad de la empresa | Estimación por Mínimos Cuadrados Ordinarios (OLS) y Variables Instrumentales (IV) | |

| Estudio | Datos a nivel de empresa | Variables endógenas | Variables independientes | Método de estimación | Otros Comentarios |
|---------------------------------|---------------------------------|--|--|---|--|
| Lööf y Heshmati (2002) | Suecia 1994-1996 | Gasto en innovación por trabajador Ventas de innovaciones por trabajador (distinguen entre innovaciones nuevas para el mercado (radicales) y nuevas para la empresa (incrementales)) Productividad del trabajo | Tamaño de empresa, capital humano, indicador de rentabilidad de la empresa, estrategia de innovación de la empresa, cooperación para innovar, fuentes de información para innovar, factores que dificultan la innovación, innovación de proceso, sector industrial Intensidad en I+D, indicador de productividad, capital humano, cooperación para innovar, fuentes de información para innovar, factores que dificultan la innovación, sector industrial, tamaño de la empresa, innovaciones de proceso Productividad del trabajo rezagada, innovación, tamaño de empresa, personal calificado, indicador de fusión con otras empresas, intensidad factorial, innovación de proceso | Máxima Verosimilitud (ML) para el modelo Tobit generalizado de gastos en I+D, las otras ecuaciones se estiman por Mínimos Cuadrados en tres etapas (3SLS) con corrección de sesgo de selección | Efecto retroalimentación desde la productividad hacia las innovaciones |
| Lööf y Heshmati (2002) | Suecia 1996-1998 | Gasto en I+D por trabajador Ventas de innovaciones por trabajador (distinguen entre innovaciones nuevas para el mercado y nuevas para la empresa) Valor agregado por trabajador | Tamaño de empresa, indicador de intensidad de capital de capital físico, capital humano, indicador de grado de competencia en el mercado de la empresa, factores que dificultan la innovación Tamaño de empresa, intensidad en I+D, intensidad de capital físico, crecimiento de mercado, fuentes de conocimiento, ratio de Mill Tamaño de empresa, ventas de innovaciones por trabajador, capital humano, tipo de innovación | Estimación por ML para el modelo Tobit generalizado de gasto en I+D, mientras que las otras ecuaciones se estiman por Mínimos Cuadrados en dos etapas (2SLS) con corrección de sesgo de selección | También se estima el modelo sólo para innovaciones radicales; la ecuación de productividad es estimada en niveles y tasas de crecimiento; efecto retroalimentación desde la productividad hacia las innovaciones |

| Estudio | Datos a nivel de empresa | Variables endógenas | Variables independientes | Método de estimación | Otros Comentarios |
|---------------------------------------|--|---|---|---|--|
| Lööf, Heshmati, Apslund y Nääs (2001) | Finlandia, Noregaa y Suecia 1994-1996 | Gasto en I+D por trabajador Ventas de innovaciones por trabajador (distinguen entre innovaciones incrementales y radicales) Productividad del trabajo | Sector industrial, tamaño de la empresa, exportaciones, patentes, proporción de administradores y no ingenieros en actividades de I+D, intensidad factorial (conocimiento, capital y trabajo), fusiones de empresas Mismas que ecuación anterior y se adicionan: intensidad en I+D, obstáculos a la innovación, fuentes de información para innovar, acuerdos de cooperación Innovación, sector industrial, intensidad factorial, innovación de proceso, razones para innovar, capital humano, fuentes conocimiento, factores que dificultan la innovación, tamaño de empresa, fusiones empresariales | ML para el Tobit generalizado de gasto en I+D, las otras ecuaciones se estiman por 2SLS y 3SLS con corrección de sesgo de selección | Estimación conjunta de todas las innovaciones y, luego, únicamente para las innovaciones radicales; efecto retroalimentación desde productividad hacia la innovación |
| Duguet (2006) | Francia 1986 1990 | Innovación radical e Innovación incremental Crecimiento de la PTF | Ventas, cuota de mercado, diversificación productiva, empuje tecnológico, arrastre de la demanda, I+D interno, I+D externo, patentes, tipos de bienes Innovación, sector industrial, PTF rezagada, arrastre de la demanda, empuje tecnológico, insumos para innovar, I+D | ML para las innovaciones, 2SLS y GMM para el crecimiento de la PTF | Estimación separada para varias oportunidades tecnológicas |
| van Leeuwen y Klomp (2006) | Holanda 1994-1996 | Intensidad de I+D: gasto total en I+D interno como porcentaje del total de ventas Intensidad de innovación Rentabilidad | Tamaño de empresa, cuota de mercado; arrastre de demanda, empuje tecnológico, variables de oportunidad tecnológica: CIENCIA (información tecnológica de la ciencia) y OTROS (información de otras empresas); facilidades permanentes en I+D, cooperación en I+D, indicadores financieros I+D más variables de ecuación anterior sin indicadores financieros Innovaciones, capital físico, materiales, trabajo, innovación de proceso, sector industrial, tamaño de empresa | OLS y 3SLS (con o sin corrección para la selectividad) | Efecto retroalimentación desde productividad a la ecuación de insumos para innovar |

| Estudio | Datos a nivel de empresa | Variables endógenas | Variables independientes | Método de estimación | Otros Comentarios |
|---|---|---|---|---|--|
| van Leeuwen (2002) | Holanda 1992-1996 | Intensidad en I+D: gasto total en I+D interno como porcentaje del total de ventas Participación de las ventas de innovaciones en el total de ventas Productividad (ingresos de la empresa por trabajador y valor agregado por trabajador) | Intensidad en I+D rezagada, tamaño de empresa, cuota de mercado, objetivos del proceso innovador, aspectos organizacionales y de entorno tecnológico, oportunidad tecnológica: facilidades permanentes en I+D, cooperación en I+D; variables financieras Mismas que ecuación anterior sin facilidades para act. de I+D, I+D, innovación de proceso, período Innovaciones, capital físico, materiales, trabajo, innovación de proceso, sector industrial, tamaño de la empresa, inversa del mark-up | ML con Información Completa | Efecto retroalimentación desde productividad a la ecuación de insumos para innovar |
| Criscuolo y Haskel (2003) | Reino Unido 1994-1996 1998-2000 (Panel de datos) | Gasto en I+D en el total de ventas Innovación PTF | Trabajadores calificados, rigidez organizacional, indicador de grado de regulación, grupo empresarial, ayuda financiera y no financiera, cuota de mercado, concentración industrial, protección estratégica y comercial, PTF relativa a la empresa en la frontera tecnológica al inicio del período, apropiabilidad de resultados, mercado de destino de la producción I+D, ayuda financiera, cooperación en I+D, I+D interna, grupo empresarial, personal calificado, fuentes de información Innovación de producto y de proceso, innovador original, imitador, capital humano, capital físico | Estimación por Mínimos Cuadrados Ordinarios (OLS) y Variables Instrumentales (IV) | |
| Parisi, Schiantarelli y Sembenelli (2005) | Italia 1992-1994 y 1995-1997 (Panel de datos) | Innovación de producto y de proceso Índice de Tornquist de crecimiento de la PTF | Tamaño de empresa, I+D, capital fijo, edad de la empresa, grupo empresarial, sector industrial, localización geográfica, crecimiento de la productividad (rentabilidad esperada de la empresa) y flujo de caja (condiciones financieras) Innovación, capital fijo (método de inventario permanente), trabajo (incluye todos los empleados excepto aquellos involucrados en actividades de I+D para evitar la doble contabilización), materiales | Generación de innovaciones de producto y proceso estimadas a través de Logit o Logit condicional. Efectos de la innovación sobre el crecimiento de la productividad estimada por IV | |

| Estudio | Datos a nivel de empresa | Variables endógenas | Variables independientes | Método de estimación | Otros Comentarios |
|---|---|---|---|---|--------------------------|
| Griffith, Huergo, Mairesse y Peters (2006) | Francia; Alemania, España y Reino Unido (CIS3) 1998 - 2000 | I+D Gasto en I+D por trabajador Innovaciones de producto o proceso Ventas por trabajador | Competencia internacional, ayuda pública, condiciones de apropiabilidad, tamaño de empresa, sector industrial, dummy para Alemania del Este Condiciones de demanda, cooperación, fuentes de información para innovar, ayuda pública Intensidad de I+D, indicadores de demanda y condiciones de apropiabilidad, fuentes de información, inversión por trabajador en innovación de proceso Innovaciones, inversión en capital físico | Tobit generalizado por ML para ecuaciones de I+D. Dos ecuaciones Probit para innovaciones de proceso y de producto | |

III.3.1. Algunos resultados generales

Las relaciones encontradas por las diferentes investigaciones se pueden sintetizar en tres ítems relacionados con los determinantes de las actividades de investigación y de innovación, y el desempeño económico de las empresas:

- La realización de actividades de I+D se incrementa significativamente con el tamaño de empresa y su diversificación productiva. Luego de controlar por tamaño y sector, se observa que la probabilidad de realizar tales actividades también se incrementa con su cuota de mercado.

Los indicadores relacionados con la demanda de mercado y nivel tecnológico sectorial tienen un impacto positivo y significativo sobre la decisión de realizar I+D.

Por su parte, tanto la cuota de mercado como su diversificación productiva tienen un impacto significativo y positivo sobre la intensidad del esfuerzo en I+D, al igual que los indicadores de arrastre de la demanda de mercado y empuje tecnológico.

Sin embargo, una vez que la empresa ha decidido realizar este tipo de actividades, la intensidad de las mismas no depende de su tamaño, lo cual ratifica uno de los hechos estilizados de *Cohen y Klepper (1996)*.

Existen trabajos que consideran variables adicionales para explicar tanto la decisión de realizar tanto actividades de I+D como su intensidad. *Lööf y Heshmati (2002)* encuentran que ambas decisiones dependen positivamente de la calificación de los recursos humanos que dispone la empresa, de su rentabilidad corriente y de la existencia de una estrategia para innovar. Por su parte, *Griffith et al (2006)* observan que las empresas que reciben fondos de fuentes gubernamentales, que operan en mercados internacionales y que hacen un mayor uso de métodos formales o estratégicos de protección de sus innovaciones, incrementan su probabilidad de realizar actividades de I+D.

- Se utilizan diferentes medidas de resultados del proceso de innovación. Por ejemplo, variables binarias de innovación (innovar o patentar durante el período o no), o medidas de intensidad con la cual se innovó (ventas de productos nuevos o mejorados, o el número de patentes obtenidas).

En términos generales, se encuentra una fuerte asociación entre la realización de actividades de I+D y la obtención de innovaciones. Sin embargo, una vez que se controla por sector y tamaño, la I+D no tiene un impacto significativo sobre la intensidad con la cual se innova.

Los efectos de los indicadores de demanda y de tecnología son significativos al explicar las ventas de innovaciones pero no sobre las patentes. Este resultado se puede explicar por la naturaleza dinámica del proceso de innovación para el cual es difícil esperar resultados instantáneos.

Las diferentes investigaciones confirman el impacto positivo y significativo para innovar de la existencia de mecanismos de protección tanto formales como informales.

Ejemplos de variables explicativas adicionales, con efectos significativos y positivos sobre la probabilidad de innovar y su intensidad, son las relacionadas con la cooperación universidades y proveedores de tecnología y, los intercambios de información con competidores y consumidores.

- Las diferentes estimaciones confirman la importancia de las innovaciones para un mejor desempeño económico de las empresas. En la gran mayoría de los estudios se encuentra un impacto significativo de alguna medida de innovación sobre el desempeño económico de las empresas.

A partir del modelo original CDM, se han realizado diversas extensiones. En el caso de la presente investigación, únicamente interesa comentar los trabajos que analizan el impacto diferenciado de los diferentes tipos de innovaciones sobre el desempeño económico de las empresas.²⁰

III.3.2. Extensiones del modelo estructural CDM: impacto diferenciado de las innovaciones

Si bien existen una serie importantes de trabajos que analizan separadamente los determinantes y el impacto de la generación de innovaciones nuevas para la empresa (“innovaciones incrementales”) o nuevas para el mercado (“innovaciones radicales”);²¹

²⁰ Para un mayor detalle ver trabajo de 4 créditos.

²¹ Por ejemplo *Lööf y Heshmati (2002)*, *Lööf et al. (2001)* o *Duguet (2006)*.

en la presente investigación interesa comentar únicamente aquellos trabajos que distinguen entre *innovaciones de producto* y *de proceso*.

En cuanto a la diferenciación de sus determinantes, el trabajo de *Griffith et al.* (2006) señala que son los proveedores una fuente importante de información para las *innovaciones de proceso*, mientras que los consumidores tienen una influencia significativa en las *innovaciones de producto*. A su vez, encuentra que la habilidad por parte de la empresa para proteger la innovación a través de métodos estratégicos o formales es menos importante para las innovaciones de proceso que para las de producto.

Con relación al impacto de estas innovaciones sobre el desempeño de las empresas existen resultados mixtos. Mientras que en el trabajo de *Criscuolo y Haskel* (2003) son las innovaciones de producto las que tienen un impacto significativo sobre el crecimiento de la Productividad Total de Factores de las empresas del Reino Unido. Los resultados de las estimaciones del trabajo de *Parisi et al.* (2005), muestran que son las innovaciones de proceso las que tienen un mayor impacto sobre la productividad de las empresas francesas. Lo cual es consecuente con los resultados obtenidos por el estudio de *Griffith et al.* (2006) para el caso de Francia.

III.3.3. Aplicaciones del modelo CDM a España

En el caso de España, no se han encontrado muchos trabajos que estimen de forma completa el modelo CDM utilizando datos provenientes de la Encuesta sobre innovación tecnológica en las empresas.

Existen sí una variedad de estudios empíricos que, utilizando datos provenientes de la Encuesta sobre Estrategias Empresariales (ESEE), caracterizan el comportamiento innovador de las empresas manufactureras españolas a través de metodologías muy similares al modelo estructural.²²

Dentro de estos últimos, es posible distinguir una variedad importante de trabajos que se focalizan en el análisis por separado de cada una de las tres variables endógenas del modelo estructural. Por ejemplo, *Huergo* (2002) describe algunos de los factores que

²² La encuesta ESEE además de estar orientada a la obtención de información sobre las características de los mercados en los que opera la empresa y, a datos contables de las mismas, indaga sobre las estrategias empresariales de corto y largo plazo, lo cual ha permitido la elaboración de estudios relacionados con las actividades de I+D de las empresas y su comportamiento innovador.

influyan la generación de innovaciones en la industria manufacturera española. *Labeaga y Ros (2003)* investigan la existencia de complementariedades entre las decisiones de llevar adelante tanto innovaciones de producto como de proceso focalizándose en el comportamiento de los gerentes. *Huergo y Jaumandreu (2004)*, controlando por el efecto edad de la empresa, analizan directamente el impacto de las innovaciones de proceso sobre el crecimiento de la productividad.

En base a datos de la encuesta ESSE para el período 1990-2002, el estudio de *Huergo y Moreno (2004)* es capaz de estimar de forma completa un modelo multi-ecuacional, al estilo CDM, para explicar el comportamiento innovador de las empresas manufactureras españolas. En este modelo, la primera ecuación describe la decisión de participación de las empresas en actividades tecnológicas. La segunda refiere a la intensidad de los insumos tecnológicos, medidos básicamente por la intensidad del gasto en I+D. La tercera aborda la generación de innovaciones a partir de factores tecnológicos, tanto internos como externos; y finalmente, la cuarta ecuación recoge el impacto de las innovaciones exitosas sobre el crecimiento de la empresa medido por el residuo de Solow.

Controlando por tamaño de empresa y sector industrial, las variables explicativas incluidas en las ecuaciones de investigación se relacionan con el entorno tecnológico de la empresa, las condiciones de la demanda y el mercado, la capacidad para apropiarse de los beneficios de las inversiones tecnológicas y las restricciones financieras.

Estiman dos funciones separadas de innovación, innovaciones de producto y proceso. Como variables explicativas incluyen: la predicción de la intensidad inversora en I+D, la intensidad de capital físico, e indicadores de la evolución expansiva de los mercados, de la evolución recesiva de los mercados, del grado homogeneidad producto y del grado de movilidad empresarial.

Consideran dos ecuaciones de crecimiento de la productividad, utilizando por separado la predicción de la innovación de proceso y la de producto y, controlando por indicadores de movimiento empresarial, sector de actividad, tamaño de empresa y año. Además, consideran la variación del grado de utilización de la capacidad instalada para recoger el impacto de los cambios en el grado de utilización de los insumos y un indicador de la variación ponderada de los factores.

Sus resultados están en la línea del trabajo de *Griffith, Huergo, Mairesse y Peters (2006)* (de ahora en adelante GHMP) para el caso de España. Es este último trabajo el que estima directamente el modelo completo CDM utilizando datos provenientes de la encuesta CIS para el período 1998 a 2000. Su objetivo es estimar este modelo básico para cuatro países (Francia, Alemania, España y Reino Unido), con el fin de poder comparar y determinar si existen diferencias destacables entre estos cuatro países en los vínculos entre esfuerzo en I+D, innovaciones obtenidas (de producto y de proceso) y productividad.

En el caso de España, los resultados señalan que:

- 1) La decisión de emprender o no actividades internas de I+D de forma estable depende de: la presencia de la empresa en mercados internacionales, el uso de métodos formales y estratégicos de protección de la propiedad intelectual, y la obtención de financiación pública. En España la importancia del tamaño de la empresa parece ser superior a la de otros países.
- 2) La intensidad del esfuerzo en I+D depende positivamente del financiamiento que obtenga la empresa y de medidas de protección estratégica, aunque no formales, de las innovaciones. A su vez, el tamaño de la empresa no influye significativamente en la magnitud de dicho esfuerzo.
- 3) La capacidad de introducir *innovaciones de producto* se relaciona con la intensidad de los esfuerzos en investigación, con el uso de métodos de protección tanto formales como estratégicos y, con la capacidad de utilizar información de los clientes y competidores. En el caso de las *innovaciones de proceso* dicha probabilidad, además de estar asociada a la intensidad de investigación y las medidas de protección formal, aunque no estratégicas, se relaciona con la información proveniente de proveedores y competidores y con la intensidad de la inversión en capital físico.
- 4) Por último, la productividad media por trabajador depende significativamente de una medida aproximada de capital físico y de la introducción de nuevos productos por parte de la empresa. Aunque, no se encuentra un efecto significativo de las innovaciones de proceso.

El estudio GHMP es uno de los pocos que permite comparar el comportamiento innovador de las empresas entre países y, por tanto, realizar indirectamente inferencias

sobre el papel que pueden tener las diferencias institucionales. Sin embargo, cuenta con tres limitaciones importantes. Por un lado, la representatividad de la muestra no es la misma para todos los países. La encuesta es obligatoria en los casos de España y Francia, pero en Alemania y Reino Unido es voluntaria, no quedando establecido el grado de inconsistencia de las estimaciones que de ello se puede derivar. En segundo lugar, posiblemente la necesidad de utilizar exactamente las mismas variables explicativas para cada país ha restringido las opciones de sofistificación del planteamiento teórico y empírico. Por último, probablemente algunas de las variables explicativas incluidas estén correlacionadas con el término aleatorio, ya sea por errores de medición o simultaneidad.

Dado que el estudio GHMP es el único antecedente analizado que aplica de forma completa el modelo estructural CDM a datos de empresas manufactureras provenientes de la encuesta CIS para el caso de España, se creyó que una metodología empírica válida sería reproducir dicho análisis, realizado para el período 1998-2000, para datos del período 2002-2004.

Al realizar un análisis de este tipo se persiguen dos objetivos. En una primera instancia, realizar un análisis comparativo que permita comprobar si se mantienen los resultados obtenidos por el estudio GHMP y, por ende, analizar la estabilidad temporal de las relaciones estimadas. Luego, en una segunda instancia, se pretende avanzar en la sofistificación teórica y empírica del modelo GHMP, limitado por el objetivo de realizar comparaciones entre diferentes países.

IV. Los datos

IV.1. La Encuesta sobre innovación tecnológica en las empresas del INE

La aplicación empírica de los modelos estructurales exige disponer de definiciones estandarizadas y de datos a nivel de empresa que brinden información detallada de sus procesos de innovación y su impacto sobre el desempeño económico.

Durante la década de los noventa, bajo la línea del Manual de Oslo (OECD, 1992), las agencias estadísticas de los diferentes países comenzaron a conducir encuestas en las que se pregunta directamente a las empresas acerca de sus innovaciones. En particular, ha sido en la Unión Europea donde se han llevado a cabo sucesivas Encuestas

Comunitarias armonizadas de Innovación (CIS), que supervisadas por su Oficina de Estadísticas brindan información detallada sobre los determinantes de los gastos de innovación, los productos de estas actividades, así como del impacto de la innovación en el mercado.

En España, el Instituto Nacional de Estadística (INE), utilizando el marco poblacional del Directorio Central de Empresas, realiza la “Encuesta sobre innovación tecnológica en las empresas” siguiendo las recomendaciones del Manual de Oslo de la OCDE (1992). Dicha encuesta, cuya continuidad queda reflejada en las sucesivas realizaciones desde 1996 a 2005, está dirigida a todas las empresas industriales, de construcción y de servicios con al menos diez personas ocupadas remuneradas cuya actividad principal se corresponde con las agrupaciones de ramas de actividad de la CNAE-93 Revisión 1. En este se pregunta directamente a las empresas acerca de características de sus procesos de innovación, para obtener información relacionada con los insumos, productos y, las dimensiones organizacionales y de comportamiento de sus actividades innovadoras.

En el caso de la encuesta 2004, las principales dimensiones relevadas refieren a:

1. *Datos generales de la empresa*: actividad económica principal, tamaño, personal empleado, cifra de negocios e inversión bruta en materiales.
2. *Actividades de Investigación y Desarrollo realizadas en 2004*. En dicho apartado se incluye: características del personal dedicado a las actividades de I+D internas, gasto en tales actividades y compra de servicios en I+D.
3. *Actividades y gastos de innovación realizados durante el período 2002-2004*. Aquí se incluye el gasto correspondiente a cada una de estas actividades y las ayudas financieras recibidas por la empresa para innovar.
4. *Innovación de productos y procesos realizados por la empresa durante el período 2002-2004*. Se distingue el tipo de innovación realizada por la empresa, el impacto económico de las mismas, las actividades de innovación en curso o abandonadas, las fuentes de información para innovar, la cooperación para realizar actividades de innovación y los efectos de la innovación sobre la actividad económica de la empresa.
5. *Factores que dificultan las actividades de innovación durante el período 2002-2004*.

6. *Solicitud de derechos de propiedad intelectual e industrial durante el período 2002-2004.*

7. *Introducción de innovaciones organizativas y de comercialización durante el período 2002-2004.*

A pesar de brindar información detallada de los procesos de innovación, dichas encuestas cuentan con una serie de problemas importantes, tanto de procedimiento de muestreo como de composición de cuestionario, que deben ser tenidas en cuenta al estimar el modelo estructural.

Dada la metodología con la que se elabora el cuestionario, parte de la información, se encuentra disponible únicamente para las empresas innovadoras. Luego de existir una serie de preguntas comunes a todas las empresas sobre información general, sólo aquellas empresas que han realizado actividades innovadoras durante el período encuestado deben contestar una serie de preguntas adicionales relacionadas con estas actividades. Esto determina que exista un importante problema de *sesgo de selección*.

En España, a diferencia por ejemplo del Reino Unido, la respuesta al cuestionario es obligatoria, por tanto, eventualmente, no existirían problemas de baja respuesta o no respuesta. Sin embargo, un problema importante a tener en cuenta es la necesaria subjetividad por parte de los empresarios que responden al cuestionario. Dado que las empresas son cuestionadas acerca de productos o procesos que son “tecnológicamente nuevos”, existe un espacio obvio para las diferencias de interpretación. Lo cual podría introducir problemas de *errores de medida* en el nivel de innovaciones sesgando, por ejemplo, los resultados de la relación entre crecimiento de la productividad e innovaciones.

Por problemas de confidencialidad no se proporcionan las variables que corresponden a datos de identificación de la empresa como: número de empleados, cifra de negocios, volumen de exportaciones, inversión bruta en bienes materiales, gasto en innovación y el personal en actividades de I+D. En su lugar se proporcionan medias muestrales. Para cada variable los valores se colocan en orden ascendente y, los registros se agrupan de tres en tres, reemplazando cada valor por la media de estos tres. Dicha metodología de reporte también puede conducir a importantes *errores de medida*.

Otra limitación importante del cuestionario es que no recaba información adecuada tanto de desempeño económico de las empresas, como de utilización de factores

productivos. Una alternativa para superar este problema consistiría en fusionar los datos de esta encuesta con otras.

Sin embargo, la mayor debilidad de los datos provenientes de esta encuesta es su naturaleza de sección cruzada, no se pueden observar a las mismas empresas repetidas en el tiempo. Esto implica ciertas precauciones al interpretar los resultados de la estimación de las diferentes ecuaciones del modelo estructural; todo lo que se puede recobrar son sus correlaciones pero no necesariamente relaciones causales. Se debe tener en cuenta que el impacto de la realización de actividades de I+D en la generación de innovaciones, y de estas sobre el desempeño de la empresa, exige cierto lapso temporal que los datos de sección cruzada no pueden capturar.

Al permitir el análisis de modelos dinámicos, la construcción de paneles de datos fruto de las sucesivas encuestas permitirían superar, al menos en parte, este tipo de problemas. Sin embargo, en el caso de España, la metodología con la que se elabora el cuestionario implica que exista un problema importante de solapamiento temporal de gran parte de las variables analizadas.

Por ejemplo, en la encuesta del año 2003 las variables relacionadas con: actividades y gastos de innovación, innovaciones de productos y procesos, factores que dificultan las actividades de innovación y solicitud de derechos de propiedad intelectual e industrial, refieren al período 2001-2003. Por tanto, al eventualmente fusionarla con la encuesta correspondiente al año 2004 existiría un problema importante de solapamiento temporal al referirse esta última al período 2002-2004.

IV.2. Características de la muestra en el período 2002-2004

En el Anexo II se reportan estadísticos descriptivos básicos de las variables objeto de análisis.

Composición sectorial de la base de datos

Conjuntamente, las ramas de productos químicos (13.9%), componentes eléctricos (12.4%), maquinaria (12.1%), metales básicos (11.7%) y alimentos (11.9%) representan más de la mitad de las empresas del sector manufacturero.

Por su parte, las empresas integrantes del sector textil y, papel y derivados representan conjuntamente el 13.8%. El sector residual Nec y el de plásticos y derivados son los de

menor frecuencia relativa en el sector con una representación de empresas de 5.3 y 5.4% respectivamente.

Distribución por tamaños

La distribución según tamaño muestra que las empresas de 20 a 49 trabajadores representan la cuarta parte del sector, siendo el tamaño más frecuente. El resto muestra un grado de participación similar: las empresas de 0 a 19 trabajadores alcanzan el 17.1%; las de 50 a 99 trabajadores el 17.5%; las de 100 a 249 trabajadores el 19%, cifra similar a las de 250 a 999 trabajadores que representan el 18.6% del sector.

Son las de 1000 trabajadores o más las de menor participación, alcanzando el 2.3% de las empresas del sector manufacturero.

Los insumos de la innovación: Quién hace I+D?

Considerando el conjunto de la industria manufacturera, aproximadamente el 54% de las empresas reportan realización de actividades de I+D de forma continua durante el período 2002-2004.

Resultados: La introducción de innovaciones de proceso y de producto

Las empresas que declaran haber introducido *innovaciones de proceso* durante el período 2002-2004 representan el 57% del sector manufacturero, mientras que las de *producto* el 58%.

Dentro de las que declaran haber introducido algún tipo de innovación en el período (74% del sector manufacturero), el 56% declara haber introducido conjuntamente ambos tipos de innovaciones. De las empresas que declaran haber introducido *innovaciones de proceso* en el período, el 32.8% no realizó actividades de I+D internas de forma continua; este porcentaje alcanza el 27.7% para las *innovadoras de producto*. Si consideramos al conjunto de las empresas innovadoras, con independencia de si son de producto o proceso, el 33.8% declara no haber realizado este tipo de actividades de forma continua durante el período 2002-2004.

¿Cómo se innova?

La cooperación

Aproximadamente el 31.3 % de las empresas cuentan con algún acuerdo de cooperación para innovar durante el año 2004.

Los mercados

Un alto porcentaje de las empresas (77.25%) han señalado que durante el año 2004, el mercado más significativo ha sido el internacional.

La protección de las innovaciones

Una tercera parte de las empresas protegieron sus innovaciones ya sea de manera formal o estratégica en el período 2002-2004. Aunque, con un porcentaje levemente superior de protección de tipo formal (34.9%) con relación a la de tipo estratégico (32%).

El apoyo público: la financiación pública

Durante el período 2002-2004, cerca de una cuarta parte de las empresas recibió apoyo financiero de las administraciones locales autonómicas, el 22% por parte del gobierno central y sólo el 5.6% de por parte de la unión europea.

Las fuentes de información

Son las fuentes de información internas a la empresa o grupo empresarial las que cuentan con una mayor proporción de empresas que las consideran de importancia elevada para innovar (46%), seguidas de los consumidores (21%) y, en menor medida los proveedores (15%) y los competidores (10%). Siendo las de menor importancia las fuentes universitarias y gubernamentales (6 y 3%, respectivamente).

Las regulaciones y la demanda de innovaciones.

Los requisitos normativos cuentan con una importancia intermedia o elevada para el 40.8% de empresas, mientras que este porcentaje en el caso del impacto medioambiental o mejora en la salud o seguridad asciende al 37.2%.

V. Réplica del modelo GHMP

En el presente apartado se estima el modelo estructural para el período 2002-2004 reproduciendo la metodología del trabajo GHMP.

En la primera sección se detalla el modelo econométrico utilizado. A continuación, se definen las variables utilizadas en la regresión. Finalmente, se presentan y comentan los resultados de la estimación realizando un análisis comparativo con los obtenidos por el trabajo de GHMP para el caso de la industria manufacturera española.

V.1. Modelo Econométrico

El modelo considerado consiste en un sistema ecuacional no-lineal y recursivo sin efecto retroalimentación que se formaliza en el sistema de cuatro ecuaciones explicitado en la sección III.2.

Dado que las variables dependientes de cada ecuación se observan de forma diferente, el tratamiento econométrico que se brinda a cada una de ellas es distinto.

La primera, ecuación 13, describe la decisión de la empresa de dedicar de forma continua un esfuerzo suficiente a actividades de I+D internas, lo cual resulta en una inversión observable en estas actividades (hacer y/o reportar I+D interna o no). La segunda, ecuación 14, describe la intensidad con la cual la empresa realiza estas actividades. La tercera, ecuación 15, es la conocida función de producción de innovaciones, que supone que puede tomar dos formas diferentes: innovaciones de producto y de proceso. Finalmente, la ecuación 16 es la función de producción ampliada de la empresa donde el conocimiento es considerado como un insumo adicional.

En el caso de las ecuaciones relacionadas con las actividades de I+D, se supone que los términos de error (u_{i0}, u_{i1}) siguen una distribución Normal Bivariada N_2 , con media cero y varianzas respectivas $\sigma_{u_{i0}}^2 = 1$ y $\sigma_{u_{i1}}^2$, siendo $\rho_{u_{i0}, u_{i1}}$ coeficiente de correlación de ambos errores.

El sistema de ecuaciones (13) y (14) se estima como un modelo Tobit tipo 2 por el método de máxima verosimilitud robusta a heteroscedasticidad (utilizando STATA y el procedimiento de Heckman para escoger los valores iniciales de los parámetros).²³

La función de innovación se estima a través de dos ecuaciones probit discretas separadas para indicadores binarios de innovaciones de producto y de proceso por máxima verosimilitud robusta a heteroscedasticidad.

En el caso del esfuerzo innovador interno de las empresas k_i^* , se toma el valor predicho por el modelo Tobit tipo 2 como variable explicativa y se estiman las funciones de innovación para todas las empresas, y no solo para la sub-muestra de aquellas que

²³ Se utiliza el comando “Heckman” de STATA.

reportan gastos en actividades de I+D internas. Esto intenta reflejar el hecho que todas las empresas realizan algún esfuerzo innovador, pero no todas lo reportan.²⁴

Al utilizar el valor predicho de la ecuación de I+D interna, se instrumenta el esfuerzo innovador k_i^* y se toma en cuenta que posiblemente sea endógeno en la función de producción de innovaciones. Parece probable que características no observables de las empresas (y por tanto omitidas) pudieran incrementar su esfuerzo innovador y, por tanto, su “innovatividad” (productividad al producir innovaciones). Esto significaría que al estimar el parámetro γ_k de la ecuación (15) se pudiera obtener una estimación sesgada hacia arriba (dado que la variable latente k_i^* y el término de error u_{i2} podrían estar positivamente correlacionados). Sin embargo, las ecuaciones de selección y esfuerzo innovador corrigen este sesgo al ser las variables x_{i0} y x_{i1} independientes del término de error u_{i2} .

En el caso de la ecuación de productividad, el producto de la empresa, y_i , se mide a través de la productividad del trabajo (logaritmo del producto por trabajador). El vector x_{i3} incluye únicamente el logaritmo del capital físico por trabajador (aproximado por la inversión en capital físico por trabajador), siendo I_i^* el insumo de conocimiento aproximado por dos indicadores: innovaciones de producto y de proceso. Se toma en cuenta la endogeneidad de la variable I_i^* en esta ecuación utilizando los valores predichos al estimar ambas funciones de innovación.

En resumen, dado que se supone una estructura de modelo recursivo y no se toman en cuenta efectos retroalimentación, se sigue un proceso de estimación de tres etapas. En la primera etapa se estima un modelo Tobit tipo 2 para las ecuaciones relacionadas con la realización de actividades de investigación interna. En una segunda etapa, se estiman separadamente las funciones de producción de conocimientos para las innovaciones de producto y de proceso como dos ecuaciones probit discretas utilizando el valor predicho de esfuerzo innovador de la primera etapa (por tanto, al estimar dichas funciones se toma en cuenta la endogeneidad y selectividad de la variable k_i^*). En la última etapa, se estima la ecuación de productividad a través de una regresión lineal, utilizando los

²⁴ Por ejemplo, se puede pensar que los empleados pasan parte del día reflexionando sobre cómo conseguir mejoras de eficiencia en el proceso productivo en el que están trabajando. Sin embargo, por debajo de cierto umbral la empresa no será capaz de recoger información explícita sobre este esfuerzo y, por tanto, no dará parte de él.

valores estimados de la segunda etapa (para tomar en cuenta la endogeneidad de I_i^* en esta ecuación).

V.2. Definición de variables

A los efectos de implementar el modelo se trabaja con datos de las empresas manufactureras españolas provenientes de la “Encuesta sobre innovación tecnológica en las empresas 2004”. En el Anexo I se presenta la definición de las variables construidas para el presente análisis.

En el caso de la decisión de realizar actividades de I+D internas de forma continua durante el período 2002-2004, reflejada por una variable binaria, se utilizan como variables explicativas un indicador binario que señala si el mercado internacional ha sido el más importante para la empresa durante este período, de forma de capturar su exposición a la competencia internacional. Tres indicadores que consideran si la empresa ha recibido financiamiento público (apoyo financiero local o autonómico, del gobierno central o de la unión europea) para llevar adelante actividades innovadoras durante el período 2002-2004. Adicionalmente, se consideran dos indicadores relacionados con las condiciones de apropiabilidad de los resultados de sus innovaciones (protección formal o estratégica).

Para explicar la intensidad de las actividades de investigación, medida como los gastos en actividades de I+D internas por trabajador (en logaritmos), conjuntamente con los tres indicadores de financiamiento público, se consideran cuatro indicadores que reflejan las condiciones de demanda a las que se enfrenta la empresa, uno referente a los arreglos de cooperación de la empresa para realizar actividades de innovación, y, un conjunto de seis variables binarias relacionadas con las diferentes fuentes de información para innovar.

Se distinguen dos tipos diferentes de innovaciones: de producto y de proceso. Cada una de ellas se mide como una variable binaria que señala si la empresa ha introducido al menos un producto o proceso innovador durante el período 2002-2004. Como variables explicativas, además de la intensidad en I+D predicha, se consideran indicadores de las condiciones de demanda de mercado y de apropiabilidad de los resultados de la innovación.

Se espera que las empresas sean más exitosas al obtener innovaciones de producto si utilizan a los consumidores o competidores como fuente de información y, en el caso de las innovaciones de proceso, si utilizan la información proveniente de sus proveedores o competidores, por lo que se utilizan indicadores para cada una de estas fuentes.

En el caso de las innovaciones de proceso, se adiciona como variable explicativa la inversión en capital físico, dado que se quiere testear la complementariedad existente entre este tipo de innovaciones y la inversión en capital que involucran los nuevos procesos tecnológicos. Dicha variable no se incluye en el caso de las innovaciones de producto porque no se encontraron antecedentes de complementariedad en este caso.

Finalmente, la productividad de la empresa, medida como la productividad del trabajo (logaritmo del salario por trabajador), depende del conocimiento medido en términos de innovaciones de producto y de proceso. Y, dado que en la encuesta no es posible observar directamente el capital físico, se lo aproxima a través de una medida continua de inversión en capital físico.

En todas las ecuaciones se controla por características inobservables de industria y,²⁵ por tamaño de empresa. En el caso de la ecuación de intensidad en I+D no se considera el tamaño pues ya está implícitamente escalada por el tamaño de la empresa.

V.3. Resultados de la Estimación

Antes de interpretar los resultados de la estimación es importante señalar que el estudio cuenta con una debilidad importante: sólo se tienen datos de sección cruzada correspondientes al período 2002-2004.

Por tanto, se necesita tener bastante precaución al interpretar los resultados de las estimaciones, todo lo que se obtienen son correlaciones que no necesariamente son relaciones causales

V.3.1. La realización de actividades internas de I+D

Las primeras dos columnas de la tabla 1 brindan las estimaciones por máxima verosimilitud robusta a heteroscedasticidad de un modelo probit discreto de los determinantes que conducen a las empresas manufactureras a realizar actividades internas de I+D de forma continua durante el período 2002-2004. Las segundas dos

²⁵ Se consideran las agrupaciones de actividad de la CNAE-93 Revisión 1. En el Anexo I se detallan las ramas integrantes de cada agrupación.

columnas muestran las estimaciones correspondientes a los determinantes de cuento esfuerzo invierten las empresas en estas actividades condicional a realizarlas.

A diferencia de los modelos lineales, los parámetros estimados de un modelo no lineal no tienen una interpretación directa. Dado que el modelo es una probabilidad, el valor absoluto de los coeficientes da una fotografía distorsionada de la respuesta de la variable dependiente ante un cambio en uno de sus estímulos. Por ello, los coeficientes aquí reportados son los efectos marginales en la media muestral.

En el modelo probit, en el caso que las variables explicativas sean binarias el efecto marginal es el cambio en la probabilidad generado de pasar de no tener una característica dada a tenerla (pasar de 0 a 1); mientras que, si la variable es continua, es el impacto en la probabilidad de un cambio marginal en la media de la variable.

En el caso de la intensidad en investigación interna, es el efecto sobre el valor esperado de este tipo de actividades condicional a un gasto positivo.

Al igual que en el modelo GHMP se consideran las empresas con un tamaño de 20 a 49 trabajadores como categoría de referencia (excluida) para las variables binarias de tamaño y,²⁶ la variable binaria de la rama de alimentos como categoría de referencia para las ramas de actividad.

²⁶ En el estudio GHMP no se toman en cuenta las empresas con un tamaño de 0 a 19 trabajadores dado que en Francia la encuesta CIS no incluye este tamaño de empresa.

Tabla 1
Determinantes de la decisión de llevar a cabo actividades de I+D interna de forma estable e intensidad de dicho esfuerzo.

Efectos marginales

| | Tipo de variable | Hacer o no de forma continua I+D 1998-2000 | Hacer o no de forma continua I+D 2002-2004 | Intensidad de I+D 1998-2000 | Intensidad de I+D 2002-2004 |
|--|------------------|--|--|-----------------------------|-----------------------------|
| Observaciones | | 3588 | 4437 | 750 | 2378 |
| Mercado internacional | Binaria | 0.073*** (0.018) | 0.434*** (0.051) | 0.132* (0.076) | -0.003 (0.148) |
| Cooperación | Binaria | - | - | 0.169* (0.090) | 0.135 (0.097) |
| Protección formal | Binaria | 0.129*** (0.027) | 0.288*** (0.044) | 0.015 (0.085) | 0.214*** (0.090) |
| Protección estratégica | Binaria | 0.200*** (0.029) | 0.331*** (0.045) | -0.149* (0.077) | -0.024 (0.089) |
| Financiamiento local o autonómico | Binaria | 0.081*** (0.023) | 0.289*** (0.051) | -0.078 (0.084) | 0.354*** (0.1028) |
| Financiamiento gobierno central | Binaria | 0.273*** (0.030) | 0.566*** (0.056) | 0.293*** (0.090) | 0.598*** (0.102) |
| Financiamiento UE | Binaria | 0.101** (0.048) | 0.286*** (0.106) | 0.147 (0.139) | 0.272* (0.152) |
| Tamaño 0-19 | Binaria | - | -0.208*** (0.066) | - | - |
| Tamaño 50-99 | Binaria | 0.101*** (0.020) | 0.219*** (0.063) | - | - |
| Tamaño 100-249 | Binaria | 0.237*** (0.025) | 0.247*** (0.061) | - | - |
| Tamaño 250-999 | Binaria | 0.418*** (0.029) | 0.047 (0.063) | - | - |
| Tamaño 1000 o más | Binaria | 0.683*** (0.058) | 0.211 (0.156) | - | - |
| Constante | | - | -0.938*** (0.080) | - | - |
| W_arrastre demanda | | - | - | 0.312 | 0.000 |
| W_fuentes | | - | - | 0.159 | 0.2318 |
| W_industria | | 0.000 | 0.000 | 0.000 | 0.000 |
| Rho | | - | - | 0.745 (0.056) | 0.112 (0.023) |
| W_Rho | | | | | 0.000 |
| Log-likelihood | | - | - | -2,135.7 | -7,760.831 |

Notas: Entre paréntesis se muestran las desviaciones típicas robustas. El término W_Rho brinda el valor de probabilidad del test de significación del término de correlación entre los residuos de ambas ecuaciones. El término W_industria lo hace para las binarias de industria; W_arrastre de la demanda lo hace para estas variables binarias y W_fuentes para las binarias de fuentes de información.

* Significación al 10%, ** Significación al 5%, ***Significación al 1%

Es importante notar que la significación del término de correlación entre los residuos de ambas ecuaciones (Rho) refleja la necesidad de estimar un modelo de selección para la intensidad observada en investigación interna.

Hacer (y/o reportar) I+D interna de forma continua.

Controlando por rama industrial y tamaño de empresa, la probabilidad que una empresa decida realizar actividades de I+D internas de forma continua durante el período 2002-2004 se incrementa significativamente con la superación de restricciones financieras. Este resultado confirma los obtenidos por trabajos previos para el caso de España.²⁷

Sin embargo, la magnitud de dicho impacto difiere entre las fuentes de financiamiento público consideradas. Si la empresa recibe apoyo financiero por parte del gobierno central dicha probabilidad, condicional al valor medio del resto de variables explicativas, se incrementa en un 56.6%. Si los fondos provienen de la unión europea lo hace en 28.6% y en el caso de gobiernos locales o autonómicos en 28.9%.

Si la empresa opera en mercados internacionales la probabilidad de realizar tales actividades, condicional al valor medio del resto de variables explicativas, se incrementa en un 43.4%. Resultado de acuerdo con la hipótesis que señala que el intento por obtener innovaciones es una herramienta estratégica tanto para el desarrollo de las empresas como para su supervivencia en el mercado.

Las medidas de protección de las innovaciones, tanto formal como estratégica, aumentan la probabilidad que la empresa realice actividades internas de investigación en el período. Resultado en línea con los trabajos *Arrow (1962)* y *Spence (1984)* que señalan la importancia de la capacidad de apropiarse de los resultados de la innovación para que las empresas se vean estimuladas a realizar actividades internas relacionadas con la innovación.

Son las medidas de protección estratégica las que cuentan con un impacto superior sobre dicha probabilidad. Mientras que, condicional al valor medio del resto de variables explicativas, las medidas de protección estratégica la incrementan en un 33.1%, las de protección formal lo hacen en un 28.8% (cerca de un 5% menos).

En concordancia con las hipótesis Schumpeterianas, al incrementarse el tamaño de las empresas es más probable que decidan realizar actividades de I+D internas de forma continua. Aunque, en el caso del período 2002-2004, no es posible observar dicha relación para los tamaños mayores de empresa.

²⁷ Por ejemplo, *Huergo (2002)* encuentra que en el año 2000 el esfuerzo inversor de las empresas españolas, al margen de su tamaño, fue superior en las que lograron financiamiento público con relación a las que lo solicitaron sin éxito y, mayor en estas últimas que en las empresas que ni lo buscaron.

Intensidad del esfuerzo en I+D interno

Una vez que la empresa ha tomado la decisión de realizar internamente actividades de I+D de forma continua en el período, la intensidad con la cual las realiza se incrementa de manera significativa con el apoyo financiero que reciba por parte del gobierno central y de las administraciones locales y autonómicas y, en menor medida, con el apoyo que reciba por parte de la unión europea.

Las medidas de protección formal impactan de forma positiva y significativa en la intensidad con la cual la empresa realice estas actividades. Sin embargo, una vez que la empresa ha decidido realizar actividades de I+D, las medidas de protección estratégica no impactan en el monto de recursos destinado a estas actividades.

No se observa una relación significativa y creciente de la intensidad del esfuerzo en investigación y el tamaño de empresa, lo cual corrobora el segundo hecho estilizado de *Cohen y Klepper (1996)*. La elasticidad del capital de I+D interno respecto al tamaño empresarial es uno.

Los resultados obtenidos en la presente estimación de las ecuaciones de I+D son cualitativamente muy similares a los obtenidos por GHMP para el caso de España. Teniendo en cuenta que en este estudio los datos refieren al período 1998- 2000, y en nuestro caso a 2002-2004, se puede concluir que se observa una estabilidad notable en cuanto a la influencia de determinadas variables en las decisiones de llevar a cabo I+D interna de forma continua y en la intensidad de dicho esfuerzo.

V.3.2. La producción de innovaciones

La tabla 2 presenta las estimaciones de las funciones de producción de innovaciones. Las primeras dos columnas reportan los valores estimados de un modelo probit discreto para las *innovaciones de proceso*, mientras que las dos columnas restantes hacia la derecha lo hacen para las *innovaciones de producto*. Los valores reportados son nuevamente los efectos marginales evaluados en las medias muestrales.

Al igual que en el caso de las ecuaciones de investigación se consideran las empresas con un tamaño de 20 a 49 trabajadores como categoría de referencia (excluida) para las variables binarias de tamaño y, la variable binaria de la rama de alimentos como categoría de referencia para las ramas de actividad.

Tabla 2
Funciones de Producción de Innovaciones de Producto y Proceso.
Efectos marginales

| | Tipo de variable | Innovaciones de proceso 1998-2000 | Innovaciones de proceso 2002-2004 | Innovaciones de producto 1998-2000 | Innovaciones de producto 2002-2004 |
|--|------------------|-----------------------------------|-----------------------------------|------------------------------------|------------------------------------|
| Observaciones | | 3588 | 4325 | 3588 | 4325 |
| Intensidad I+D | Continua | 0.281*** (0.025) | 0.046*** (0.005) | 0.296*** (0.026) | 0.063*** (0.005) |
| Intensidad Inversión | Continua | 0.029*** (0.012) | 0.015*** (0.002) | - | - |
| Protección formal | Binaria | -0.031 (0.031) | -0.017 (0.020) | 0.077** (0.034) | 0.090*** (0.019) |
| Protección estratégica | Binaria | 0.068** (0.034) | 0.231*** (0.017) | 0.059* (0.034) | 0.140*** (0.018) |
| Fuente Proveedores | Binaria | 0.405*** (0.028) | 0.224*** (0.020) | - | - |
| Fuente Competidores | Binaria | 0.187*** (0.046) | 0.060** (0.028) | 0.089* (0.048) | 0.078** (0.031) |
| Fuente Consumidores | Binaria | - | - | 0.381*** (0.030) | 0.208*** (0.019) |
| Aspectos medioambientales reducidos | Binaria | 0.916*** (0.254) | 0.214 (0.568) | 0.032 (0.248) | 1.672*** (0.554) |
| Aspectos medioambientales elevados | Binaria | -0.003 (0.234) | 0.143 (0.418) | -0.198 (0.226) | 1.165*** (0.411) |
| Estándares reducidos | Binaria | -0.474* (0.263) | 0.022 (0.547) | 0.306 (0.254) | -0.855 (0.537) |
| Estándares elevados | Binaria | -0.258 (0.231) | 0.737* (0.389) | 0.622*** (0.222) | -0.282 (0.378) |
| Tamaño 0-19 | Binaria | - | 0.024 (0.025) | - | 0.044* (0.025) |
| Tamaño 50-99 | Binaria | 0.015 (0.023) | -0.008 (0.025) | -0.044* (0.022) | -0.039 (0.026) |
| Tamaño 100-249 | Binaria | -0.001 (0.026) | -0.035 (0.025) | 0.070*** (0.026) | -0.063** (0.026) |
| Tamaño 250-999 | Binaria | 0.101*** (0.031) | 0.007 (0.024) | 0.097*** (0.031) | -0.030 (0.024) |
| Tamaño 1000 o más | Binaria | 0.255*** (0.081) | 0.01 (0.057) | 0.259*** (0.083) | 0.005 (0.057) |
| W_industria | | 0.000 | 0.000 | 0.000 | 0.000 |
| Pseudo- R2 | | 0.225 | 0.152 | 0.249 | 0.1869 |
| Log-likelihood | | -1,796.0 | -2,492.5289 | -1,719.1 | 2,372.282 |

Notas: El término W-industria brinda el valor de probabilidad del test de significación conjunta de las variables binarias de industria. * Significación al 10%, ** Significación al 5%, ***Significación al 1%.

Innovaciones de proceso

La probabilidad que una empresa manufacturera obtenga una innovación de proceso en el período 2002-2004, evaluado en el valor medio del resto de las variables, se incrementa en un 4.6% con la realización de actividades internas de I+D de forma continua durante este período.

Únicamente las medidas de protección estratégica que realice la empresa, y no las de protección formal, incrementan la probabilidad de obtener innovaciones de proceso, y lo hacen, condicional al valor medio del resto de las variables explicativas, en un 23.1%.

Los proveedores son una fuente de información significativa y positiva para la obtención de estas innovaciones (incrementan dicha probabilidad en un 22.4%), mientras que los consumidores como fuente de información, aunque son significativos y tienen un efecto positivo, su impacto es marcadamente menor (apenas 6%).

Al igual que en el caso del estudio de GHMP se verifica la complementariedad existente entre la obtención de estas innovaciones y la intensidad de la inversión en capital. Lo cual es coherente con la idea que parte de estas innovaciones se producen por compra de maquinaria.

Sin embargo, a diferencia del estudio GHMP, no es posible observar una relación significativa positiva entre tamaño de empresa y obtención de innovaciones de proceso.

Innovaciones de producto

La probabilidad de obtener de innovaciones de producto en el período 2002-2004, condicional la valor medio del resto de la variables explicativas, se incrementa en un 6.2% con la realización de las actividades de I+D internas, un impacto cerca de 2% superior que en el caso de las innovaciones de proceso.

Ambas medidas de protección de las innovaciones incrementan la probabilidad de obtener este tipo de innovaciones. Mientras que las medidas de protección estratégica, condicional al valor medio del resto de las variables, lo hacen en un aproximadamente un 14%, las medidas de protección formal en un 9%.

La probabilidad de obtener innovaciones de producto depende de las fuentes de información provenientes de los consumidores (incremento de la probabilidad en un 20.8%), y de la competencia (incremento de un 7.8%).

Los valores obtenidos en el caso de las variables relacionadas con el arrastre de la demanda con coeficientes significativos y con signo contrario al esperado, y en algunos casos mayores a uno, podrían sugerir un problema de mala especificación en este sentido.

No se observa una relación significativa entre tamaño de empresa y probabilidad de obtener innovaciones de producto.

En resumen, al igual que en el caso de las ecuaciones de investigación, las estimaciones obtenidas para ambas funciones de innovación son consecuentes con las obtenidas por el estudio de GHMP para el caso de España. Lo que nuevamente vendría a comprobar la estabilidad temporal de las relaciones estimadas.

V.3.3. La Función de Producción

En la tabla 3 se reportan las estimaciones de la ecuación de productividad. Dicha regresión lineal se estima por MCO. Al ser la variable dependiente el logaritmo de las ventas por trabajador, los coeficientes aquí reportados son: una elasticidad en el caso de la variable de inversión y semielasticidades en el resto de las variables.

Tabla 3
Función de Productividad Media

| | Productividad 1998-2000 | Productividad 2002-2004 |
|-------------------------------|----------------------------|----------------------------|
| Observaciones | 3588 | 4325 |
| Intensidad Inversión | 0.061*** (0.006) | 0.038*** (0.004) |
| Innovación de proceso | -0.038 (0.043) | 0.047 (0.035) |
| Innovación de producto | 0.176*** (0.034) | 0.062* (0.108) |
| Tamaño 0-19 | - | -0.271*** (0.041) |
| Tamaño 50-99 | 0.108** (0.045) | 0.171*** (0.035) |
| Tamaño 100-249 | 0.152*** (0.056) | 0.243*** (0.034) |
| Tamaño 250-999 | 0.350*** (0.061) | 0.376*** (0.036) |
| Tamaño 1000 o más | 0.510*** (0.109) | 0.544*** (0.085) |
| W_industria | | 0.000 |
| Constante | 3.692*** (0.078) | 11.700*** (0.063) |
| R² | 0.18 | 0.194 |

El término W-industria brinda el valor de probabilidad del test de significación conjunta de las variables binarias * Significación al 10%, ** Significación al 5%, ***Significación al 1%

Controlando por rama industrial y tamaño de empresa, se encuentra que la elasticidad de las ventas por trabajador depende positivamente de la medida aproximada de inversión en capital físico, y de las innovaciones de producto que obtenga la empresa. Sin embargo, al igual que en el caso de estudio de GHMP, no existe un impacto significativo de las innovaciones de proceso.

Capacidad predictiva del modelo estructural estimado

No existe una única medida estadística simple para evaluar la adecuación del modelo estructural estimado. El paquete estadístico STATA reporta varias, por ejemplo los estadísticos: R^2 o Pseudo- R^2 , o las respuestas predichas.

En el modelo de regresión lineal la bondad de ajuste se puede evaluar comparando los valores estimados con los valores observados. Sin embargo, para los modelos binarios el valor estimado de la variable dependiente es binario pues la variable observada lo es.

Por tanto, el criterio $\sum_i (y_i - \hat{y}_i)$ nos brinda el número de predicciones incorrectas, que

surge si (y_i, \hat{y}_i) es igual a $(1,0)$ o $(0,1)$.²⁸

En términos generales, se observa una buena capacidad predictiva del modelo estructural anteriormente estimado. En el caso de las ecuación de la decisión de realizar I+D el modelo predice de forma correcta el 65.1% de los ceros y el 73.0% de los unos.

Al considerar las innovaciones de proceso el modelo predice de forma correcta el 60.4% de los ceros y el 77.8% de los unos. Mientras que, en el caso de las innovaciones de producto el modelo predice de forma correcta el 55.8% de los ceros y el 77.0% de los unos.

En la ecuación de productividad media la varianza explicada por el modelo es escasa, el coeficiente R^2 apenas alcanza el 19.4%. Este resultado pone en evidencia, la necesidad de mejorar la encuesta sobre innovación tecnológica en lo que respecta a las medidas de

²⁸ Una regla de predicción es fijar $\hat{y} = 1$, cuando $\hat{p} = F(X' \beta) > 0.5$. Sin embargo, esto tiene la debilidad que

si la mayoría de la muestra tiene $y = 1$ entonces $\sum_i (y_i - \hat{y}_i)^2 = n(1 - \bar{y})$ dado que $\hat{p} > 0.5$ y, por tanto,

$\hat{y} = 1$ para todas las observaciones. Surgen problemas similares si la mayoría de la muestra tiene $y = 0$.

desempeño económico y utilización de factores de producción por parte de las empresas, o fusionar dicha encuesta con otras que contengan este tipo de información.

VI. Contrastes de la especificación GHMP y extensiones alternativas

En este apartado se discuten una serie de debilidades teóricas y empíricas que surgen al aplicar el modelo estructural siguiendo la metodología del estudio GHMP.

Básicamente, se trataron tres tipos de problemas. En primer lugar, la posible existencia de variables endógenas en las ecuaciones de investigación. En segundo lugar, el surgimiento de un problema de inconsistencia al estimar las funciones de innovación debido a la aplicación incorrecta de métodos de estimación en etapas a sistemas de ecuaciones no lineales. En tercer lugar, se indaga de forma más profunda en la no significación del indicador de innovaciones de proceso para explicar la variación de la productividad media de las empresas manufactureras españolas.

VI.1. Problemas de endogeneidad en la ecuación de realización de actividades de I+D

En la especificación del modelo estructural planteado por GHMP es posible intuir que ciertas variables explicativas pueden estar correlacionadas con variables no observadas incluidas en el término aleatorio, lo cual generaría problemas de endogeneidad por omisión de variables relevantes y, por tanto, inconsistencia de las estimaciones de los coeficientes asociados a las variables explicativas que se trasladaría al conjunto del sistema de ecuaciones.

En particular, si consideramos la decisión de llevar adelante actividades de I+D internas de forma continua, las tres variables relacionadas con las diferentes fuentes de financiación pudieran estar relacionadas con la “capacidad empresarial” de los directivos (incluida en el término de error) que influiría tanto sobre la realización de I+D como sobre la búsqueda de fuentes de financiación pública.

Siguiendo la metodología de dos etapas de *Rivers y Vuong (1988)* se contrasta la endogeneidad de las tres variables relacionadas con las diferentes fuentes de financiación en la ecuación probit discreta que especifica la probabilidad de llevar adelante actividades de I+D internas de forma continua.

En este caso, el sistema de ecuaciones para la empresa i , especifica una ecuación probit discreta para la decisión de realizar (y/o reportar) I+D interna de forma continua y tres ecuaciones probit para cada una de las fuentes de financiamiento (f_1, f_2, f_3):

$$\left\{ \begin{array}{l} g_i = 1 (g_i^* = x_{i0}b_0 + f_{i1}b_{1f} + f_{i2}b_{2f} + f_{i3}b_{3f} + u_{i0} > 0) \quad (17) \\ f_{i1} = 1[x_i b + v_1] \quad (18) \\ f_{i2} = 1[x_i b + v_2] \quad (19) \\ f_{i3} = 1[x_i b + v_3] \quad (20) \end{array} \right.$$

Para cada una de las diferentes fuentes de financiamiento se especifican el mismo conjunto de variables exógenas expresadas en el vector x .

A su vez, $x_{i0} \in x_i$ y los vectores $(u_{i0}, v_j) \sim N_2[0, \Sigma]$, con coeficientes de correlación $\rho_j = \text{Corr}(u_{i0}, v_j)$ para $j = 1, 2, 3$. Además, los vectores (u_{i0}, v_j) para $j = 1, 2, 3$, son independientes del vector x_i .

La ecuación (17) es la ecuación estructural objeto de interés, y cada una de las ecuaciones probit para las fuentes de financiamiento (ecuaciones 18, 19 y 20) son “ecuaciones reducidas” que proceden de un modelo de decisión de solicitar y obtener financiación pública.

Cada una de las fuentes es endógena en la ecuación estructural si cada coeficiente de correlación $\rho_j \neq 0$ para $j = 1, 2, 3$.

Bajo normalidad conjunta de los términos de error, y suponiendo $V(u_{i0}) = 1$, podemos escribir el término de error de la ecuación estructural como:

$$u_{i0} = \theta_1 v_1 + \theta_2 v_2 + \theta_3 v_3 + e_{i1} \quad (21)$$

El nuevo término de error e_{i1} es independiente de x_i, v_1, v_2 y v_3 . Por tanto, sustituyendo (21) en (17) podemos escribir:

$$g_i = 1 (g_i^* = x_{i0}b_0 + f_{i1}b_{1f} + f_{i2}b_{2f} + f_{i3}b_{3f} + \theta_1 v_1 + \theta_2 v_2 + \theta_3 v_3 + e_{i1} > 0) \quad (22)$$

El procedimiento se realiza en dos etapas:

1. Correr una regresión probit discreta para f_{il} sobre x_i y obtener los residuos estimados \hat{v}_1 . Siguiendo el mismo procedimiento para f_{i2} y f_{i3} , obteniendo estimaciones de los residuos \hat{v}_2 y \hat{v}_3 , respectivamente.
2. Correr una regresión probit discreta de g_i sobre $(x_{i0}, f_{i1}, f_{i2}, f_{i3}, \hat{v}_1, \hat{v}_2, \hat{v}_3)$, para luego testear, a través de una prueba t, la significación conjunta de los coeficientes asociados a las variables \hat{v}_1 , \hat{v}_2 y \hat{v}_3 en esta regresión.

Dicho procedimiento se plantea para las tres fuentes de financiamiento consideradas. En una primera etapa, se estiman tres “ecuaciones reducidas” para cada una de las fuentes de financiación proyectadas sobre el conjunto de variables exógenas. Estas últimas incluyen, además de las variables exógenas de la “ecuación estructural”, una serie de variables adicionales que eventualmente se podrían correlacionar con la solicitud y obtención de fuentes de financiamiento pero no con el término de error en la ecuación de probabilidad de realizar I+D:²⁹

- Una variable relacionada con aspectos organizativos de las empresas: pertenencia o no a un grupo empresarial.
- Una variable que señala si la empresa es de reciente creación. La evidencia empírica señala que las empresas entrantes suelen encontrarse entre las más innovadoras.
- Cuatro variables referidas a la propiedad del capital de la empresa (empresa pública, privada nacional, multinacional o asociación de investigación).
- Tres variables que reflejan la percepción empresarial sobre los obstáculos que limitan su capacidad para realizar innovaciones.³⁰

²⁹ Por razones de tiempo no se ha explorado a fondo la exogeneidad de estas nuevas variables y, por tanto, su calidad como instrumentos. Sin embargo, es importante tener en cuenta que la encuesta no proporciona variables alternativas a las aquí consideradas.

³⁰ En el caso de la construcción de estas variables relacionadas con los obstáculos para innovar se siguió la metodología establecida en el trabajo de Cassiman y Veugelers (1999). La Encuesta sobre innovación tecnológica en las empresas 2004 releva información referida a la importancia de diferentes obstáculos para innovar. Las empresas deben ubicar el obstáculo considerado en una escala de 1 (importancia elevada) a 4 (no relevante). A los efectos de comparabilidad, en el presente trabajo se agregaron las respuestas de ítems relacionadas y se reescalaron en un número continuo comprendido entre 0 y 1.

En una segunda etapa, la “ecuación estructural” de I+D se estima adicionando como variables explicativas las estimaciones de los errores correspondientes a las tres ecuaciones reducidas. Luego, se procede a testear si los coeficientes estimados asociados a dichos residuos en la ecuación estructural son significativamente iguales a cero.

En la tabla 4 se presentan los resultados de la estimación de esta última “ecuación estructural” que incluye los términos de error de las tres ecuaciones reducidas:

Tabla 4
Estimación de ecuación estructural de I+D continua

| | I+D continua |
|--|----------------------|
| Observaciones | 4445 |
| Mercado internacional | 0.144*** (0.023) |
| Protección formal | 0.115*** (0.017) |
| Protección estratégica | 0.131*** (0.017) |
| Financiamiento local o autonómico | 0.107*** (0.019) |
| Financiamiento gobierno central | 0.212*** (0.020) |
| Financiamiento UE | 0.122*** (0.040) |
| Residuos Financiación local | 7.501*** (2.886) |
| Residuos financiación nacional | -2.159 (2.426) |
| Residuos financiación UE | -3.925*** (1.522) |
| Tamaño 0-19 | -0.072*** (0.026) |
| Tamaño 50-99 | 0.082*** (0.024) |
| Tamaño 100-249 | 0.075*** (0.025) |
| Tamaño 250-999 | -0.027 (0.029) |
| Tamaño 1000 o más | -0.053 (0.078) |
| W_industria | 0.000 |
| Pseudo- R2 | 0.161 |
| Log-likelihood | -2574.726 |

El término W-industria brinda el valor de probabilidad del test de significación conjunta de las variables binarias de industria * Significación al 10%, ** Significación al 5%, ***Significación al 1%

La hipótesis nula plantea que los coeficientes asociados a cada uno de los residuos de las tres formas reducidas son iguales a cero en la ecuación estructural:

$$H0: \theta_1 = \theta_2 = \theta_3 = 0$$

El estadístico de dicho test sigue una distribución chi-cuadrado con grados de libertad igual al número de supuestas variables endógenas. En el caso del presente test toma el valor:

$$\text{Chi2 (3)} = 9.03$$

A un 95% de significación nos brinda un valor de probabilidad asociado a dicho coeficiente de 0.0289.³¹ Se rechaza hipótesis nula de exogeneidad de las tres variables de financiación consideradas.

Por tanto, el resultado del test nos señala que las estimaciones del modelo estructural en el que se reproduce la metodología del estudio de GHMP están sesgados y, por ende, dicho sesgo se traslada a todas las ecuaciones del modelo.

VI.2. Inconsistencia de los métodos de estimación en etapas en sistemas ecuacionales no lineales

En la sección anterior se trataron problemas de endogeneidad en las ecuaciones relacionadas con actividades de investigación interna. Sin embargo, aún persisten problemas econométricos importantes en la estimación del modelo estructural que pueden derivar, en este caso, en estimaciones inconsistentes de las funciones de innovación.

Si las diversas variables dependientes del modelo estructural fuesen directamente observables, se podría obtener una estimación consistente de los parámetros estructurales mediante mínimos cuadrados ordinarios en el caso de la primera ecuación de I+D, y mediante un método en etapas en las siguientes. Sin embargo, las características de observabilidad parcial de parte de las variables endógenas convierten este sistema en no lineal. En este caso, la aplicación de un método de estimación en etapas no produciría estimaciones consistentes de las funciones de innovación.³²

En el caso de las ecuaciones de investigación e innovación, la utilización de un método bietápico comprende los siguientes pasos:

³¹ Recordar que el valor de probabilidad, asociado con el estadístico de la prueba de hipótesis, nos da la probabilidad de obtener un valor de dicho estadístico al menos tan grande en valor absoluto como el obtenido realmente, suponiendo que la hipótesis de relevancia (hipótesis alternativa) es correcta. Por tanto, cuanto menor es este valor, mayor es la evidencia contra la hipótesis irrelevancia (hipótesis nula).

³² Ver Wooldridge (2002), capítulo 15 página 478.

1. En una primera instancia las ecuaciones de investigación se estiman a través de un modelo tobit tipo 2. Una ecuación de selección describe si la empresa realiza (y/o reporta) o no actividades de I+D:

$$g_i = 1 (g_i^* = x_{i0}b_0 + u_{i0} > 0)$$

Luego, las empresas deciden el monto a invertir en estas actividades de investigación internas, donde k_i^* se observa de forma condicional a que ya han tomado la decisión de realizar tal inversión:

$$k_i = 1(k_i^* = x_{i1}b_1 + u_{i1} > 0)$$

En este caso, los parámetros b_0, b_1 se estiman de forma consistente obteniendo una estimación del valor esperado de la intensidad de estas actividades de investigación internas:

$$\hat{E}(k_i / x_{i0}, x_{i1}, g_i^* > 0)$$

2. En una segunda etapa se estiman los modelos probit discretos de la variable observada I_i sobre: x_{i2} y la estimación del esperado de la variable endógena

$$\hat{E}(k_i / x_{i0}, x_{i1}, g_i^* > 0)$$

Sin embargo, para que este procedimiento brinde estimaciones consistentes, la función de producción de innovaciones debería ser del tipo:

$$P(I_i = 1 / k_i, x_{i2}) = \Phi[\gamma_k E(k_i / x_{i0}, x_{i1}, g_i^* > 0) + x_{i2}b_2] \quad (23)$$

Pero, dicha función es:

$$P(I_i = 1 / \gamma_k, x_{i2}) = E(I_i / \gamma_k, x_{i2}) = E[1(k_i^* \gamma_k + x_{i2}b_2 + u_{i2} > 0)] \quad (24)$$

Al ser la función indicador $1(\cdot)$ no lineal no se puede pasar dentro el operador esperanzas. Por lo cual, la estimación en dos etapas de las funciones de producción de innovaciones se podría realizar de forma inconsistente.

En el caso de la estimación en etapas de la ecuación de productividad, considerando los valores estimados de las funciones de innovación, no nos enfrentaríamos a problemas de este tipo pues, dicha función es lineal y, por tanto, es posible estimarla de forma consistente utilizando los valores predichos de las funciones de producción de innovaciones.

Ante los problemas de endogeneidad de las fuentes de financiamiento público en las ecuaciones de investigación e inconsistencia al estimar las funciones de innovación intentando reproducir de forma incorrecta un método bietápico, se exploraron especificaciones alternativas al modelo estructural utilizado por GHMP.

Ante el problema de endogeneidad en las funciones de investigación, aunque una posibilidad era ampliar el modelo estructural incorporando una ecuación para cada una de las variables de financiación, en la presente investigación se consideró pertinente plantear las ecuaciones de I+D en función únicamente de variables exógenas.

En segundo lugar, a los efectos de superar los problemas de inconsistencia en la estimación de las funciones de innovación se decidió especificar dos modelos de ecuaciones simultaneas recursivos que incluyan: una ecuación relacionada con la decisión de realizar actividades de I+D internas de forma continua y la función de innovación (producto o proceso).³³

La clave aquí es que la función de máxima verosimilitud se construye considerando la función de densidad condicional conjunta: $f(I_i / k_i, x_{i2})f(k_i / x_{i0}, x_{i1})$. Esto implica que los estimadores que surjan al maximizar dicha función, además de contar con las propiedades deseables de consistencia, serán más eficientes que los obtenidos en el procedimiento de dos etapas pues, estas estimaciones utilizan de forma simultanea la información contenida tanto en $f(I_i / k_i, x_{i2})$ como en $f(k_i / x_{i0}, x_{i1})$. Mientras que, el procedimiento en dos etapas se focaliza exclusivamente en la función de probabilidad $f(I_i / k_i, x_{i2})$.³⁴

³³ Otra alternativa considerada fue la aproximación de *Función de Control*, pues dicha aproximación considera la endogeneidad de la variable de intensidad en I+D en la estimación de las funciones de producción de innovaciones. Sin embargo, esta aproximación exige la imposición de supuestos muy restrictivos en cuanto a la distribución de probabilidad de las variables exógenas y endógenas. En particular, exige una distribución normal homoscedástica de la forma reducida para la ecuación de investigación: $k_i = x_{i1}b_1 + u_{i1}$, con $u_{i1} \sim N(0, \tau_2^2)$. Esto implica que la distribución condicional de la variable observada k_i sea normal homoscedástica, lo cual claramente no se cumple en el caso de la variable censurada de intensidad en I+D como la aquí analizada.

³⁴ Debido a ello Wooldridge (2002, página 476) denomina a la estimación de dos etapas un “procedimiento de información limitada”.

VI.3. La relación entre I+D interna e Innovaciones. Modelos Probit Bivariantes

Dada la eventual existencia de una relación importante entre la decisión de realizar actividades de I+D internas y la generación de innovaciones por parte de las empresas, que hacen que la variable binaria de I+D sea endógena en la ecuación de innovación debido a problemas de simultaneidad, se consideró pertinente ajustar un modelo recursivo de ecuaciones simultáneas para explicar el comportamiento de ambas variables.

Al ser ambas variables dependientes binarias se optó por plantear un modelo específico de ecuaciones simultáneas denominado “probit bivariante recursivo”.³⁵

Con el objeto de implementar el modelo se consideran las ecuaciones (13) y (15) expresadas en términos de variables latentes:

$$g_i^* = x_{i0}b_0 + u_{i0}, \text{ donde } g_i = 1 \text{ si } g_i^* > 0, 0 \text{ en caso contrario}$$

$$I_i^* = g_i^*\gamma_g + x_{i2}b_2 + u_{i2}, \text{ donde } I_i = 1 \text{ si } I_i^* > 0, 0 \text{ en caso contrario}$$

A su vez, se supone que:

$$E(u_{i0} / x_{i0}, x_{i2}) = E(u_{i2} / x_{i0}, x_{i2}) = 0$$

$$Var(u_{i0} / x_{i0}, x_{i2}) = Var(u_{i2} / x_{i0}, x_{i2}) = 1$$

$$Cov(u_{i0}, u_{i2} / x_{i0}, x_{i2}) = \rho_i$$

La endogeneidad de la variable latente g_i^* surge sí y sólo sí los términos de error de ambas ecuaciones se encuentran correlacionados. Al suponer un modelo probit bivariante los términos de error: (u_{i0}, u_{i1}) siguen una distribución normal bivariante N_2 , con vector de medias nulo, varianzas unitarias y coeficiente de correlación ρ_i .

En estos casos, el término de error u_{i1} se escribe como una proyección sobre u_{i0} : $u_{i1} = \rho u_{i0} + \xi_i$. Esta última expresión se interpreta en el sentido que el efecto no observable u_{i0} ejerce una efecto sobre ambas variables latentes, aunque de diferente magnitud.

³⁵ Es en el trabajo de Greene (1998) donde originalmente se plantea este tipo de modelo recursivo.

Dicho modelo recursivo se estima por máxima verosimilitud robusta a heteroscedasticidad, considerando la función de distribución conjunta del vector (u_{i0}, u_{i2}) como: $\Phi(0, 0; \rho_i)$.

Por tanto, la función de probabilidad puede expresarse como:

$$P_{0,0} = \Pr ob(g_i = 1, I_i = 1) = N_2(x_{i0}^\top b_0, x_{i2}^\top b_2 + \gamma_{g_i}, \rho_i)$$

$$P_{1,0} = \Pr ob(g_i = 1, I_i = 0) = N_2(x_{i0}^\top b_0, -x_{i2}^\top b_2 - \gamma_{g_i}, -\rho_i)$$

$$P_{0,1} = \Pr ob(g_i = 0, I_i = 1) = N_2(-x_{i0}^\top b_0, x_{i2}^\top b_2, -\rho_i)$$

$$P_{1,1} = \Pr ob(g_i = 0, I_i = 0) = N_2(-x_{i0}^\top b_0, -x_{i2}^\top b_2, \rho_i)$$

En el caso de las innovaciones, el valor esperado de la variable binaria observada condicional a x_{i0} y x_{i2} , es:

$$\begin{aligned} E(I_i / x_{i0}, x_{i2}) &= \Pr ob(g_i = 1)E(I_i / g_i = 1, x_{i0}, x_{i2}) + \Pr ob(g_i = 0)E(I_i / g_i = 0, x_{i0}, x_{i2}) = \\ &= \Pr ob(g_i = 1)\Pr ob(I_i = 1 / g_i = 1, x_{i0}, x_{i2}) + \Pr ob(g_i = 0)\Pr ob(I_i = 1 / g_i = 0, x_{i0}, x_{i2}) = \\ &= \Pr ob(I_i = 1, g_i = 1) + \Pr ob(I_i = 1, g_i = 0) = N_2(x_{i0}^\top b_0, x_{i2}^\top b_2 + \gamma_{g_i}, \rho_i) + N_2(-x_{i0}^\top b_0, x_{i2}^\top b_2, -\rho_i) \end{aligned} \quad (25)$$

En las tablas 5 y 6 se presentan los coeficientes estimados de ambos modelos. En los dos casos, las variables binarias excluidas (variables de referencia) son: empresa privada nacional, empresa de 20 a 49 trabajadores y rama de alimentos.

Tabla 5
Modelo probit bivariante: I+D continua e innovaciones de producto
Coeficientes estimados

| | Hacer o no de forma continua I+D 2002-2004 | Innovaciones de producto 2002-2004 |
|---|--|---------------------------------------|
| Observaciones | 4445 | |
| I+D | - | 1.863*** (0.053) |
| Mercado internacional | 0.314*** (0.050) | - |
| Protección formal | 0.167*** (0.045) | 0.279*** (0.049) |
| Protección estratégica | 0.137*** (0.046) | 0.334*** (0.047) |
| Empresa pública | 0.090 (0.238) | - |
| Asociación de investigación | -0.356 (0.562) | - |
| Empresa multinacional | 0.002 (0.059) | - |
| Grupo empresarial | 0.035 (0.050) | - |
| Empresa de nueva creación | 0.272* (0.163) | - |
| Obstáculos de costes | 0.159*** (0.074) | - |
| Obstáculos de información | 0.327*** (0.087) | - |
| Obstáculos Nec | -0.903*** (0.074) | - |
| Fuente interna | 0.600*** (0.040) | |
| Fuente universitaria | 0.420*** (0.058) | - |
| Fuente pública | 0.148** (0.063) | - |
| Fuente competencia | -0.087 (0.081) | 0.250*** (0.082) |
| Fuente consumidores | 0.223*** (0.060) | 0.299*** (0.061) |
| Aspectos medioambientales bajos | 3.134** (1.397) | 1.534 (1.404) |
| Aspectos medioambientales elevados | 0.980 (1.104) | 1.577 (0.100) |
| Estándares bajos | -0.804 (1.342) | -1.011 (1.342) |
| Estándares elevados | 0.086 (1.012) | -0.389 (1.020) |
| Tamaño 0-19 | -0.238*** (0.068) | 0.108* (0.064) |
| Tamaño 50-99 | 0.148** (0.065) | -0.087 (0.062) |
| Tamaño 100-249 | 0.191*** (0.066) | -0.160*** (0.061) |
| Tamaño 250-999 | 0.212 (0.071) | -0.061 (0.061) |
| Tamaño 1000 o más | 0.141 (0.151) | 0.024 (0.149) |
| Constante | -1.692*** (0.400) | -1.617*** (0.383) |
| W_industria | 0.000 | 0.000 |
| Rho | | -0.726 |
| W_Rho | | 0.000 |

El término W_Rho brinda el valor de probabilidad del test de significación del término de correlación entre los residuos de ambas ecuaciones, W-industria lo hace para las variables binarias de industria.

* Significación al 10%, ** Significación al 5%, ***Significación al 1%

La significación del término de correlación entre los residuos de ambas ecuaciones (Rho) refleja la adecuación al estimar un modelo simultáneo para explicar su comportamiento.

Hacer (y/o reportar) I+D interna

Los resultados en cuanto a la significación de ciertas variables son muy similares a los obtenidos para el caso del estudio GHMP.

Se observan impactos positivos de las variables relacionadas con la competencia en mercados internacionales, medidas de protección y fuentes de información.

A su vez, es posible observar un impacto significativo de variables no consideradas anteriormente. Por ejemplo, si la empresa es de reciente creación se incrementa la probabilidad de realizar estas actividades internas de forma continua, lo cual evidencia un comportamiento inversor de las empresas cuando son jóvenes en el mercado.

Los obstáculos para innovar relacionados con la existencia de innovaciones anteriores y/o falta de demanda para las innovaciones tienen un impacto significativo y negativo sobre la decisión de realizar estas actividades.

Sin embargo, en el caso de los obstáculos relacionados con los costes o la falta de información relevante para innovar, aunque cuentan con un impacto significativo, su signo es contrario al esperado. Este resultado se ha obtenido también en investigaciones realizadas con datos de la encuesta CIS para otros países.³⁶ En el caso de los costes, existen dos interpretaciones a dicho resultado: 1) las empresas que consideran que los costes son muy importantes son las que han valorado más seriamente la posibilidad de hacer I+D y acaban haciendo I+D; 2) la presencia de causalidad inversa: hacer I+D comporta darse cuenta de la importancia de los costes.

Innovaciones de producto

Los resultados son muy similares a los obtenidos al reproducir el estudio GHMP. La realización de actividades de I+D de forma continua demuestra ser una variable significativa y con impacto positivo para explicar la probabilidad de obtener este tipo de innovaciones.

Por su parte, tanto las medidas de protección formal como estratégica impactan de forma significativa y positiva sobre la probabilidad de obtener este tipo de

³⁶ Ver Cassiman y Veugelers (1999).

innovaciones. Al igual que lo hacen las fuentes de información provenientes de consumidores y competidores.

Por último, no es posible detectar una relación significativa y positiva entre tamaño de empresa y probabilidad de obtener innovaciones de producto.

Tabla 6
Modelo probit bivariante: I+D continua e innovaciones de proceso
Coefficientes estimados

| | Hacer o no de forma continua I+D 2002-2004 | Innovaciones de proceso 2002-2004 |
|---|--|--------------------------------------|
| Observaciones | 4445 | 4445 |
| I+D | - | 1.526*** (0.053) |
| Inversión | 0.036*** (0.006) | 0.014** (0.006) |
| Mercado internacional | 0.278*** (0.049) | - |
| Protección formal | 0.155*** (0.046) | -0.019 (0.045) |
| Protección estratégica | 0.134*** (0.047) | 0.511*** (0.046) |
| Empresa pública | 0.063 (0.237) | - |
| Asociación de investigación | -0.319 (0.547) | - |
| Empresa multinacional | -0.041 (0.060) | - |
| Grupo empresarial | 0.040 (0.049) | - |
| Empresa de nueva creación | -0.146 (0.168) | - |
| Obstáculos de costes | 0.167** (0.073) | - |
| Obstáculos de información | 0.283*** (0.087) | - |
| Obstáculos Nec | -0.848*** (0.074) | - |
| Fuente interna | 0.642*** (0.040) | |
| Fuente universitaria | 0.411*** (0.057) | - |
| Fuente pública | 0.188*** (0.061) | - |
| Fuente proveedores | -0.001 (0.065) | 0.491*** (0.065) |
| Fuente competencia | -0.024 (0.077) | 0.098 (0.071) |
| Aspectos medioambientales bajos | 3.920*** (1.390) | -1.751 (1.295) |
| Aspectos medioambientales elevados | 1.150 (1.062) | -0.793 (0.983) |
| Estándares bajos | -1.162 (1.336) | 1.051 (0.128) |
| Estándares elevados | -0.058 (0.972) | 2.032** (0.921) |
| Tamaño 0-19 | -0.206** (0.068) | 0.073 (0.062) |
| Tamaño 50-99 | 0.142** (0.065) | -0.018 (0.060) |
| Tamaño 100-249 | 0.182*** (0.066) | -0.081 (0.059) |
| Tamaño 250-999 | -0.033 (0.070) | 0.056 (0.059) |
| Tamaño 1000 o más | 0.002 (0.147) | 0.068 (0.138) |
| Constante | -1.979*** (0.413) | -1.304*** (0.390) |
| W_industria | 0.000 | 0.000 |
| Rho | | -0.738 |
| W_Rho | | 0.000 |

El término W_Rho brinda el valor de probabilidad del test de significación del término de correlación entre los residuos de ambas ecuaciones, W-industria lo hace para las variables binarias de industria.

* Significación al 10%, ** Significación al 5%, ***Significación al 1%

Nuevamente, la significación del término de correlación entre los residuos de ambas ecuaciones (Rho) refleja la adecuación al estimar un modelo simultáneo para explicar su comportamiento.

Hacer (y/o reportar) I+D interna

En el caso de la ecuación de investigación los resultados van en la línea de los obtenidos al reproducir el estudio GHMP.

A su vez, se observa la significación con un impacto negativo de la variable relacionada con los obstáculos para innovar derivados de la existencia de innovaciones anteriores y/o falta de demanda para las innovaciones. Al igual que en el caso de las innovaciones de producto, se observa un impacto significativo pero con signo contrario al esperado de las otras dos variables de obstáculos.

Innovaciones de proceso

La realización de actividades de I+D internas de forma continua demuestra ser una variable altamente significativa y con un impacto positivo para explicar la probabilidad de obtener innovaciones de proceso en el período 2002-2004.

También se evidencia una relación significativa de complementariedad entre este tipo de innovaciones y la inversión en capital físico.

Únicamente las medidas de protección formal impactan de forma significativa (y positiva) sobre la obtención de este tipo de innovaciones. Solamente los proveedores son una fuente significativa para obtener innovaciones de proceso.

Finalmente, al igual que en el caso de las innovaciones de producto, no es posible detectar una relación significativa y positiva entre tamaño de empresa y probabilidad de obtener innovaciones de proceso.

Las estimaciones obtenidas al utilizar modelos bivariantes recursivos son muy similares a los que se arribaron en la sección V.3 al reproducir la metodología del estudio GHMP para datos del período 2002-2004.

De todas formas, es importante tener en cuenta que al aplicar el modelo bivariante eventualmente se soluciona el problema de inconsistencia al estimar las funciones de innovación. Por otra parte, es importante destacar que la estimación de un modelo bivariante de este tipo cuenta con la debilidad de no considerar la intensidad con la cual

las empresas realizan actividades de investigación internas y como ello repercute en la generación de innovaciones.

Como ya fue señalado en reiteradas ocasiones a lo largo de la presente investigación, a diferencia de los modelos lineales, los parámetros estimados de un modelo no lineal no tienen una interpretación directa. Para averiguar tales efectos es necesario realizar transformaciones adicionales que permitan calcular efectos marginales.

En el Anexo III se detallan los cálculos necesarios para obtener los efectos marginales del presente modelo probit recursivo. Sin embargo, se requiere la utilización de un programa propio para calcular tales efectos pues, el programa estadístico STATA si bien estima el modelo adecuadamente no permite un cálculo directo de los mismos. En la presente investigación no se generó el *do.file* para calcular tales efectos por lo que queda pendiente como futura línea de investigación.

VI.4. El impacto diferenciado de las innovaciones sobre la productividad

Otro punto a tener en cuenta en la estimación del modelo estructural reproduciendo la metodología de GHMP refiere al impacto de las innovaciones sobre la productividad de las empresas.

Sorprende el resultado, obtenido tanto en nuestro trabajo como en la investigación de GHMP, que únicamente sean las innovaciones de producto las que impactan sobre la productividad media de las empresas manufactureras españolas, no verificándose este resultado para las innovaciones de proceso. En el caso de aquel estudio, los autores argumentan que la no significatividad de las innovaciones de proceso podría corresponder al hecho que se está midiendo la productividad de las ventas de las empresas. Sin embargo, en la presente investigación se consideraron otras alternativas para intentar explicar tales resultados.

En la sección IV.2 se señaló que cerca de un 56% de las empresas del sector manufacturero que innovaron durante el período 2002-2004 obtuvieron conjuntamente tanto innovaciones de proceso como de producto. Considerando, a su vez, el aporte de *Huergo y Moreno (2004)*, en la presente investigación se optó por estimar la ecuación de productividad media a través de dos ecuaciones. En la primera de ellas se analiza únicamente el impacto de las innovaciones de producto mientras que en la otra se analizan únicamente las innovaciones de proceso.

En la siguiente tabla se presentan los resultados de ambas estimaciones:

Tabla 7
Función de Productividad Media
Coeficientes estimados

| | Productividad 2002-2004 | Productividad 2002-2004 |
|-------------------------------|----------------------------|----------------------------|
| Observaciones | 4445 | 4445 |
| Intensidad Inversión | 0.038*** (0.004) | 0.039*** (0.004) |
| Innovación de proceso | 0.058*** (0.014) | - |
| Innovación de producto | - | 0.045*** (0.012) |
| Tamaño 0-19 | -0.284*** (0.041) | -0.284*** (0.041) |
| Tamaño 50-99 | 0.178*** (0.178) | 0.181*** (0.034) |
| Tamaño 100-249 | 0.251*** (0.033) | 0.254*** (0.034) |
| Tamaño 250-999 | 0.375*** (0.035) | 0.382*** (0.035) |
| Tamaño 1000 o más | 0.571*** (0.084) | 0.578*** (0.085) |
| W_industria | 0.000 | 0.000 |
| Constante | 11.708*** (0.054) | 11.705*** (0.054) |
| R² | 0.1876 | 0.1872 |

El término W-industria brinda el valor de probabilidad del test de significación conjunta de las variables binarias de industria.* Significación al 10%, ** Significación al 5%, ***Significación al 1%

A diferencia de la estimación del modelo siguiendo la metodología de GHMP, al estimar por separado el impacto de los indicadores de innovaciones de producto y de proceso se observa en ambos casos un impacto significativo y positivo sobre la productividad media de las empresas. También se observa un impacto significativo y positivo de la medida aproximada de inversión en capital físico.

Sin embargo, la varianza explicada en ambas ecuaciones continúa siendo bastante reducida, lo que nuevamente brinda indicios de la debilidad de las medidas de desempeño económico y de utilización de factores productivos tradicionales provenientes de la encuesta utilizada.

VII. Nuevas extensiones del modelo estructural: la decisión de realizar actividades de I+D

En esta sección se indaga de forma más profunda en la decisión de la empresa en cuanto a la realización de actividades de investigación internas. De hecho, las empresas pueden

decidir no realizar tales actividades, realizarlas de forma ocasional o hacerlas de forma estable durante el período.

El estudio de los factores que explican el comportamiento de las empresas con relación a la elección de estas tres categorías, exhaustivas y mutuamente excluyentes, pueden permitir describir de forma más adecuada el comportamiento de las empresas con relación a estas actividades.

Con este objeto, se construyó una variable categórica relacionada con la realización de actividades de investigación internas a la empresa durante el período 2002-2004 que toma tres valores: no realización de actividades de investigación, realización en forma ocasional y realización en forma continua.

Se formularon dos tipos de modelos para esta variable categórica, cada uno de los cuales refleja distintos supuestos sobre el comportamiento empresarial. Uno es un modelo logit multinomial, y el otro un probit ordenado.

VII.1. Un modelo multinomial de decisión de I+D

En una primera instancia, se optó por estimar un modelo logit multinomial para explicar la decisión de realizar actividades de I+D internas. Al aplicar un modelo de este tipo la teoría implícita, basada en un modelo de utilidad aleatoria, es que la empresa valora las tres alternativas, para luego elegir una de ellas en función de cuál es la que le da mayor beneficio.

Sin embargo, los resultados obtenidos sugirieron que la aplicación de un modelo de este tipo no resultaba muy satisfactoria para explicar el comportamiento de la empresa. El modelo no explicaba la probabilidad de realizar actividades de investigación internas de forma ocasional. Por tanto, imponer esta teoría de utilidad aleatoria parecería no ser adecuada para explicar el comportamiento de la empresa con relación a las actividades de I+D internas.

VII.2. Un modelo ordenado de decisión de I+D

En una segunda instancia, se optó por considerar modelos adicionales construyendo una estructura del proceso de toma de decisiones, como por ejemplo un ordenamiento natural de las alternativas. Puede ocurrir que los tres estados de la variable “hacer I+D interna” reflejen la intensidad de una variable no observada continua, que justificaría la utilización de un modelo ordenado.

La idea básica de este tipo de modelos es suponer la existencia de una variable latente continua, relacionada con un índice simple de variables explicativas y un término de error, y obtener una variable categórica efectivamente observada al discretizar la línea real en un número finito de intervalos.

A diferencia de los modelos no ordenados, a través de un modelo ordenado se obtiene un modelo más parsimonioso y sensible que toma en cuenta el orden de las alternativas, por lo que su análisis puede ser más informativo desde el punto de vista económico.

Formalmente, se especificó un modelo para la variable categórica ordenada de actividades de I+D internas a la empresa codificada un ranking ordenado de tres categorías $g_i \in \{0,1,2\}$ y, donde el vector x_{i0} representa el conjunto de variables explicativas.

En este modelo estándar, las probabilidades acumuladas de la variable observada se relacionan con un índice simple de variables explicativas:

$$\Pr[g_i \leq j / x_{i0}] = F(l_j - x_{i0}^\top b_{i0}), \quad j = 0, 1, 2 \quad (26)$$

Siendo l_j los parámetros umbral que discretizan la línea real, representada por la variable latente g_i^* , en 3 categorías.

Se considera esta variable latente conjuntamente con el siguiente mecanismo de respuesta:

$$g_i = j \Leftrightarrow l_{j-1} \leq g_i^* = b_0 x_{i0} + u_{i0} < l_j, \quad j = 0, 1, 2$$

Se reemplazó la función F de la ecuación (26) por una función de distribución acumulativa normal estándar, lo cual condujo a la utilización de un modelo probit ordenado. Al igual que en el caso anterior, la estimación se realiza por máxima verosimilitud robusta a heteroscedasticidad.

En el caso de un modelo ordenado el signo de los coeficientes estimados determina si la variable latente cruza o no determinado umbral al variar el regresor. Si el signo es positivo entonces la variable latente se incrementa, cruza los umbrales, al aumentar el valor del regresor. Sin embargo, dado que el interés radica en el cambio de la variable

observada, discreta y ordenada, cuando varían uno o más de los regresores, se optó por analizar los efectos marginales de probabilidad.³⁷

En la tabla 8 se presentan los efectos marginales únicamente de las variables estadísticamente significativas.

En el caso que las variables explicativas sean binarias el efecto marginal es el cambio en la probabilidad de pasar de 0 a 1, mientras que si la variable es continua el efecto marginal es el impacto en la probabilidad de un cambio marginal.

Tabla 8
Modelo Probit ordenado de I+D interna
Efectos marginales

| | Tipo de variable | No realizar I+D 2002-2004 | Realizar I+D de forma ocasional 2002-2004 | Realizar I+D de forma continua 2002-2004 | |
|--|------------------|------------------------------|---|--|--|
| Observaciones | | 4445 | | | |
| Mercado internacional | Binaria | -0.150*** (0.018) | -0.016*** (0.002) | 0.166*** (0.018) | |
| Protección formal | Binaria | -0.090*** (0.013) | -0.017*** (0.003) | 0.107*** (0.017) | |
| Protección estratégica | Binaria | -0.096*** (0.013) | -0.019*** (0.003) | 0.115*** (0.016) | |
| Obstáculos de costes | Binaria | -0.084*** (0.025) | -0.014*** (0.004) | 0.098*** (0.029) | |
| Obstáculos de información | Binaria | -0.111*** (0.030) | -0.019*** (0.005) | 0.130*** (0.035) | |
| Obstáculos Nec | Binaria | 0.376*** (0.025) | 0.064*** (0.006) | -0.440*** (0.030) | |
| Fuente universidades | Binaria | -0.168*** (0.017) | -0.036*** (0.004) | 0.204*** (0.021) | |
| Fuente organismos gubernamentales | Binaria | -0.074*** (0.019) | -0.014*** (0.004) | 0.088*** (0.023) | |
| Tamaño 0-19 | Binaria | 0.059*** (0.021) | 0.008*** (0.002) | -0.067*** (0.024) | |
| Tamaño 50-99 | Binaria | -0.057*** (0.019) | -0.011*** (0.004) | 0.068*** (0.023) | |
| Tamaño 100-249 | Binaria | 0.049** (0.020) | -0.009** (0.004) | 0.058** (0.024) | |
| Prob. Media estimada | | 0.28 | 0.18 | 0.54 | |
| Prob. Muestral observada | | 0.32 | 0.14 | 0.54 | |

* Significación al 10%, ** Significación al 5%, ***Significación al 1%

³⁷ Tener en cuenta que los efectos marginales en este tipo de modelos ordenados estándar tienen dos propiedades restrictivas. En primer lugar, el ratio de probabilidades marginales de dos variables independientes continuas distintas sobre la misma respuesta observada, efectos marginales relativos, son constantes entre individuos. Además, las probabilidades marginales cambian de signo una vez que nos movemos desde la categoría más baja a hacia la más alta. Este último resultado proviene directamente de la forma acampanada de las funciones de densidades normales o logísticas.

Si una empresa está presente en mercados internacionales, manteniendo todo lo demás constante, la probabilidad de que no realice actividades de I+D internas en el período 2002-2004 se reduce en 15.0 puntos porcentuales en relación a una empresa que no esté presente en dichos mercados. También se reduce la probabilidad de que haga I+D ocasional en 1.6 puntos porcentuales, mientras que aumenta la probabilidad de que lo haga de forma continua en 16.6 puntos porcentuales.

En resumen, al estimar el modelo ordenado se observa que la utilización de ambos mecanismos de protección de la innovación son los factores más relevantes para explicar la realización de actividades de I+D internas en el período 2002-2004, seguidos de la importancia que da la empresa a fuentes de información académicas y al carácter exportador de la misma. Sin embargo, la dimensión empresarial tiene una importancia menor.

VIII. Comentarios finales y futuras líneas de investigación

La estimación del modelo estructural básico de función de producción, aplicada a datos provenientes de la Encuesta sobre innovación tecnológica en las empresas 2004, ha demostrado ser una herramienta analítica de suma utilidad para caracterizar el comportamiento innovador de las empresas manufactureras españolas.

Utilizando la metodología del estudio GHMP, fue posible demostrar que en un período posterior al analizado por estos autores se obtienen resultados muy similares en relación a los lazos existentes entre investigación-innovación-productividad a nivel empresarial, lo que brinda indicios de estabilidad temporal en las relaciones estimadas.

Según estos resultados la realización de forma continua de actividades de investigación internas se correlaciona significativa y positivamente con: la obtención de fuentes de financiamiento públicas, la utilización de medidas de protección de los resultados de innovación, la competencia en mercados internacionalizados y el tamaño empresarial.

La intensidad de estas actividades se correlaciona con la obtención de las tres fuentes de financiamiento consideradas y con medidas de protección estratégica aunque, no con mecanismos formales de protección. A su vez, una vez que la empresa ha decidido realizar este tipo de actividades, la intensidad de las mismas no depende de su tamaño. La elasticidad del capital de I+D interno con respecto al tamaño empresarial es uno.

La capacidad de introducir innovaciones de producto por parte de una empresa, además de correlacionarse positivamente con la intensidad con la cual realiza actividades internas de I+D, lo hace con el uso de métodos de protección estratégicos, no con mecanismos formales, con la intensidad de la inversión en capital físico y con la importancia de las fuentes de información provenientes de proveedores y consumidores.

En el caso de las innovaciones de producto, además de correlacionarse positivamente con la intensidad de actividades de investigación internas, lo hace con ambas medidas de protección de las innovaciones y, con la importancia de los consumidores y la competencia como fuentes de información para innovar.

La probabilidad de obtener innovaciones, tanto de producto como de proceso, no se correlaciona de forma significativa con el tamaño de la empresa. No existen evidencias en los datos aquí analizados que las empresas grandes cuenten con ventajas para innovar.

Finalmente, la productividad media de la empresa se correlaciona positivamente con: la introducción de nuevos productos, el tamaño empresarial y la intensidad de capital físico. En esta primera instancia no se encuentra un efecto significativo de las innovaciones de proceso sobre la productividad media de las empresas.

Sin embargo, fue posible demostrar que la metodología utilizada al reproducir el modelo GHMP sufre de una serie de debilidades teóricas y empíricas importantes que, entre otras cosas, pueden dar lugar a estimaciones inconsistentes de los parámetros del modelo.

En primer lugar, las variables relacionadas con las fuentes de financiamiento público son endógenas en la ecuación de decisión de realizar actividades de investigación internas de forma continua, lo cual genera problemas de inconsistencia que se trasladan al resto del modelo. En segundo lugar, se demostró analíticamente que resulta erróneo aplicar un método de estimación en etapas a sistemas de ecuaciones no lineales como el aquí considerado.

La estimación de un modelo probit recursivo para las ecuaciones de investigación e innovación en el que se consideran el conjunto de variables exógenas, no sólo nos permitió obtener estimaciones consistentes y más eficientes de los parámetros del modelo, sino que, a su vez, nos permitió evaluar el impacto significativo en ambas ecuaciones de nuevas variables. Por ejemplo la importancia de: las fuentes de

información provenientes de ámbitos académicos y organismos gubernamentales para innovar y, los obstáculos relacionados la existencia de innovaciones anteriores y/o falta de demanda para las innovaciones.

Por su parte, al analizar separadamente el impacto de los indicadores de innovaciones de producto y proceso, fue posible observar un impacto estadístico significativo y positivo de ambos tipos de innovación sobre la productividad media de las empresas.

Por último, la estimación de un modelo probit ordenado para explicar la decisión de realizar actividades de investigación internas, en el que no sólo se considera la decisión de su realización en forma continua sino que además en forma ocasional, nos permitió obtener información más rica desde el punto de vista económico para comprender la toma de decisiones de las empresas con relación a estas actividades. Se observa que ambos mecanismos de protección de la innovación son los factores más relevantes para explicar la realización de actividades de I+D internas por parte de las empresas, seguidos de la importancia que da la empresa a fuentes de información académicas y al carácter exportador de la misma. En este caso, la dimensión empresarial tiene una importancia menor.

Sin embargo, las diferentes variantes de planteamiento teórico y metodología empírica utilizadas aún se encuentran sometidas a una serie importante de problemas cuya superación marcan las futuras líneas de investigación en este campo.

Una de las principales ventajas del modelo estructural de función de producción es que permite capturar las principales características del comportamiento innovador de la empresa, siendo al mismo tiempo parsimonioso y capaz de adecuarse a distintas aplicaciones teóricas que impliquen, por ejemplo, la consideración de nuevos supuestos. Entre los principales supuestos a tratar se pueden citar: la existencia de retornos no constantes escala de los factores productivos y/o la existencia de competencia perfecta en el mercado de producto. En el trabajo de revisión teórica, anterior a la presente investigación, se realiza un análisis detallado en este sentido.

Desde el punto de vista empírico, la presente investigación al trabajar con datos de sección cruzada únicamente pudo constatar correlaciones, que no necesariamente son relaciones causales, entre las variables analizadas. Para poder considerar los rezagos existentes entre la realización de actividades de I+D, su impacto sobre las generación de innovaciones y de estas últimas sobre la productividad, es necesario disponer de un

panel de datos de empresas que tenga en cuenta la variación de la dimensión temporal de las variables analizadas.

La utilización de técnicas de datos de panel además de permitir el análisis de modelos dinámicos, es capaz de controlar por los efectos experiencia de la empresa y heterogeneidad no observable, lo que eventualmente solucionaría ciertos problemas de endogeneidad e inconsistencia en las estimaciones del modelo.

A su vez, trabajar con un panel de datos de empresas permitiría, por ejemplo, la aplicación de un análisis derivado de los modelos convencionales de costos de ajuste de la literatura de la inversión. Esta metodología ha demostrado ser una herramienta analítica sumamente adecuada para analizar la variedad y complejidad de los factores que influyen sobre la decisión de realizar actividades innovadoras por parte de las empresas.³⁸

Sin embargo, es importante tener en cuenta que para construir dicho panel de datos en España es necesario superar un problema importante de solapamiento temporal de la gran mayoría de las variables que surgen de la Encuesta sobre innovación tecnológica en las empresas.

Otro problema empírico a solucionar refiere a la obtención de medidas adecuadas tanto de desempeño económico de las empresas como de factores de producción utilizados. Un camino posible en este sentido es fusionar esta encuesta con otras que contengan este tipo de información.

Finalmente, unos de los problemas importantes al estimar el modelo estructural básico fue la imposición de especificaciones paramétricas a sus diferentes ecuaciones. La limitación principal de esta aproximación es que se basa en fuertes supuestos distribucionales. Otra posible línea futura de investigación consistiría en intentar minimizar o salvar estos supuestos distribucionales a través especificaciones semiparamétricas.

³⁸ Para un detalle sobre este punto se recomienda leer trabajo de 4 créditos en el que analiza de forma detallada el estudio de *Criscuolo y Haskel (2003)* que aplica dicha metodología a un panel de datos de empresas del Reino Unido.

Referencias

- Acs, Z.J. and Audretsch, D.B., 1988, "Innovation in Large and Small Firms: An Empirical Analysis", *The American Economic Review*, 78, 678–690.
- Amemiya, T., "Advanced Econometrics", *Harvard University Press*, Cambridge, MA.
- Arrow, K., 1962; "Economic welfare and the allocation of resources for invention", In: R. Nelson (Ed.), *The Rate and Direction of Inventive Activity*. *Princeton University Press*, Princeton, NJ, pp. 609-626
- Benavente, J., 2006, "The role of research and innovation in promoting productivity in Chile", *Economics of Innovation and New Technology*, 2006, Vol. 15(4/5), June/July, pp. 301–315. Special Issue "On Empirical Studies of Innovation in the Knowledge Driven Economy"
- Bozeman and Link, A., 1983, "Investments in technology: corporate strategies and public policy alternatives", *Praeger, New York*.
- Cameron, A.C and Trivedi, P.K, 2005, "Microeconometrics. Methods and Applications", *Cambridge University Press*.
- Cassiman, B. and Veugelers, R., 1999, "Make and buy in innovation strategies: evidence from Belgian manufacturing firms", *Research Policy* 28, pp. 63-80.
- Cohen, W. and Klepper, S., 1996, "A Reprise of Size and R&D", *The Economic Journal*, Vol. 106, Nº. 437, 925-951
- Cohen, W. and Levinthal, D., 1989, "Innovation and learning: the two faces of R&D", *The Economic Journal* 99, pp. 569-596
- Cragg, J.G., 1971, "Some Statistical Models for Limited Dependent Variables with Application to the Demand for Durable Goods", *Econometrica*, 39, 829-844.
- Crépon, B.; Duguet, E.; Mairesse, J., 1998, "Research, Innovation, and Productivity: An Econometric Analysis at the Firm Level", *NBER Working Paper Series*. Working Paper 6696.
- Criscuolo, C. and Haskel, J. 2003, "Innovations and Productivity Growth in the UK: Evidence from CIS2 and CIS3", *CeRiBA Discussion Paper*, EBPF03-3(10), London
- Denison, E. (1985), "Trends in American Economic Growth, 1929-1982", *Brookings Institution, Washington, D.C.*
- Duguet, E., 2006, "Innovation Height, Spillovers and TFP Growth at the Firm Level: Evidence From French Manufacturing", *Economics of Innovation and New Technology*, Vol. 15(4/5), pp. 415-442.
- Huergo, Elena, 2002, "Determinantes de la innovación tecnológica en la industria manufacturera española", *Revista del Instituto de Estudios Económicos*, nº 3, Págs. 121-141.
- Huergo, E. and Jaumandreu, J., 2004, "Firm's age, process innovation and productivity growth", *International Journal of Industrial Organization*, Nº 22/4, pp. 541-559
- Huergo, E. y Moreno, L., 2004, "La innovación y el crecimiento de la productividad en España", *Ekonomia* N° 56.

- Griffith, R., Huergo, E., Mairesse, J. and Peters, B., 2006, "Innovation and Productivity Across Four European Countries", *NBER Working Paper Series*, Working Paper 12722. National Bureau of Economic Research.
- Greene, W. H., 1998, "Gender Economics Course in Liberal Arts College: Further Results", *Journal of Economic Education*, pp. 291-300
- Greene, W.H., 2003, "Econometric Analysis", fifth edition, *Upper Saddle River*, NJ, Prentice-Hall.
- Griliches, Zvi. 1979. "Issues in Assessing the Contribution of R&D to Productivity Growth.", *Bell Journal of Economics* (10), pp. 92-116.
- Jefferson, G., Huambo, B., Xiaojing, G. and Xiaoyun, Y., 2006, "R&D and Economic Performance in Chinese Industry", *Economics of Innovation and New Technology*, Vol. 15(4/5), pp.345-366. Special Issue "On Empirical Studies of Innovation in the Knowledge Driven Economy"
- Labeaga, J. and Martinez Ros, E., 2003, "Persistence and ability in the innovation decisions", *Working Paper 03-01. Business Economics Series 01*. Universidad Carlos III de Madrid. Departamento de Economía de la Empresa.
- Levin, R. and Reiss, P., 1988, "Cost reducing and demand creating R&D with spillovers", *Rand Journal of Economics* 19, pp. 538-556.
- Lööf, H. and A. Heshmati, 2002, "Knowledge Capital and Performance Heterogeneity: An Innovation Study at Firm Level", *International Journal of Production Economics*, 76(1), 61-85
- Lööf, H. and Heshmati, A., 2002, "On the Relationship between Innovation and Performance: A Sensitivity Analysis", *Economics of Innovation and New Technology Volume 15, Number 4-5*, pp. 317 - 344
- Lööf, H., Heshmati, A., Asplund, R. and Nääs, S. (2001), "Innovation and Performance in Manufacturing Industries: a Comparison of the Nordic Countries", *Stockholm School in Economics: Swedish Working Paper Series in Economics and Finance*, 2001, p. 457.
- Mairesse, J. and Mohnen, P., 2005, "Accounting for Innovation and Measuring Innovativeness: An Illustrative Framework and an Application", *American Economic Review*, 92(2), 226–230.
- Mansfield, E., 1985, "How Rapidly Does New Industrial Technology Leak Out?", *The Journal of Industrial Economics* 34, pp. 217-223.
- Olley, G. S. and Pakes, A., 1996, "The dynamics of productivity in the telecommunications equipment industry". *Econometrica* 64(6), 1263-1297.
- Pakes A. and Z. Griliches, 1984, "Patents and the R&D at Firm Level: A First Look", in Griliches (ed.), R&D, Patents and Productivity, Chicago: *University of Chicago Press*, pp. 55-72.
- Parisi, M., Schiantarelli, F. and Sembenelli, A. (2005), "Productivity, Innovation and R&D: Micro Evidence for Italy", *European Economic Review*.
- Romer, Paul, 1994, "The origins of Endogenous Growth", *Journal of Economics Perspectives*, Volume 8, Number 1.
- Scherer, F. M., 1965, "Firm size, market structure, opportunity and the output of patented innovations", *American Economic Review* 55, pp. 1097-1125

Schmookler, J., 1966, "Invention and Economic Growth", *Cambridge, MA: Harvard University Press.*

Schumpeter, J. A., 1942, "Capitalism, Socialism and Democracy", *Unwin University Books, London.*

Solow, Robert M., 1956, "A Contribution to the Theory of Economic Growth.", *Quarterly Journal of Economics*, 70:65-94

Solow, R., 1957, "Technical Change and the Aggregate Production Function", *The Review of Economics and Statistics*, 39, 312–320.

Spence, M., 1984, "Cost Reduction, Competition and Industry Performance", *Econometrica*, pp.101-121

Van Leeuwen, G. and Klomp, L., 2006, "On the contribution of Innovation to Multi-Factor Productivity Growth", *Economics of Innovation and New Technology*, Vol. 15(4/5), pp.367-390. Special Issue "On Empirical Studies of Innovation in the Knowledge Driven Economy"

Van Leeuwen, G., 2002, "Linking innovation to productivity growth using two waves of the community Innovation Survey".

Winter, S.G., 1987, "Knowledge and Competence as Strategic Asconjuntos", In Teece, D.J. (ed.) *The Competitive Challenge*. NewYork: Harper and Row, pp. 159–183.

Wooldridge, J. M., "Econometric analysis of Cross Section and Panel Data", *The MIT Press*, Cambridge, Massachusetts, London, England.

Anexo I. Definición de variables

| VARIABLE | DESCRIPCIÓN | PERÍODO DE REFERENCIA |
|-----------------------------------|---|-----------------------|
| Conocimiento/Innovación | | |
| Intensidad de gastos en I+D | Gasto en actividades internas de I+D por trabajador (en logaritmos) | 2002-2004 |
| Gastos continuos en I+D interna | Variable dicotómica que toma el valor 1 si la empresa reporta gastos en actividades de I+D interna de forma continua | 2002-2004 |
| Gastos en I+D interna | Variable categórica que toma el valor 0 si la empresa no realizó actividades de I+D interna, el valor 1 si lo hizo en forma ocasional y valor 2 si lo hizo en forma continua. | 2002-2004 |
| Innovaciones de proceso | Variable dicotómica que toma el valor 1 si la empresa reporta que ha introducido procesos de producción nuevos o significativamente mejorados | 2002-2004 |
| Innovaciones de producto | Variable dicotómica que toma el valor 1 si la empresa ha introducido bienes o servicios nuevos o mejorados de manera significativa | 2002-2004 |
| Financiamiento público | | |
| Financiamiento local o autonómico | Variable dicotómica que toma el valor 1 si la empresa recibió apoyo financiero de las administraciones locales o autonómicas | 2002-2004 |
| Financiamiento gobierno central | Variable dicotómica que toma el valor 1 si la empresa recibió apoyo financiero de las administraciones del Estado | 2002-2004 |
| Financiamiento UE | Variable dicotómica que toma el valor 1 si la empresa recibió apoyo financiero de la UE | 2002-2004 |

| VARIABLE | DESCRIPCIÓN | PERÍODO DE REFERENCIA |
|--|--|-----------------------|
| Arrastre de la Demanda | | |
| Estándares elevados | Porcentaje de empresas para las cuales el cumplimiento de los requisitos normativos ha tenido una importancia intermedia/elevada | 2002-2004 |
| Estándares reducida | Porcentaje de empresas para las cuales el cumplimiento de los requisitos normativos ha tenido una importancia reducida | 2002-2004 |
| Aspectos medioambientales elevados | Porcentaje de empresas para las cuales el menor impacto medioambiental o mejora en la salud y la seguridad ha tenido una importancia intermedia/elevada | 2002-2004 |
| Aspectos medioambientales reducidos | Porcentaje de empresas para las cuales el menor impacto medioambiental o mejora en la salud y la seguridad ha tenido una importancia reducida | 2002-2004 |
| Fuentes de Información para innovar | | |
| Fuentes internas | Variable dicotómica que toma el valor 1 si la información de fuentes internas a la empresa o grupo fue de importancia elevada | 2002-2004 |
| Fuentes universitarias | Variable dicotómica que toma el valor 1 si la información de fuentes universitarias u otros centros de enseñanza superior grupo fue de importancia elevada | 2002-2004 |
| Fuente organismos públicos | Variable dicotómica que toma el valor 1 si la información de organismo públicos de investigación fueron de importancia elevada | 2002-2004 |
| Fuente proveedores | Variable dicotómica que toma el valor 1 si la información de proveedores de equipo material, componentes o software fue de importancia elevada | 2002-2004 |
| Fuente competidores | Variable dicotómica que toma el valor 1 si la información de competidores u otras empresas de su misma rama de actividad fue de importancia elevada | 2002-2004 |
| Fuente consumidores | Variable dicotómica que toma el valor 1 si la información de clientes fue de importancia elevada | 2002-2004 |

| VARIABLE | DESCRIPCIÓN | PERÍODO DE REFERENCIA |
|---|---|-----------------------|
| Condiciones de apropiabilidad de los resultados de la innovación | | |
| Protección formal | Variable dicotómica que toma el valor 1 si la empresa utilizó registros de modelos de utilidad, marcas de fabrica o derechos de autor para proteger sus invenciones o innovaciones | 2002-2004 |
| Protección estratégica | Variable dicotómica que toma el valor 1 si la empresa utilizó modificaciones significativas del diseño o envasado de un bien o servicio, reducción del período de respuesta a un cliente o proveedor o cambios significativos en las relaciones con otras empresas o instituciones para proteger sus invenciones o innovaciones | 2002-2004 |
| Acuerdos de cooperación para innovar | | |
| Cooperación | Variable dicotómica que toma el valor 1 si la empresa tuvo algún acuerdo de cooperación en actividades innovadoras | 2002-2004 |
| Mercado | | |
| Mercado internacional | Variable dicotómica que toma el valor 1 si el mercado más significativo para la empresa es el internacional | 2004 |
| Obstáculos para innovar | | |
| Obstáculos de información | Variable continua de 0 a 1 que toma en cuenta la importancia que tienen los obstáculos para innovar relacionados con: falta de personal cualificado, falta de información sobre tecnología y sobre mercados | 2002-2004 |
| Obstáculos de costes | Variable continua de 0 a 1 que toma en cuenta la importancia que tienen los obstáculos para innovar relacionados con: falta de fondos dentro de la empresa o grupo, falta de financiación externa a la empresa y la existencia de costes de innovación elevados. | 2002-2004 |
| Obstáculos Nec | Variable continua de 0 a 1 que toma en cuenta la importancia que tienen los obstáculos para innovar debido a la existencia de innovaciones anteriores y la falta de demanda para las innovaciones. | 2002-2004 |

| VARIABLE | DESCRIPCIÓN | PERÍODO DE REFERENCIA |
|--|--|-----------------------|
| Productividad | | |
| Productividad media | Ventas por trabajador (en logaritmos) | 2004 |
| Inversión en capital físico | | |
| Intensidad Inversión | Inversión bruta media en bienes materiales en 2004 en la agrupación a que pertenece la empresa (en logaritmos) | 2004 |
| Grupo empresarial | | |
| Grupo empresarial | Variable dicotómica que toma el valor 1 si la empresa pertenece a un grupo empresarial | 2002-2004 |
| Propiedad capital de la empresa | | |
| Empresa pública | Variable dicotómica que toma el valor 1 si la empresa es pública | 2004 |
| Empresa privada nacional | Variable dicotómica que toma el valor 1 si la empresa es privada nacional | 2004 |
| Empresa multinacional | Variable dicotómica que toma el valor 1 si la empresa es una multinacional | 2004 |
| Asociación de investigación | Variable dicotómica que toma el valor 1 si la empresa es una asociación de investigación | 2004 |
| Otras variables | | |
| Empresa de nueva creación | Variable dicotómica que toma el valor 1 si la empresa pertenece a un grupo empresarial | 2002-2004 |
| Tamaño | Conjunto de 6 variables dicotómicas de acuerdo al número de empleados de la empresa. Las categorías son: 0-19, 20-49, 50-99, 100-249, 250-999, 1000 o más. | 2004 |
| Ramas de actividad | Conjunto de variables dicotómicas de acuerdo a la rama de actividad del sector manufacturero de la empresa (11 sectores CNAE-93) | 2002-2004 |

Anexo II. Descriptivos de variables

| Variable | Media | Desviación estándar | Mínimo | Máximo |
|-------------------------------------|--------------|----------------------------|---------------|---------------|
| I+D | | | | |
| Intensidad de gastos en I+D | 7.6869 | 6.0531 | 0 | 16.2121 |
| Gastos continuos en I+D interna | 0.5367 | 0.4987 | 0 | 1 |
| Gastos en I+D interna | 1.2125 | 0.9033 | 0 | 2 |
| Innovación | | | | |
| Innovaciones de proceso | 0.5667 | 0.4955 | 0 | 1 |
| Innovaciones de producto | 0.5799 | 0.4936 | 0 | 1 |
| Financiamiento | | | | |
| Financiamiento local o autonómico | 0.2533 | 0.4350 | 0 | 1 |
| Financiamiento gobierno central | 0.2242 | 0.4171 | 0 | 1 |
| Financiamiento UE | 0.0561 | 0.2299 | 0 | 1 |
| Arrastre de demanda | | | | |
| Estándares elevados | 0.4083 | 0.0986 | 0.1526 | 0.625 |
| Estándares reducida | 0.1599 | 0.0331 | 0 | 0.25 |
| Aspectos medioambientales elevados | 0.3721 | 0.1012 | 0.1369 | 0.75 |
| Aspectos medioambientales reducidos | 0.1910 | 0.0507 | 0 | 0.3218 |
| Fuente de información | | | | |
| Fuentes internas | 0.4629 | 0.4986 | 0 | 1 |
| Fuentes universitarias | 0.0672 | 0.2505 | 0 | 1 |
| Fuente organismos públicos | 0.0351 | 0.1840 | 0 | 1 |
| Fuente proveedores | 0.1509 | 0.3580 | 0 | 1 |
| Fuente competidores | 0.1023 | 0.3031 | 0 | 1 |
| Fuente consumidores | 0.2161 | 0.4117 | 0 | 1 |
| Protección innovación | | | | |
| Protección formal | 0.3489 | 0.4766 | 0 | 1 |
| Protección estratégica | 0.3201 | 0.4665 | 0 | 1 |
| Acuerdos de cooperación | | | | |
| Cooperación | 0.3139 | 0.4707 | 0 | 1 |
| Mercado de destino | | | | |
| Mercado internacional | 0.7725 | 0.4192 | 0 | 1 |
| Obstáculos para innovar | | | | |
| Obstáculos de información | 0.3924 | 0.2711 | 0 | 1 |
| Obstáculos de costes | 0.5333 | 0.3236 | 0 | 1 |
| Obstáculos Nec | 0.2258 | 0.2714 | 0 | 1 |

| Variable | | Media | Desviación estándar | Mínimo | Máximo |
|--|-----------------|--------------|----------------------------|---------------|---------------|
| Productividad | | | | | |
| Productividad media | | 11.8499 | 0.8542 | 4.9182 | 16.2100 |
| Inversión | | | | | |
| Intensidad Inversión | | 6.7845 | 3.5274 | 0 | 15.9562 |
| Grupo empresarial | | | | | |
| Grupo empresarial | | 0.3939 | 0.4886 | 0 | 1 |
| Propiedad capital de la empresa | | | | | |
| Empresa pública | | 0.0056 | 0.0747 | 0 | 1 |
| Empresa privada nacional | | 0.8407 | 0.3659 | 0 | 1 |
| Empresa multinacional | | 0.1516 | 0.3587 | 0 | 1 |
| Asociación de investigación | | 0.0020 | 0.0449 | 0 | 1 |
| Otras | | | | | |
| Empresa de nueva creación | | 0.3939 | 0.4886 | 0 | 1 |
| Tamaño de empresa | | | | | |
| 0 a 19 trabajadores | | 0.1712 | 0.3767 | 0 | 1 |
| 20 a 49 trabajadores | | 0.2539 | 0.4353 | 0 | 1 |
| 50 a 99 trabajadores | | 0.1750 | 0.3800 | 0 | 1 |
| 100 a 249 trabajadores | | 0.1903 | 0.3926 | 0 | 1 |
| 250 a 999 trabajadores | | 0.1862 | 0.3893 | 0 | 1 |
| 1000 o más trabajadores | | 0.0231 | 0.1504 | 0 | 1 |
| Ramas de actividad | CNAE -93 | | | | |
| Alimentos y bebidas | 15-16 | 0.1196 | 0.3246 | 0 | 1 |
| Textil | 17-19 | 0.0726 | 0.2596 | 0 | 1 |
| Papel y derivados | 20-22 | 0.0668 | 0.2497 | 0 | 1 |
| Químicos | 23-24 | 0.1397 | 0.3467 | 0 | 1 |
| Plásticos | 25 | 0.0544 | 0.2269 | 0 | 1 |
| No metales | 26 | 0.0621 | 0.2413 | 0 | 1 |
| Metales básicos | 27-28 | 0.1172 | 0.3217 | 0 | 1 |
| Maquinaria | 29 | 0.1210 | 0.3262 | 0 | 1 |
| Electrónicos | 30-33 | 0.1241 | 0.3298 | 0 | 1 |
| Vehículos | 34-35 | 0.0681 | 0.2520 | 0 | 1 |
| Nec | 36 | 0.0539 | 0.2260 | 0 | 1 |

Anexo III. Los efectos marginales en los modelos probit bivariantes recursivos

En primer lugar, se calcula el efecto de la realización de actividades de I+D internas continuas sobre la probabilidad que la empresa i obtenga innovaciones de producto (o proceso). Este efecto se obtiene por la diferencia de probabilidades condicionales de obtener una innovación de producto (o proceso) realizando y no realizando este tipo de actividades:

$$H(g_i) = \Pr ob(I_i = 1/x_{i0}, x_{i2}, g_i = 1) - \Pr ob(I_i = 1/x_{i0}, x_{i2}, g_i = 0)$$

En segundo lugar, se calculan los efectos marginales de las variables exógenas sobre la probabilidad de obtener innovaciones.

1. Si la variable exógena es binaria:

Consideremos los vectores x_{i0}^0, x_{i2}^0 en los cuales la variable binaria es igual a 0, y los vectores para los cuales es igual a 1, x_{i0}^1, x_{i2}^1 , manteniendo el resto de las variables en su valor medio muestral.

Por tanto, el efecto de la variable binaria x_i está dado por:

$$\begin{aligned} G(x_i) &= E(I/x_{i0}^1, x_{i2}^1) - E(I/x_{i0}^0, x_{i2}^0) = N_2(b_0^1 x_{i0}^1, b_2^1 x_{i2}^1 + \gamma_{g_i}, \rho_i) + N_2(-b_0^1 x_{i0}^1, b_2^1 x_{i2}^1, -\rho_i) - \\ &\quad - N_2(b_0^0 x_{i0}^0, b_2^0 x_{i2}^0 + \rho_i) - N_2(-b_0^0 x_{i0}^0, b_2^0 x_{i2}^0, -\rho_i) \end{aligned}$$

Obsérvese que el efecto puede descomponerse en dos partes:

$$G(x_i) = G_1(x_i) + G_2(x_i)$$

$$G_1(x_i) = N_2(b_0^1 x_{i0}^1, b_2^1 x_{i2}^1 + \gamma_{g_i}, \rho_i) - N_2(b_0^0 x_{i0}^0, b_2^0 x_{i2}^0 + \rho_i)$$

$$G_2(x_i) = N_2(-b_0^1 x_{i0}^1, b_2^1 x_{i2}^1, -\rho_i) - N_2(-b_0^0 x_{i0}^0, b_2^0 x_{i2}^0, -\rho_i)$$

La primera parte, $G_1(x_i)$, es el efecto de la variable binaria sobre la probabilidad de innovar habiendo realizado actividades de I+D, y la segunda, $G_2(x_i)$, es el efecto sobre la probabilidad de innovar sin realizar actividades de I+D.

2. En el caso de una variable exógena continua, x_h , su efecto marginal viene dado por la expresión:

$$\frac{\partial E(I/x_{i0}, x_{i2})}{\partial x_h} = G_1(x_h) + G_2(x_h)$$

$$G_1(x_h) = \phi(b_0^x x_{i0}) \Phi \left(b_2^x x_{i2} + \gamma_{g_i} - \frac{\rho b_0^x x_{i0}}{\sqrt{1-\rho_i^2}} \right) b_0^h + \phi(b_2^x x_{i2} + \gamma_{g_i}) \Phi \left(b_0^x x_{i0} - \frac{\rho(b_2^x x_{i2} + \gamma_{g_i})}{\sqrt{1-\rho_i^2}} \right) b_2^h$$

$$G_2(x_h) = \phi(-b_0^x x_{i0}) \Phi \left(b_2^x x_{i2} - \frac{\rho b_0^x x_{i0}}{\sqrt{1-\rho_i^2}} \right) b_0^h + \phi(b_2^x x_{i2}) \Phi \left(-b_0^x x_{i0} + \frac{\rho b_2^x x_{i2}}{\sqrt{1-\rho_i^2}} \right) b_2^h$$