



---

**Universidad de Valladolid**

**Determinantes de la inversión en I+D:  
un análisis empírico en empresas europeas  
de *software* y servicios informáticos**

**JAVIER FERNÁNDEZ LÓPEZ**

**Máster de Investigación en Contabilidad y Gestión  
Financiera**

**Facultad de C. Económicas y Empresariales**

**Trabajo fin de master dirigido por**

**Dr. José Miguel Rodríguez Fernández**

**Julio 2013**

# INDICE

<b>INDICE</b> .....	<b>2</b>
<b>INTRODUCCIÓN</b> .....	<b>3</b>
1.    OBJETIVOS .....	4
1.1. <i>General</i> .....	4
1.2. <i>Específicos</i> .....	4
2.    METODOLOGÍA EMPLEADA Y ESTRUCTURA DEL TRABAJO .....	5
<b>CAPÍTULO I. MARCO TEÓRICO: UNA REVISIÓN DE LA LITERATURA</b> .....	<b>6</b>
1.    LA I+D COMO UNA PARTE DEL PROCESO DE INNOVACIÓN .....	6
2.    EL “PROBLEMA” DE LA I+D .....	8
3.    LA INDUSTRIA DEL <i>SOFTWARE</i> : UN EJEMPLO DE INDUSTRIA INTENSIVA EN I+D .....	8
3.1. <i>La industria del software: un poco de historia</i> .....	9
3.2. <i>Sectores y compañías principales</i> .....	10
3.3. <i>Tendencias actuales de la industria del software</i> .....	10
3.4. <i>Amenazas en la industria del software</i> .....	11
4.    FACTORES DETERMINANTES DE LA INVERSIÓN EN I+D .....	12
4.1. <i>Tamaño</i> .....	14
4.2. <i>Concentración de la propiedad</i> .....	16
4.3. <i>Endeudamiento</i> .....	17
4.4. <i>Desempeño</i> .....	17
<b>CAPÍTULO II. DISEÑO DE LA INVESTIGACIÓN EMPÍRICA</b> .....	<b>18</b>
1.    HIPÓTESIS .....	18
2.    VARIABLES UTILIZADAS .....	22
3.    FUENTES DE DATOS .....	24
4.    MUESTRA DE EMPRESAS .....	25
5.    MODELOS Y MÉTODOS DE ESTIMACIÓN .....	27
5.1. <i>Modelo de efectos fijos (fixed effects)</i> .....	27
5.2. <i>Modelo de efectos aleatorios (random effects)</i> .....	28
5.3. <i>Elección entre efectos fijos y aleatorios: test de Hausman</i> .....	29
5.4. <i>Modelo de mínimos cuadrados generalizados factibles</i> .....	30
5.5. <i>Regresión con errores estándar corregidos por autocorrelación y heterocedasticidad (Prais-Winsten)</i> .....	31
6.    ESTRATEGIA ECONOMETRICA DE LA INVESTIGACIÓN .....	33
<b>CAPÍTULO III. RESULTADOS EMPÍRICOS Y CONTRASTE DE HIPÓTESIS</b> .....	<b>34</b>
1.    ESTADÍSTICOS DESCRIPTIVOS .....	34
2.    ANÁLISIS DE CORRELACIÓN .....	35
3.    ESTIMACIONES CON EFECTOS ALEATORIOS Y EFECTOS FIJOS. TEST DE HAUSMAN .....	35
4.    ANÁLISIS DE AUTOCORRELACIÓN Y HETEROCEDASTICIDAD: EFECTOS FIJOS Y ALEATORIOS. ...	39
5.    CONTRASTE DE AUTOCORRELACIÓN .....	39
6.    CONTRASTES DE HETEROCEDASTICIDAD (EFECTOS FIJOS) .....	40
7.    CONTRASTES DE HETEROCEDASTICIDAD (EFECTOS ALEATORIOS) .....	41
8.    RECONSIDERACIÓN DEL TEST DE HAUSMAN Y CONTRASTE DE RESTRICCIONES SOBREIDENTIFICADAS .....	41
9.    REGRESIÓN DE PANEL CON ERRORES ESTÁNDAR CORREGIDOS POR AUTOCORRELACIÓN Y HETEROCEDASTICIDAD (PRAIS-WINSTEN) .....	42
10.   RESULTADOS GLOBALES DEL CONTRASTE DE LAS HIPÓTESIS DE PARTIDA .....	47
<b>CAPÍTULO IV. A MODO DE CONCLUSIÓN</b> .....	<b>50</b>
<b>REFERENCIAS BIBLIOGRÁFICAS</b> .....	<b>52</b>
<b>ÍNDICE DE CUADROS</b> .....	<b>57</b>

## Introducción

Actualmente, no existe discusión acerca de la influencia en la capacidad de innovación de una región a la hora de lograr un crecimiento económico, son muchas las causas de la actual crisis económica y de liderazgo que sufrimos en Europa. En un contexto de fuerte recesión, no cabe duda que una de las mejores formas de creación de valor añadido, y por tanto de riqueza, es por medio de la innovación.

Desde el descubrimiento del fuego pasando por la revolución industrial, llegando a la actual revolución tecnológica e informática, las naciones han logrado su desarrollo por medio de la innovación, y para ello hay que invertir recursos de manera continuada a todos los niveles, tanto en educación como a nivel empresarial e institucional. Estos son los pilares que sustentan la creación de valor añadido por medio de la investigación, que como se ha demostrado a lo largo de la historia es un crecimiento de calidad y sostenido.

Cabe destacar, por otra parte, que la Unión Europea se ha percatado de que la investigación es una de las claves para el futuro desarrollo de la zona, y está llevando a cabo un plan de desarrollo de I+D a escala europea llamado “Horizonte 2020” cuya finalidad es potenciar el desarrollo de la innovación, tratando de incentivarlo hasta el punto que éste sea un 3% del PIB de la zona para el año 2020.

La actual revolución tecnológica se caracteriza por la imponente expansión de sectores de proceso de la información. La informática e Internet han cambiado los modelos empresariales y la sociedad en su conjunto. Por ello que nadie duda de que uno de los sectores con una notable importancia en nuestros días, y con unas mayores posibilidades de futuro, es el relacionado con la creación de *software*. Es un fenómeno tan reciente que, pese a ser uno de los más punteros, a su vez, es uno de los sectores menos estudiados por la academia.

Tras contextualizar el problema, todo ello nos ha llevado a plantearnos cuál es la importancia de la inversión en I+D, y más concretamente en la industria del *software*, dentro del marco de la Unión Europea, una de las zonas con más recursos del planeta, y actualmente con una grave crisis económica.

Por lo tanto, parece un requisito fundamental tanto para la academia como para el ámbito empresarial conocer los determinantes fundamentales que incentivan la inversión en I+D en las empresas de *software* en la Unión Europea.

## **1. Objetivos**

El presente proyecto se enmarca en una investigación cuya finalidad es la explicación, la inferencia causal entre la variable dependiente y las variables independientes seleccionadas, que posteriormente se justificarán. Se encuentra, por ende, estructurado en dos partes.

### **1.1. General.**

Desarrollar un modelo que proporcione una idea sobre los factores que determinan la intensidad de inversión en I+D en empresas de *software*, las cuales, por las características de su negocio tienen una especial singularidad. Como referencia del estudio, hemos tomado empresas que se encuentran dentro del área geográfica de Europa.

### **1.2. Específicos.**

Desarrollar un estudio empírico que explique cuáles son las variables que determinan la intensidad de la inversión en I+D en las empresas de alta tecnología.

Otorgar pautas para las empresas e instituciones europeas que les sirvan de apoyo en el cumplimiento de las directrices pactadas por la Comisión Europea (2010) con la hoja de ruta llamada “Horizonte 2020”.

Reconocer la relevancia del sector del *software* en el desarrollo económico actual y futuro como uno de los más importantes para la mejora de la sociedad en su conjunto y a todos los niveles.

Por un lado, se ahonda en la conceptualización del término de I+D haciendo especial hincapié en la industria del *software*, el objeto de esta investigación. Para ello se realiza una revisión literaria relativa a los estudios realizados sobre los factores determinantes a la hora de tener una mayor intensidad en I+D.

Por otra parte, y mediante herramientas econométricas, se hace un estudio empírico que persigue conocer si una serie de variables independientes anteriormente estudiadas en otros trabajos son a su vez determinantes para realizar una inversión en I+D en algunas de las empresas de *software* más grandes de Europa entre 2004-2010.

Con lo cual, no se persigue explicar cuáles son los efectos o consecuencias económicas que la intensidad de la inversión en I+D provoca en las empresas. Como así tampoco es objeto de investigación el proceso de toma de decisiones por el cual unas empresas apuestan por esta forma de creación de riqueza mediante la innovación. Sino que la

intención es conocer cuáles son las bases que influyen en la inversión en I+D en las empresas de *software*.

## **2. Metodología empleada y estructura del trabajo**

La presente investigación tiene un enfoque teórico y empírico y está estructurada de la siguiente manera:

### **Capítulo I.**

En este capítulo se procederá a realizar, una revisión literaria sobre estudios empíricos previos que sirven de base para posteriormente realizar la investigación que nos concierne en este trabajo. A su vez definiremos y veremos cuáles son las variables incluidas en otros trabajos que se enmarcarán dentro de nuestro proyecto.

### **Capítulo II.**

Tras la revisión de la literatura pasamos a realizar un diseño de la investigación empírica. En este punto formularemos las hipótesis que trataremos de contrastar, variables a utilizar, bases de datos, muestra realizada y la estrategia econométrica utilizada para llegar a los resultados.

### **Capítulo III.**

Este capítulo incluye un análisis empírico, el cual busca explicar los factores determinantes de la intensidad de la inversión en I+D en compañías europeas de *software*, mediante el contraste de las correspondientes hipótesis, la aplicación de las oportunas técnicas econométricas y la ayuda del paquete estadístico STATA. Se formulan cuatro hipótesis sometidas a comprobación.

### **Capítulo IV.**

Finalmente, el último apartado de este trabajo son las conclusiones del mismo, viendo si se han cumplido las hipótesis previstas y daremos respuesta a la pregunta de este proyecto, qué variables son susceptibles de ser determinantes a la hora de realizar una inversión en I+D.

## **CAPÍTULO I. MARCO TEÓRICO: UNA REVISIÓN DE LA LITERATURA.**

El estudio de la innovación ha adquirido un papel protagonista en la literatura económica empresarial desde el siglo XIX, algunos de los más importantes autores de la historia en la ciencia económica como Adam Smith (2001) entre otros, ya enmarcaban a la innovación como uno de los pilares de la creación de riqueza, y cómo gracias a la innovación se podría no solo conseguir el éxito empresarial, sino que a su vez lograr que los ciudadanos gocen de una mayor prosperidad, mejorando la calidad de vida y los niveles de desarrollo de las sociedades.

Toda innovación se deriva de una idea creativa, y de implementación exitosa de las mismas, es la clave de la innovación en las organizaciones y con ello del éxito, no sólo de las empresas, sino de la sociedad en su conjunto. Desde este punto de vista, la creatividad de los individuos y los equipos son un punto de partida para la innovación, la primera es condición necesaria pero no suficiente para la segunda (Amabile, 1982).

Como veremos la academia es consciente de la necesidad del estudio de la innovación para lograr una mejora de la sociedad. Así, existe una amplia literatura que permite evidenciar dos tipos de estudios. Aquellos de naturaleza macroeconómica y aquellas de índole microeconómica. Para con ello, y como veremos más adelante, lograr determinar qué factores determinan la creación de innovación y así mejorar la calidad de vida.

### **1. La I+D como una parte del proceso de innovación**

La definición del término innovación que se utiliza como referencia en numerosos trabajos de investigación es, las actividades de innovación tecnológica son el conjunto de etapas científicas, tecnológicas, organizativas, financieras y comerciales, incluyendo las inversiones en nuevos conocimientos, que llevan o que intentan llevar a la implementación de productos y de procesos nuevos o mejorados (OCDE, 2002). A su vez, existen otras definiciones igualmente aceptadas como válidas. La búsqueda y el descubrimiento, el desarrollo, la mejora, la aprobación y comercialización de nuevos procesos, nuevos productos y nuevas estructuras organizativas y procedimientos (Dosi, 1988; Wennberg, 2009).

En este sentido, y a colación con lo anterior, es importante distinguir el término de innovación del término de Investigación y el Desarrollo (de ahora en adelante I+D).

- La innovación comprende una serie de actividades científicas, tecnológicas, organizativas, financieras y comerciales de nuevos productos (Sánchez Bueno, 2008: 21).
- La I+D es una de las partes que comprenden el proceso complejo de innovación, mas no la única tal y como se ha explicado anteriormente.

Es por ello que, a partir de la literatura consultada, se estima que su estudio deviene relevante por cuatro razones esenciales.

En primer lugar, como se ha señalado, se trata de un tema ampliamente abordado por la literatura durante los últimos veinte años, que soporta en sí mismo toda una serie de externalidades positivas a dos niveles: (a) macroeconómico debido al importante desarrollo económico que conlleva y (b) empresarial.

En segundo lugar, desde un análisis macroeconómico, numerosos autores sostienen que el esfuerzo de innovación lleva consigo un mayor desarrollo económico (Baumol, 2002). No obstante, la academia no encuentra consenso en lo que a este punto se refiere, ya que no es segura la causalidad directa entre la innovación y el crecimiento.

Por último, y desde el análisis que nos interesa, el examen empresarial, la innovación produce un *pseudocontagio* positivo hacia otras empresas, el efecto derrame se extiende a terceras empresas y a las nuevas. Así también, el Manual de Oslo (1997) apoya esta hipótesis recogiendo algunos de los efectos positivos que los procesos de investigación e innovación pueden generar a las empresas. Entre ellos destaca, el desarrollo y mejora de productos que pueden incidir directamente en su poder en el mercado, o eficiencia en los costes soportados en los procesos productivos. Es decir, I+D se traduce en más ventajas competitiva de la empresa (González y Pargas, 2010).

En definitiva, tanto la academia como los centros privados argumentan la existencia de un impacto positivo en el ámbito empresarial (Woolridge, 1988).

A su vez, la Unión Europea, especialmente a partir de la Comisión Europea, ha dedicado esfuerzos al desarrollo tecnológico y la investigación científica. Así, en el año 2002 señala a la I+D como una de las áreas prioritarias a potenciar la inversión en conocimiento para asegurar la futura competitividad y el empleo. La Unión Europea debe intensificar los esfuerzos en materia de investigación, innovación, educación y formación, y aumentar su impacto mediante la aplicación de un enfoque más integrado y agrupando estas políticas bajo una bandera común: un espacio europeo del conocimiento. (OCDE, 2002). En el año 2010, con la iniciativa de “Unión por la innovación”; publica la “*Nueva estrategia europea de crecimiento para el 2020*” donde especifica que la UE tiene como objetivo invertir el 3 % de su PIB en I+D, Comisión Europea (2010: 15), centrándose en el impacto y la composición de su inversión en investigación y mejorando las condiciones de la I+D del sector privado en la UE.

## 2. El “problema” de la I+D

Como argumenta Griliches (1979), otro aspecto ligado a la actividad de I+D es la distinción entre I+D del *input* y de *output*. Mientras que la I+D del *input* hace referencia a los esfuerzos realizados por las empresas para llegar a nuevos productos, nuevas formas de producir su producto o para administrar sus negocios de manera más eficiente y para conquistar nuevos mercados, el *output* está relacionada con los nuevos productos o procesos introducidos con éxito, el aumento de las utilidades o eficiencia. En el lado de entrada, la primera medida que viene a la mente es I + D (Mohnen y Hall, 2013).

De esta manera, y desde un marco teórico más actual, la I+D como *input* es considerada como un nexo de la innovación a pesar de sus limitaciones. López-Mielgo et al. (2005) reconocen dos:

En primer lugar, esgrimen la necesidad de identificar qué importes incluyen las empresas en las partidas de I+D. Mairesse y Hall (1996), ya apuntaron al problema de la doble vía posible para la contabilización de la I+D, es decir, como gasto o como inversión. Esto sucede porque, en casos como el de Francia, las empresas contabilizan como I+D aquellos gastos de los que pueden obtener beneficios fiscales Lhuillery y Templé (1995); en otros países, como es el caso de España, generalmente las empresas contabilizan como I+D los conceptos más evidentes y de cálculo más sencillo.

En segundo lugar, la inversión en I+D minusvaloran las actividades de innovación de las pequeñas y medianas empresas, ya que en sí mismos son una función de la especialización y tamaño empresarial. Patel y Pavitt (1994), estiman que cuando tratan el cambio técnico como sinónimo de I+D las Pymes, están dejando de considerar, en algunos sectores, cerca del 40% de la inversión en las empresas de menor tamaño (López-Mielgo et al., 2005).

## 3. La industria del *software*: un ejemplo de industria intensiva en I+D.

La industria del *software* ha crecido de manera exponencial en los últimos años, según la empresa de investigación de mercados Datamonitor, quien sostiene que el tamaño de la industria del *software* a nivel mundial en 2008 fue de 303.8 mil millones de dólares EE.UU., lo cual supone un aumento del 6,5% respecto a 2007. Su crecimiento, por lo tanto, está siendo imparable.

Los datos mencionados no sólo hacen de esta industria una de las más prósperas, sino también que se caracterizan por ser una fuente de crecimiento tanto en empleo como de surgimiento de nuevas compañías.

Sin embargo, a pesar de su rápido crecimiento y la importancia que obtiene en el mercado internacional en términos de competitividad para otras industrias de alta tecnología y de crecimiento nacional, la industria del *software* ha recibido



sorprendentemente poca atención de los académicos (Mowery, 1995). Es por ello, que no puede entenderse esta industria sin realizar una breve referencia a su historia.

### **3.1. La industria del *software*: un poco de historia.**

La industria del *software* se desarrolló en los años 60, especialmente en EEUU. En un principio cada compañía que adquiría un ordenador contrataba un programador para desarrollar su propio *software*, pero los programadores se dieron cuenta de que era posible utilizar el mismo programa en varios ordenadores. Con lo que se fueron formando empresas de base comercial en las que los programadores se agrupaban para desarrollar *software* para todo tipo de compañías (Alberts et al., 2004). A su vez, los fabricantes de hardware comenzaron a agrupar los sistemas operativos, *software* de sistemas y entornos de programación para sus máquinas. Y su evidente necesidad mutua a la hora de progresar tecnológicamente.

No obstante, el punto de inflexión de la industria se sitúa a mediados de la década de los '70 con el surgimiento del ordenador personal. Tal aparición hundió los precios del hardware, gracias a lo cual se produjo una gran innovación en términos de nuevos lenguajes de programación y metodologías de desarrollo del *software*, provocando la utilización masiva de ordenadores en el ámbito laboral. Durante los años siguientes, los juegos, aplicaciones y diversas utilidades fueron los protagonistas de un mercado creciente.

El increíble desarrollo que se produjo en la industria del *software* a partir de 1995, el año de la “liberación” de Netscape y el nacimiento de la Web, cambió el modo de utilización y de negocio de las empresas, ya que variaron tanto sus modelos de producción como de distribución de productos (Campbell-Kelly, 2012).

Llegados al principio del siglo XXI, “*software-As-A-Service*” o “SaaS” se postuló como otro exitoso modelo de negocio denominado “informática en la nube”. Desde el punto de vista de algunos propietarios de *software*, tal modelo reduce las preocupaciones acerca de la copia no autorizada, ya que sólo se puede acceder a través de la Web, y por definición no hay *software* de cliente que se cargue en el PC del usuario final. Pero, lo más relevante es que supuso una completa redefinición del modelo de negocio del *software*. Es decir, los programas informáticos en la nube establecen un sistema de pago por uso, en lugar del clásico pago por licencia, suponiendo un mayor beneficio empresarial.

### 3.2. Sectores y compañías principales

Actualmente podemos distinguir tres tipos de negocios esenciales dentro de la industria del *software*. Son los siguientes.

En primer lugar, lo que se denomina *software* de Infraestructura, donde a su vez se pueden diferenciar los sistemas operativos (por ejemplo, Windows), las bases de datos (por ejemplo, ASSET 4) y el *middleware* (por ejemplo, Android). En ellos, destacan compañías tales como Apple, Google, Microsoft, SAP AG, o EMC.

En segundo lugar, el *software* de la empresa, el cual automatiza los procesos de negocio en finanzas, producción, logística, ventas y marketing, donde destacan Oracle, Sage e Infor.

Por último, el *software* de seguridad con Manage Engine Symantec, Trend Micro y Kaspersky.

Sin embargo, a su vez existe una subindustria, el *software* libre creado sin ánimo de lucro, y que cada vez tiene más adeptos, haciendo más accesible el desarrollo del *software* a modo individual. También se destaca la independencia de los intereses comerciales y en su importancia como base de un enorme ecosistema de varios cientos distribuciones derivados, que incluye las más populares distribuciones de GNU / Linux de hoy. (Zacchiroli, 2011).

### 3.3. Tendencias actuales de la industria del *software*

Algunos autores identifican cuatro tendencias que denominan actualmente la industria del *software* (Tucker et al., 2011).

En primer lugar, la innovación y el desarrollo de nuevos productos acompañados de nuevas aplicaciones, servicios y plataformas en las que se utiliza el *software*. Las aplicaciones SaaS para el mercado empresarial se han expandido rápidamente, porque muchas empresas los perciben como ofrecer un costo total de propiedad más bajo, lo que podría reducir la necesidad de personal de TI (*Technology Integration*) para gestionar las aplicaciones. Así mismo, es necesario destacar Platform as a Service, - PaaS-, una variante de SaaS mediante la cual los clientes o empresas de *software* de terceros añaden sus propias aplicaciones para una plataforma SaaS asociada.

En segundo lugar, el desarrollo de aplicaciones móviles motivado por las nuevas plataformas y dispositivos baratos fácilmente disponibles, como los teléfonos inteligentes y Tablet PC. Las empresas y los consumidores están recurriendo a aplicaciones móviles debido a su facilidad de uso y la constante disponibilidad. Incluso las grandes organizaciones de TI están integrando aplicaciones móviles para aumentar la productividad en sus compañías.

En tercer lugar, las fusiones y adquisiciones. Las compañías pequeñas desaparecen o se fusionan con empresas más grandes, debido a su pequeña base instalada de usuarios, bajo soporte técnico y el elevado coste de desarrollo que conlleva el *software*. Dichas fusiones y adquisiciones están impulsando el cambio en la industria, ya que de este modo ofrecen soluciones completas a los problemas de sus clientes.

Por último, la expansión y la proliferación de nuevos mercados, haciendo especial énfasis en el mercado asiático, el cual atrae la atención de las más importantes empresas de *software*. Compañías de redes sociales a su vez comparten ese interés. Con un importante desarrollo de juegos estandarizados para las mismas.

### **3.4. Amenazas en la industria del *software*.**

Expertos en la industria señalan cuatro retos a los que la industria del *software*, en constante evolución, debe enfrentarse.

- Los competidores dentro de las empresas de *software* abierto o libre, en contraposición con el *software* tradicional.
- La obsolescencia del producto, que está en constante variación, tanto por la exigencia del cliente como por las tecnologías de hardware que día tras día revolucionan la manera de entender el *software*.
- La constante piratería informática, con los consiguientes problemas de privacidad de los clientes así como riesgos de seguridad constantes con más y mejores herramientas de intrusión.
- La baja propiedad intelectual en temas referentes a *software*. Problema aún mayor en países emergentes, donde la legislación es insuficiente.

Esta son algunas de las principales amenazas que se ciernen sobre la industria del *software*. (Tucker et al., 2011).

En conclusión, sobre la industria del *software* se ciernen algunos nubarrones que la convierten en una industria en constante evolución e innovación. Tras un breve repaso por la industria del *software*, ahora pasamos a ver su relación con I+D.

#### **4. Factores determinantes de la inversión en I+D**

Numerosos estudios teóricos evidencian la existencia de una gran variedad de factores determinantes de la innovación empresarial, derivados tanto del ámbito externo de la organización como del ámbito interno (Díaz Martín, 1996). Sin embargo, habiendo un debate abierto respecto de los indicadores más apropiados para la medición del nivel de I+D de una empresa (González Bravo y Pargas Carmona, 2010).

En este sentido, cabe preguntarse cuáles son los factores determinantes que empujan a la empresas europeas de *software* a invertir en I+D.

El cuadro 1 recoge los principales estudios que señalan los factores determinantes de la inversión en I+D que se analizarán en el presente trabajo.

**Cuadro 1. Algunos estudios sobre los factores determinantes de la inversión en I+D**

<b>FACTOR</b>	<b>AUTORES</b>	<b>HIPÓTESIS</b>	<b>VARIABLE DETERMINANTE DE LA INTENSIDAD EN I+D.</b>	<b>CONTRASTE DE LA HIPÓTESIS.</b>
TAMAÑO	Schumpeter (1950); Cohen y Klepper (1996); Gumbau (1997); Tsai y Wang (2005)	H. Shumpeteriana marca I. A mayor tamaño más intensidad de la inversión en I+D.	Número de empleados.	Positivo.
	Scherer y Ross (1990) y Veciana Verges (2007)	Hay relación inversa entre el tamaño e intensidad de la inversión en I+D.	Número de empleados.	Negativo
	Cohen y Levin (1989)	El tamaño determina la intensidad de la inversión en I+D según el sector.	Número de empleados.	Positivo.
CONCENTRACIÓN DE LA PROPIEDAD	Wang (2009); Hall (2002); Moreno y Ortega-Argilés(2004)	A mayor concentración de la propiedad, menor intensidad de inversión en I+D.	Escala de concentración de la propiedad accionarial.	Positivo.
DESEMPEÑO ( <i>performance</i> )	Grabowski (1968); González y Pargas (2010)	A mayor desempeño ( <i>performance</i> ), mayor inversión en I+D.	Productividad	Positivo.
ENDEUDAMIENTO	Bhagat y Welch (1995); Bah y Dumontier (2001)	A mayor endeudamiento, menor intensidad de la inversión en I+D.	Endeudamiento	Negativo.

Fuente: Elaboración propia

## 4.1. Tamaño

La revisión de la literatura de las dos últimas décadas concerniente a la intensidad que realizan las empresas una inversión en I+D señala al tamaño de la empresa, medido a partir del número de empleados de la misma, como una de las variables explicativas de la innovación en los estudios empíricos.

No obstante, en este sentido, pueden diferenciarse tres tipos de estudios: (a) los que apuestan por la hipótesis Schumpeteriana; (b) los contrarios a la misma y (c) los que señalan la dependencia del sector.

En primer lugar, el protagonismo de esta variable se la debemos al trabajo de Schumpeter (1942), el cual señalaba la importancia del rol del tamaño tanto en la acumulación de conocimiento como en el crecimiento empresarial y económico, debido a la propia naturaleza del sistema capitalista. En este sentido, Schumpeter apuesta por que las grandes empresas generarán una ventaja competitiva en términos de innovación respecto a las pequeñas y medianas empresas, es decir, las primeras tienen más capacidad para invertir en I+D que las segundas. También denominado hipótesis schumpeteriana marca I.

Numerosos estudios confirman esta hipótesis, hallando una relación cuasi perfecta entre las empresas de gran tamaño y la inversión en I+D. Entre ellos, Cohen y Klepper (1996), Gumbau (1997) y Tsai y Wang (2005).

En segundo lugar, se encuentran aquellos estudios que señalan el papel de las pequeñas en la economía puesto que son consideradas como el motor de la innovación en determinadas industrias, a pesar de que formalmente su esfuerzo en I+D es significativamente menor. En esta línea se hace referencia a los trabajos realizados por Acs y Audretsch (1991), Cohen (1995), Mansfield (1962), Scherer (1984), Scherer y Ross (1990) y Cohen y Levintghal (1989), evidencian que existen pocos argumentos para apoyar la hipótesis schumpeteriana; o, dicho de otra manera, “no había ninguna evidencia que apoyara la existencia de resultados crecientes en función del tamaño de la empresa en cuanto a la generación de «outputs» de innovación” (Veciana Verges, 2007). En esta línea, los autores citados concluyen que las empresas pequeñas no están necesariamente en desventaja innovadora frente a las grandes, y añade que la habilidad para innovar permite a las pequeñas empresas compensar las desventajas de escala en comparación con las grandes empresas. Estas desventajas se solventan debido a la flexibilidad burocrática de las pequeñas y medianas empresas según Scherer y Ross (1990).

Por último, Cohen y Levinthal (1989) señalan que la decisión de la inversión en I+D viene determinada por las características diferenciales de la industria, es decir, por el sector y mercado analizado. Es más, afirman que si no se consideran estas variables caeríamos en un error de sesgo en el estudio, siendo la presión de la demanda, la oportunidad tecnológica y las condiciones de apropiabilidad características claves de la innovación tecnológica (Díaz Martín, 1996: 147).

Se concluye, por lo tanto, que a pesar de la tendencia de la literatura a apoyar la hipótesis schumpeteriana, se hallan así mismo numerosos estudios que ponen en tela de juicio la misma. En este sentido, Díaz Martín (1996), tal y como se muestra en el cuadro 2, realiza un análisis comparativo entre las ventajas y desventajas de las grandes y pequeñas empresas en el ámbito de la innovación.

**Cuadro 2. Análisis comparativo de dimensión empresarial / innovación tecnológica**

	<b>Empresas grandes</b>	<b>Empresas pequeñas</b>
<b>Ventajas innovación tecnológica</b>	Mayor volumen ventas / mejor reparto costes de innovación Diversificar riesgos Capacidad organización / tratamiento información Relación tamaño / cuota mercado Personal especializado / servicios técnicos avanzados Mejor acceso exterior Mayores posibilidades. Proyectos costosos	Flexibilidad / adaptación mercado Menor requerimiento burocrático Sistemas comunicación informal Rápida resolución de problemas Mayor colaboración Dirección / empleados Especialización productiva Innovar productos, segmentos despreciados grandes Existencia programas públicos de apoyo
<b>Desventajas innovación tecnológica</b>	Falta Flexibilidad / adaptación mercado Exceso burocracia Problemas comunicación Actitud negativa Dirección / empleados Desprecio segmentos específicos	Limitación de recursos Estructuras de gestión poco profesionalizadas Escaso control canales comercialización Dificultades patentes / regulaciones estatales

Fuente: Díaz Martín (1996: 148)

### 4.3. Concentración de la propiedad

Parece evidente que uno de los factores más influyentes en la determinación de la intensidad de la inversión es la forma de realizar la toma de decisiones dentro de la empresa; por lo tanto, cómo está distribuida la participación de la propiedad de la compañía (Suriñach, 2005). Actualmente hay varias corrientes de opinión dentro de la academia:

- a) La estructura de la propiedad es independiente en lo referente a la intensidad de inversión en I+D
- b) Cuanta menor concentración de propiedad hay una mayor inversión en I+D.
- c) A mayor concentración de la propiedad, más intensidad en la inversión en I+D.

Algunos de los trabajos como Acs e Isberg, (1991) llegaron a la conclusión de que la estructura de la propiedad empresarial es independiente a la hora de realizar una mayor inversión I+D.

Una segunda corriente de autores concluyen que la concentración de la propiedad aleja a la empresa de mayores beneficios y una mayor especialización por medio de una mayor inversión en I+D (Snell y Hill, 1988).

Dentro de esta corriente aunque con importantes matices cabe destacar Beyer y Kraft (2007) muestran que la concentración de la propiedad provoca un aumento de la intensidad de la inversión en I+D, aunque no de una manera lineal, sino que en forma de “U invertida”, con lo cual el momento óptimo en la concentración de la propiedad se da entre los percentiles 25 y 75 de menor a mayor concentración de la propiedad. Conviene una concentración elevada de la propiedad para una mayor intensidad en la inversión en I+D, evitando consiguientemente ambos extremos de la escala de concentración.

Del mismo modo, una elevada concentración de la propiedad causa problemas a la hora de la asunción de políticas más arriesgadas. Un incremento en la participación de los propietarios en los puestos de gestión, reducirá la probabilidad tanto de adoptar proyectos de I+D como de formalizar el resultado de innovación en el registro de patentes y modelos de utilidad (Moreno y Ortega-Artiles, 2007).

Por último, hay una tercera corriente en la academia que afirma que el mecanismo de concentración de la propiedad en pocas manos provoca la alineación de intereses del equipo directivo con los accionistas de la empresa, traduciéndose en un mecanismo que reduce el problema de agencia entre accionistas y directivos.



#### **4.4. Endeudamiento**

Hay numerosos estudios que hacen referencia a la capacidad de una empresa para realizar una inversión en I+D y sus factores determinantes. Y sin duda uno de los más importantes es la capacidad de endeudamiento de una compañía.

Existe un amplio consenso en los autores consultados que sostienen que la estructura de capital de una compañía es uno de los factores más determinantes a la hora de aumentar o disminuir la intensidad de la inversión en I+D (Marra, 2007). Trabajos como el de Gosh (2012) señalan que una empresa muy endeudada baja su intensidad de la inversión en I+D. Así, se presentan estudios que ofrecen el resultado inverso; esto es, una empresa con unos buenos ratios de solvencia provoca una mayor inversión en I+D (Pindado et al., 2011).

Es decir, “existe una relación inversa entre los niveles de deuda y los gastos de I+D” Welch y Bhagat (1995), debido a la evidente aversión de cualquier prestamista al riesgo derivado de la intensidad de la inversión en I+D y la consiguiente incertidumbre de conseguir retornar dicha inversión Bah y Dumontier (2001) la deuda, por lo tanto, tiene una desventaja estratégica en el mercado de la innovación (Aghion et al., 2004). Dado que las empresas con un elevado endeudamiento se encuentran con problemas en ofrecer garantías a los eventuales prestamistas de deuda, le resulta más complicado obtener nueva financiación para realizar proyectos, y con ello afrontar un aumento de los gastos en I+D.

#### **4.5. Desempeño**

Cuantiosos estudios sobre el desempeño avalan la consideración de esta variable como determinante de I+D. Griliches (1979) fue el primero en presentar el debate sobre la innovación como output o *input* que nos introduce directamente en esta cuestión.

Varios trabajos confirman la relación positiva entre el desempeño y la intensidad de la inversión en I+D (Mueller y Grabowski, 1978; Griliches y Mairesse, 1984; y Griliches, 1986). Sin embargo, cabe destacar la existencia de una corriente minoritaria que afirma, que dicha asociación no es directa Bean (1995), sino que debemos considerar la interposición de variables intervinientes como las ganancias que una empresa experimenta en productos y procesos o el tamaño de la misma (Cohen y Keppler, 1996) que modifican o alteran tal relación.

Sin embargo, parece evidente que al utilizar variables de desempeño existe un considerable efecto de retardo en la rentabilidad de las actividades de I + D, ya que una empresa estima el conjunto de rentabilidad de sus proyectos en curso dependiendo de sus resultados anteriores. Es por ello que probablemente uno de los determinantes más importantes en lo referido a las actividades de I+D es un indicador de la productividad de I + D de cada empresa durante un período anterior (Grabowski, 1968).

## **CAPÍTULO II. DISEÑO DE LA INVESTIGACIÓN EMPÍRICA**

### **1. Hipótesis.**

Las hipótesis han sido formuladas siguiendo dos premisas.

Por un lado, tal y como se ha detallado en el capítulo I, la existencia de una amplia literatura dota de consistencia su introducción en el modelo de investigación para su estudio. Por otro, los datos requeridos para llevar a cabo estas hipótesis se encuentran dentro de las bases de datos, anteriormente explicadas, proporcionando datos suficientes para la constatación de las siguientes hipótesis.

#### **Determinante tamaño.**

Perseguimos conocer los determinantes de la innovación. Por ello, no se puede dejar de mencionar uno de los primeros trabajos vinculados con nuestro objeto de investigación. Se hace referencia a Schumpeter (1942) quien señala que la gran empresa es el motor principal de la economía del progreso.

Como sabemos que la inversión en I+D conlleva una serie de costes fijos elevados, parece evidente que los mismos serán mejor afrontados por empresas grandes. Un ejemplo de ello son las economías de escala de innovación. De esta manera, las grandes empresas tienen un mejor acceso a la financiación externa Hall (2002), con lo cual disminuyen los costes de financiación y diversifican sus proyectos de I+D, así, los riesgos de dichas inversiones se reparten.

El número de empleados de la empresa o el tamaño de la empresa se ha configurado como una de las variables explicativas más importantes, de la cual hay una amplia literatura (López Mielgo et al., 2005). De esto se desprende que una empresa que tenga con un mayor número de trabajadores contará así mismo con una serie de recursos financieros, de especialización que le dotarán de una ventaja competitiva en la decisión de invertir en I+D.

No obstante, se encuentran aquellos estudios que señalan el papel de las pequeñas en la economía puesto que son consideradas como el motor de la innovación en determinadas industrias, a pesar de que formalmente su esfuerzo en I+D es significativamente menor (Costa et al., 2000: 4-5). En esta línea se hace referencia a los trabajos realizados por Acs y Audretsch (1993); Mansfield (1962), Scherer (1984), Scherer y Ross (1990), Cohen y Levinthal (1989).

Existen pocos argumentos para apoyar la hipótesis schumpeteriana; o, dicho de otra manera, “no había ninguna evidencia que apoyara la existencia de resultados crecientes en función del tamaño de la empresa en cuanto a la generación de «outputs» de innovación” (Veciana Verges, 2007: 108). En esta línea, los autores citados concluyen que las empresas pequeñas no están necesariamente en desventaja innovadora frente a las grandes, y añade que la habilidad para innovar permite a las pequeñas empresas

compensar las desventajas de escala en comparación con las grandes empresas. Estas desventajas se solventan debido a la flexibilidad burocrática de las pequeñas y medianas empresas según Scherer y Ross (1990).

Por último, y aunque en los trabajos empíricos se han obtenido resultados diversos, parecen predominar argumentos a favor de una relación positiva sobre tamaño e intensidad de innovación (López Mielgo et al., 2005). Es decir, hay una correlación positiva entre el número de trabajadores de la empresa, o tamaño, y la intensidad de la inversión en I+D. Esto es, al aumentar la primera aumenta también la segunda. Por lo tanto:

**H1: Cuanto mayor sea el tamaño de la empresa, mayor será la intensidad de la inversión en I+D en las empresas de software (Hipótesis 1)**

### **Determinante desempeño.**

Una de las variables más importantes a la hora de determinar la intensidad de la inversión en I+D de las empresas, es sin duda el desempeño, según la bibliografía consultada Kraens et al (2009), se observa una relación positiva entre los beneficios o retornos y un aumento de la intensidad de la inversión en investigación de las empresas tecnológicas. Esto se refleja con mayor claridad a largo plazo puesto que a corto plazo los resultados son menos concluyentes.

Cabe destacar por otra parte, existe cierta controversia a la hora de elegir cuál es el mejor indicador del desempeño empresarial, ya que Griliches (1995) realizó una revisión sobre diversos trabajos que a nivel microeconómico han relacionado la productividad de las empresas con variables explicativas de la inversión en I+D. Para este autor, una de las principales limitaciones de este enfoque es que las medidas de rendimiento empresarial utilizadas no reflejan de manera adecuada el resultado del I+D.

Es por esta razón que González y Pargas (2010) creen conveniente la utilización de la variable productividad para medir el desempeño empresarial. Así mismo, autores tanto nacionales, como Fluvía (1990), o internacionales, como Griliches y Mairesse (1984), analizan la relación entre crecimiento de la productividad e intensidad en la inversión en I+D, alegan que tal crecimiento determina positivamente el aumento en la intensidad de la inversión en I+D.

Por último, debemos considerar, como ya señalamos en el apartado anterior, que este tipo de variables son susceptibles de ser retardadas puesto que sus rendimientos se producen *a posteriori* (Grabowski, 1968). En función de lo que precede:

**H2: Cuanto mayor sea el desempeño (performance), mayor será la intensidad en la inversión en I+D en empresas de software (Hipótesis 2).**

### **Determinante concentración de la propiedad.**

Del mismo modo, y como ya hemos subrayado en el Capítulo I, existe un gran debate sobre cómo afecta la estructura del capital de la compañía a la intensidad de inversión en I+D. Aunque parece evidente, tal y como señala Hall (2002), que la estructura del capital y la concentración del mismo es un factor determinante en la intensidad de la inversión en I+D.

Pese a la polémica existente en esta materia, podemos tomar como referencia a Moreno y Ortega-Argilés (2004) que explican empíricamente la existencia de una relación inversa entre la concentración de la propiedad y la intensidad de la inversión en I+D. En este sentido, los resultados abogan por la siguiente conclusión: cuanto más diluida está la concentración de la propiedad, es mayor la intensidad de la inversión en I+D de la compañía. Y al revés. Por ello, cabe establecer la siguiente hipótesis, que para más fácil comprensión y claridad enunciamos así:

**H3: Cuanto mayor es la concentración de la propiedad, menor será la intensidad de la inversión de la empresa en I+D. (Hipótesis 3).**

### **Determinante endeudamiento.**

Numerosos autores en estudios precedentes miden el endeudamiento o apalancamiento financiero de la compañía como un factor determinante negativo a la hora de aumentar la intensidad de la inversión en I+D. Marra (2007), Martinsson (2009) o Gosh (2012) demuestran empíricamente que una estructura de capital con un elevado endeudamiento provoca una menor intensidad de la inversión en I+D.

A su vez y de manera inversamente proporcional, la solvencia de las compañías determina un aumento en la intensidad de la inversión de las mismas en I+D. Pindado et al. (2011) halló una correlación positiva entre la cantidad de cash flow libre en la estructura financiera de la empresa y el aumento en la intensidad de la inversión en I+D. Por esto:

**H4: Cuanto menor es el endeudamiento de una empresa, mayor será la intensidad de la inversión en I+D de la compañía. (Hipótesis 4).**

A continuación, se presentan en un cuadro resumen las hipótesis que la investigación contrastará en su desarrollo. Para ello, puede observarse el cuadro 3, el cual recoge de modo esquemático las relaciones esperadas entre la variable dependiente de estudio, la intensidad de la inversión en I+D y las variables independientes o explicativas que se toman en consideración.

**Cuadro 3. Hipótesis a contrastar y resultados esperados**

<b>FACTOR</b>	<b>HIPÓTESIS</b>	<b>VARIABLE</b>	<b>EFFECTOS ESPERADOS</b>	<b>ACEPTACIÓN O RECHAZO DE HIPÓTESIS</b>
Tamaño	H1: A mayor tamaño mayor intensidad de la inversión en I+D.	Logaritmo del número de empleados.	Signo positivo.	Aceptada
Concentración de la propiedad	H2: A mayor concentración de la propiedad, menor intensidad de la inversión en I+D.	Escala de concentración de la propiedad	Signo negativo.	Aceptada
Desempeño	H3: A mayor desempeño, mayor intensidad de la inversión en I+D.	Productividad.	Signo positivo.	Aceptada
Endeudamiento	H4: A mayor endeudamiento, menor intensidad de la inversión en I+D.	Pasivos totales/ Fondos propios.	Signo negativo.	Aceptada

Fuente: Elaboración propia.

## 2. Variables utilizadas.

A continuación, en el cuadro 4 se presenta el conjunto de las variables e indicadores que forjarán la estructura esencial del trabajo. Adicionalmente a las variables ya mencionadas al hilo de la presentación de las hipótesis, dicha tabla 4 refleja también la variable de control sectorial habitual en las estimaciones econométricas similares a las que aquí se van a realizar, si bien en el presente caso conviene aclarar que dicha variable es en realidad subsectorial, es decir, subsectores dentro del sector global analizado (el de *software* y los servicios informáticos). Además, se observará en su momento que la estimación econométrica final va también controlada por el efecto del tiempo (variable año) y del subsector (variable nace), así como por el efecto de la coyuntura temporal sobre el correspondiente subsector (es decir, la interacción de variable año y variable nace).

**Cuadro 4. Variables independientes y de control.**

VARIABLES		INDICADORES	DEFINICION
Variable dependiente	Intensidad de la inversión en I+D.	Gasto anual en I+D/ número de empleados de la empresa. Denominación: rdxemp	La proporción de inversión en I+D que realiza una empresa en relación al número de trabajadores de la misma.
Variables independientes o explicativas	Tamaño	Logaritmo neperiano del número de empleados. Denominación: logemp	Número de personas que trabajan para la compañía a final de cada ejercicio.
	Productividad	Ratio productividad por empleado (ventas netas / número de empleados). Denominación: productividad	Mide la productividad en términos de ventas netas generadas por cada empleado.
	Endeudamiento	Pasivos totales/Fondos propios. Denominación: endeudamiento	Mide la relación existente entre el importe de los fondos propios de una empresa con relación a las deudas.
	Concentración de la propiedad	Indicador de concentración de la propiedad de la base de datos <i>Amadeus</i> . Denominación: prop	Escala alfabética de la A a la U, que clasifica la concentración de las acciones de una compañía. Comenzando de menos concentrada, A, a más U.

Variable de control	Subsector	NACE (statistical <u>N</u> omenclature of economic <u>A</u> ctivities in the <u>E</u> uropean <u>C</u> ommunity). Denominación: nace	El número que clasifica actividades en que las empresas desarrollan su actividad.
---------------------	-----------	---	---

Fuente: Elaboración propia.

### 3. Fuentes de datos

Las bases de datos son dos, *Economics of Industrial Research and Innovation* (IRI) y *Amadeus*.

La primera, *Economics of Industrial Research and Innovation*, procede de una acción científica desarrollada por el Institute for Prospective Technological Studies (IPTS). Es de libre acceso y está disponible con datos para cada año en la siguiente dirección electrónica: <http://iri.jrc.ec.europa.eu/scoreboard.html>. Contiene información detallada de las 1000 empresas de mayor inversión en I+D a nivel mundial. Dentro de ella se utilizan los referentes al sector del *software* en la Unión Europea. Esta base aporta una clasificación de la inversión en I+D que ayuda, a su vez, a elegir una muestra de empresas con información clara acerca tanto de su inversión en I+D como de sus ventas netas y de su capitalización bursátil para cada ejercicio económico.

La segunda, *Amadeus*, elaborada por el Bureau Van Dijk (Bruselas) y con acceso adquirido por la Universidad de Valladolid, nos ofrece información más detallada de las empresas anteriormente seleccionadas dentro de IRI. Estos datos más específicos de las empresas abarcan desde los estados contables consolidados hasta los ratios estructurales, de rentabilidad y operativos, además de la composición del accionariado y la mayor o menor concentración de su propiedad.

Tras la elección de las bases de datos, hemos pasado a verificar que los datos en ambas bases son susceptibles de ser combinados, procesándolos a efectos de la adecuación de sus formatos y eliminando toda empresa en la que los datos no tuvieran una concordancia total entre las cifras de ambas bases. Todo ello para obtener una mayor y mejor calidad de información de las empresas y con ello estudiar cuáles son determinantes últimos que llevan a las empresas a invertir en I+D.

Cabe destacar la calidad y facilidad de acceso de la información, además de la posibilidad de combinar datos, han sido determinantes en la elección final de estas bases. Otras fuentes de información disponibles no aportaban una información adicional susceptible de ser útil dentro del estudio, puesto que acarrearían la pérdida de datos sobre muchas compañías.

Finalmente, explicado que utilizaremos las bases de datos *Amadeus* y *Economics of Industrial Research and Innovation*, pasamos a detallar la muestra de empresas seleccionada dentro de ellas y los datos necesarios manejados para llevar a cabo la investigación empírica. Los datos referentes a unidades monetarias están expresados en millones de euros aunque no se indique explícitamente.



#### 4. Muestra de empresas.

La muestra utilizada está formada por un panel de datos incompleto, es decir, con algún dato ausente para una determinada empresa en un cierto año. Contiene 49 empresas europeas cotizadas en Bolsa pertenecientes a los sectores de *software* y servicios informáticos, de las cuales existe información relevante para el estudio durante el periodo que abarca los años entre 2004 y 2010 ambos inclusive. La lista concreta de empresas se refleja en el cuadro 5.

**Cuadro 5: Relación de empresas seleccionadas para el estudio.**

<b>EMPRESA DE SOFTWARE</b>	<b>PAÍS</b>
ALDATA SOLUTION OYJ	FINLANDIA
ALTERIAN LIMITED	REINO UNIDO
ATARI	FRANCIA
AUTONOMY CORPORATION LIMITED	REINO UNIDO
AVANQUEST SOFTWARE	FRANCIA
AVEVA GROUP PLC	REINO UNIDO
BASWARE OYJ	FINLANDIA
BOND INTERNATIONAL SOFTWARE PLC	REINO UNIDO
COMPTTEL OYJ	FINLANDIA
DELCAM PLC	REINO UNIDO
ENEA AB	SUECIA
ESI GROUP	FRANCIA
FIDESSA GROUP PLC	REINO UNIDO
F-SECURE OYJ	FINLANDIA
INDUSTRIAL AND FINANCIAL SYSTEMS, IFS AB	SUECIA
INVENSYS	REINO UNIDO
KEWILL LIMITED	REINO UNIDO
LECTRA	FRANCIA
NEMETSCHEK AKTIENGESELLSCHAFT	ALEMANIA
NET ENTERTAINMENT	SUECIA
P & I PERSONAL & INFORMATIK AG	ALEMANIA
READSOFT AB	SUECIA
RM	REINO UNIDO
SAP AG	ALEMANIA
SDL PLC	REINO UNIDO
SIMCORP A/S	DINAMARCA
SOPHOS LIMITED	REINO UNIDO
TEKLA OYJ	REINO UNIDO

UNIT4 N.V.	HOLANDA
USU SOFTWARE AG	ALEMANIA
STATPRO GROUP PLC	REINO UNIDO
SINOSOFT TECHNOLOGY LIMITED	REINO UNIDO
LINEDATA SERVICES	FRANCIA
INVISION SOFTWARE AKTIENGESELLSCHAFT	ALEMANIA
MEVIS MEDICAL SOLUTIONS AG	ALEMANIA
INTERSHOP COMMUNICATIONS AKTIENGESELLSCHAFT	ALEMANIA
KASPERSKY LABS	REINO UNIDO
INDRA SISTEMAS, SOCIEDAD ANONIMA	ESPAÑA
SOPRA GROUP	FRANCIA
TIETO OYJ	FINLANDIA
CEGEDIM	FRANCIA
COMPUGROUP MEDICAL	ALEMANIA
CS COMMUNICATION & SYSTEMES CS	FRANCIA
LOGICA PLC	REINO UNIDO
SWORD GROUP SE	FRANCIA
ENGINEERING - INGEGNERIA INFORMATICA - S.P.A.	ITALIA
NOVABASE	PORTUGAL
GFI INFORMATIQUE	FRANCIA
MICROGEN PLC	REINO UNIDO

La referencia de la base de datos específicamente utilizada para reunir la información referente a cada una de las variables manejadas en el presente estudio se recoge en el cuadro 6.

**Cuadro 6. Bases de datos de todas las variables.**

<b>Variables</b>	<b>Base de datos utilizada</b>
NACE (nace)	<i>Amadeus.</i>
Concentración empresarial (prop)	<i>Amadeus.</i>
Productividad (product)	<i>Amadeus.</i>
Ratio endeudamiento (endeuda)	<i>Amadeus.</i>
Intensidad de la inversión en I+D (rdxemp)	I.R.I. (*)
Número de mpleados (logemp)	I.R.I. (*)

Elaboración propia. (\*) I.R.I (Economics of Industrial Research and Innovation)

## 5. Modelos y métodos de estimación

Se trata de un trabajo de índole estadístico con finalidad empírico explicativa, o de otro modo, se utilizan técnicas estadísticas para describir las relaciones entre variables. Este método es preciso para los objetivos de la investigación; así, el análisis desde el punto de vista econométrico, hay varios modelos que podríamos estimar, así como diferentes métodos aplicables.

Seguidamente, vamos a realizar un resumen teórico de dichos métodos y modelos, tomando como referencia a Greene (2002) y a Aparicio y Márquez (2005).

El enfoque más simple sería la regresión agrupada por mínimos cuadrados ordinarios, ya que consiste en omitir las dimensiones del espacio y el tiempo de los datos agrupados y calcular simplemente la regresión por Mínimos Cuadrados Ordinarios (MCO). El problema es que es posible que los residuos no sean independientes de las observaciones por lo que MCO estará sesgado. Por lo tanto, no se usará este método, debido a que existen técnicas más recientes que resultan más apropiadas en caso de datos de panel y que se detallarán en el presente epígrafe.

A continuación, se explican los métodos concretos previstos y utilizados para el presente trabajo, si bien uno de ellos finalmente no procede aplicarlo, como en su momento se explicará.

### 5.1. Modelo de efectos fijos (fixed effects)

Este modelo supone que las diferencias entre individuos no son aleatorias, sino constantes, lo cual se traduce en la siguiente ecuación:

$$Y_{it} = \alpha_i + \beta_k X_{kit} + e_{it}$$

Donde  $\alpha_i$  es un valor fijo, diferente para cada individuo pero constante –no aleatorio– a lo largo del tiempo. De este modo, el término de error solamente refleja la perturbación asociada a cada observación, ya que se supone que el término constante no es una variable aleatoria y, por lo tanto, no existe perturbación asociada al individuo. Para estimar este modelo, se supone que el error cumple todos los supuestos clásicos y que los regresores estén incorrelacionados con el error de la observación ( $E(X_{it} * e_{it}) = 0$ ), por lo cual el Estimador Lineal Insesgado Óptimo (ELIO) de  $\beta$  sería el estimador MCO incorporando N-1 variables ficticias, lo cual podría expresarse matricialmente del siguiente modo:

$$\begin{pmatrix} \hat{\beta} \\ \hat{\mu} \end{pmatrix} = (Z'Z)^{-1} Z'Y$$

Donde Z es una matriz NT x (K+(N-1)) con todas las K variables explicativas (X) y las N-1 variables ficticias. Si N o T tienden a infinito, el estimador de  $\beta$  es consistente.

## 5.2. Modelo de efectos aleatorios (random effects)

El modelo de efectos aleatorios permite suponer que cada individuo está asociado a una constante diferente dentro de la ecuación. Partamos del modelo:

$$Y_{it} = \alpha_i + \beta_k X_{kit} + e_{it}$$

Dónde:

$Y_{it}$  es el valor de la variable dependiente para el individuo  $i$  en el periodo  $t$ .

$\alpha_i$  es el valor del término independiente.

$\beta_k$  es el coeficiente asociado al regresor  $k$ .

$X_{kit}$  es el regresor  $k$  para el individuo  $i$  en el periodo  $t$ .

$e_{it}$  es el error de la observación o perturbación aleatoria para el individuo  $i$  en el periodo  $t$ .

$i$  significa la  $i$ -ésima unidad transversal (individuo) y  $t$  el tiempo (año).

Consideremos que  $\alpha_i = \alpha + u_i$ . Es decir, suponemos que  $\alpha_i$  es una variable aleatoria diferente para cada individuo, con un valor medio y una desviación típica aleatoria. Sustituyendo en la ecuación anterior, se obtiene la expresión del modelo de efectos aleatorios, tal y como se muestra a continuación:

$$Y_{it} = \alpha + \beta_k X_{kit} + u_i + e_{it}$$

De este modo,  $u_i$  forma parte del término de error, que ahora tiene dos componentes uno específico del individuo  $i$  ( $u_i$ ) y otro de la observación ( $e_{it}$ ).

Las condiciones de ortogonalidad del modelo de efectos aleatorios incluyen que los regresores estén incorrelacionados con el error de la observación ( $E(X_{it} * e_{it}) = 0$ ) y con el error derivado de cada grupo o individuo ( $E(X_{it} * u_i) = 0$ ).

El Estimador Lineal Insesgado Óptimo (ELIO) para  $\beta$  en el modelo de efectos aleatorios sería el estimador de Mínimos Cuadrados Generalizados (MCG), dado que el mencionado modelo tiene problemas de autocorrelación inherentes a su propio diseño. El modelo de MCG, en su versión matricial, sería el siguiente:

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \\ E(\mathbf{u}) &= \mathbf{0} \\ V(\mathbf{u}) &= E(\mathbf{u}\mathbf{u}') = \sigma^2 \boldsymbol{\Omega} (\delta \mathbf{V}) \end{aligned}$$

No obstante, la matriz de varianzas y covarianzas ( $\boldsymbol{\Omega}$ ) del modelo de MCG no es escalar, lo cual complica su estimación. Sin embargo, la matriz de varianzas y covarianzas ( $\boldsymbol{\Omega}$ ) se puede expresar como el producto de dos matrices:

$$\boldsymbol{\Omega} = \mathbf{P}\mathbf{P}'$$

En ese caso, el problema de no escalaridad de la matriz puede resolverse realizando las siguientes transformaciones:

$$\begin{aligned} \mathbf{P}^{-1}\boldsymbol{\Omega}(\mathbf{P}')^{-1} &= \mathbf{P}^{-1}\mathbf{P}\mathbf{P}'(\mathbf{P}')^{-1} = \mathbf{I}_T \\ (\mathbf{P}')^{-1}\mathbf{P}^{-1} &= \boldsymbol{\Omega}^{-1} \end{aligned}$$

Tras los ajustes anteriores, el modelo resultante tendría una matriz de varianzas y covarianzas escalar y matricialmente quedaría del siguiente modo:

$$\begin{aligned} \mathbf{y}^* &= \mathbf{X}^*\boldsymbol{\beta} + \mathbf{u}^* \\ E(\mathbf{u}^*) &= \mathbf{0} \\ V(\mathbf{u}^*) &= E(\mathbf{u}^*\mathbf{u}^{*\prime}) = \sigma^2 \mathbf{I}_T \end{aligned}$$

Dónde:

$$\mathbf{y}^* = \mathbf{P}^{-1}\mathbf{y}, \mathbf{X}^* = \mathbf{P}^{-1}\mathbf{X}, \mathbf{u}^* = \mathbf{P}^{-1}\mathbf{u}.$$

De tal manera que, al aplicar MCO a la transformación anterior, se obtiene el estimador MCG:

$$\hat{\boldsymbol{\beta}}_G = (\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X})^{-1} \mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{y}$$

El estimador de MCG es ELIO y eficiente.

### 5.3. Elección entre efectos fijos y aleatorios: test de Hausman

El test de Hausman es utilizado para elegir entre efectos fijos y aleatorios.

Dicho contraste compara un estimador consistente (el estimador de efectos fijos) con un estimador eficiente (el estimador de efectos aleatorios es eficiente y consistente). Es

obvio que, a ser posible, sería conveniente utilizar el estimador de efectos aleatorios en lugar del estimador de efectos fijos. No obstante, cuando los regresores y el error específico del individuo ( $u_i$ ) están correlacionados, el estimador de efectos aleatorios no es consistente y, para determinar la consistencia o no de dicho estimador, recurrimos al test de Hausman.

Hausman demostró que la diferencia entre los coeficientes de efectos fijos y aleatorios ( $\beta_{ef} - \beta_{ea}$ ) puede ser usada para probar la hipótesis nula de que y las variables regresoras no están correlacionadas ( $E(X_{it} * u_{it}) = 0$ ). Así pues, la Hipótesis nula ( $H_0$ ) de la prueba de es que los estimadores de efectos aleatorios y de efectos fijos no difieren sustancialmente. Si se rechaza la Hipótesis nula, los estimadores sí difieren, y la conclusión es efectos fijos es más conveniente que efectos aleatorios debido a que  $E(X_{it} * u_{it}) \neq 0$  y ello incumple las condiciones de efectos aleatorios, haciendo inconsistente su estimador. Si no podemos rechazar la hipótesis nula, no hay sesgo de qué preocuparnos por lo que  $E(X_{it} * u_{it}) = 0$  y se cumplen las condiciones que hacen consistente el estimador de efectos aleatorios, por lo que se elegiría este último modelo ya que su estimador es eficiente.

#### 5.4. Modelo de mínimos cuadrados generalizados factibles

La principal diferencia entre el estimador de MCG y el estimador de Mínimos Cuadrados Generalizados Factibles (MCGF) es que el estimador de MCGF reemplaza las varianzas por estimaciones consistentes de las mismas, por lo que partiendo de la ecuación, el estimador de  $\beta$  sería el siguiente:

$$\hat{\beta}_{GF} = (\mathbf{X}'\hat{\Omega}^{-1}\mathbf{X})^{-1}\mathbf{X}'\hat{\Omega}^{-1}\mathbf{y}$$

La forma de estimar la matriz de varianzas y covarianzas dependerá de si la causa de su no escalaridad es la heterocedasticidad y/o la autocorrelación.

El estimador MCGF no es ELIO, puesto que no es insesgado debido a que usa una estimación de la matriz de varianzas y covarianzas. Sin embargo, el estimador MCGF es consistente y, asintóticamente, más eficiente que el estimador de mínimos cuadrados ordinarios (MCO).

Conviene aclarar que, aunque la estimación por MCGF por defecto se asocia con efectos aleatorios, puede utilizarse con efectos fijos introduciendo en el modelo las variables que caracterizan al individuo  $t$  y al periodo de tiempo en el panel ( $i$  y  $t$ ). No obstante, en el presente trabajo, como se explicará más adelante, finalmente no se ha utilizado la estimación por MCGF.

## 5.5. Regresión con errores estándar corregidos por autocorrelación y heterocedasticidad (Prais-Winsten)

La transformación de Prais-Winsten es una variante del método de mínimos cuadrados generalizados factibles, específicamente diseñada por Prais y Winsten (1954) para resolver problemas de autocorrelación con un proceso estocástico AR(1). No obstante, el método implementado en STATA, también corrige heterocedasticidad, ya que calcula los errores estándar y la matriz de varianzas y covarianzas. Dicho método supone que las perturbaciones aleatorias presentan heterocedasticidad y autocorrelación, por lo que utiliza estimaciones consistentes que soporten ambos problemas.

Para ello, realiza una serie de transformaciones que se detallan a continuación.

Dado el siguiente modelo:

$$y_t = \alpha + X_t\beta + e_t,$$

En el que  $e_t$  es la perturbación aleatoria y suponiendo que existe autocorrelación siguiendo un proceso AR(1), debe verificarse que:

$$e_t = \rho e_{t-1} + e_t, \quad |\rho| < 1$$

Donde  $e_t$  es una perturbación aleatoria ruido blanco (esperanza nula, varianza constante y covarianza nula).

Para neutralizar el proceso que sigue esta correlación y obtener un modelo con ruido blanco, el procedimiento Prais-Winsten reproduce el proceso AR(1) en el modelo del siguiente modo:

$$y_t - \rho y_{t-1} = \alpha(1 - \rho) + \beta(X_t - \rho X_{t-1}) + e_t.$$

Siendo  $e_t$  un ruido blanco.

Posteriormente, se realiza otra transformación adicional introduciendo raíces que, para  $t=1$  sería la siguiente:

$$\sqrt{1 - \rho^2} y_1 = \alpha \sqrt{1 - \rho^2} + (\sqrt{1 - \rho^2} X_1) \beta + \sqrt{1 - \rho^2} e_1.$$

Finalmente, se procede a estimar ese modelo por el método de mínimos cuadrados que se desee.

Si traspasamos esto al modelo matricial, se obtendría el estimador de la regresión Prais-Winsten que es el siguiente:

$$\hat{\Theta} = (\mathbf{Z}'\Omega^{-1}\mathbf{Z})^{-1}(\mathbf{Z}'\Omega^{-1}\mathbf{Y}),$$

Dónde:

Z es una matriz de observaciones sobre la variable independiente ( $X_t, t = 1, 2, \dots, T$ ).

Y es el vector que recoge las observaciones sobre la variable dependiente.

$\Omega$  es la matriz de varianzas y covarianzas que se expresaría del siguiente modo:

$$\Omega = \begin{bmatrix} \frac{1}{1-\rho^2} & \frac{\rho}{1-\rho^2} & \frac{\rho^2}{1-\rho^2} & \dots & \frac{\rho^{T-1}}{1-\rho^2} \\ \frac{\rho}{1-\rho^2} & \frac{1}{1-\rho^2} & \frac{\rho}{1-\rho^2} & \dots & \frac{\rho^{T-2}}{1-\rho^2} \\ \frac{\rho^2}{1-\rho^2} & \frac{\rho}{1-\rho^2} & \frac{1}{1-\rho^2} & \dots & \frac{\rho^{T-3}}{1-\rho^2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{\rho^{T-1}}{1-\rho^2} & \frac{\rho^{T-2}}{1-\rho^2} & \frac{\rho^{T-3}}{1-\rho^2} & \dots & \frac{1}{1-\rho^2} \end{bmatrix}.$$

No obstante, dicha matriz se puede expresar como producto de dos matrices:

$$\Omega^{-1} = \mathbf{G}'\mathbf{G}$$

Donde G es la siguiente matriz:

$$\mathbf{G} = \begin{bmatrix} \sqrt{1-\rho^2} & 0 & 0 & \dots & 0 \\ -\rho & 1 & 0 & \dots & 0 \\ 0 & -\rho & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}.$$



## 6. Estrategia econométrica de la investigación

Tras la breve revisión de la teoría econométrica que se acaba de realizar, pasamos a especificar la estrategia econométrica concretamente adoptada, resumiendo los pasos que se han seguido y el detalle de los comandos empleados en el paquete estadístico STATA. Dichos pasos son los siguientes:

### a. Estadísticos descriptivos y correlaciones bivariadas.

```
summarize  
pwcorr variables dependiente e independientes, sig
```

### b. Estimación con efectos fijos y con efectos aleatorios, así como realización del test de Hausman para elegir entre ambos.

```
Xtreg variable dependiente variables independientes, fe (estimación con efectos fijos)  
estimates store fixed  
Xtreg variable dependiente variables independientes, re (estimación con efectos  
aleatorios)  
Hausman fised . , sigmamore
```

### c. Contrastes de autocorrelación y heterocedasticidad, concluyéndose que existen ambas. Esta fase se completa con la realización de un test de Hausman robusto (test de restricciones sobreidentificadas) para elegir entre efectos fijos y efectos aleatorios.

```
xtserial variable dependiente variables independientes (test de autocorrelacion válido  
para ambos modelos)  
Xtreg variable dependiente variables independientes, fe  
Xtreg variable dependiente variables independientes, re  
Predict pred, e  
Robvar pred, by(n_ent) (test de heterocedasticidad para efectos aleatorios (robvar))  
Xtoverid, robust cluster (n_ent) (test de restricciones sobreidentificadas).
```

### d. La posible estimación por mínimos cuadrados generalizados factibles (MCGF), finalmente no se realizó, tal como se explica en su momento.

### e. Regresión de panel con errores estándar corregidos por autocorrelación y heterocedasticidad (regresión Prais-Winsten).

```
xtpcse variable dependiente variables independientes, noconstant correlation(dw) hetonly
```

Utilizaremos este modelo, tomando como base los razonamientos de Beck y Katz (1995), quienes demostraron que los errores estándar corregidos para panel (PSCE) son más precisos que los estimadores de mínimos cuadrados generalizados factibles (MCGF).

## CAPÍTULO III. RESULTADOS EMPÍRICOS Y CONTRASTE DE HIPÓTESIS.

En este apartado podemos ver los principales resultados estadísticos y econométricos alcanzados a efectos del propósito de la presente investigación.

### 1. Estadísticos descriptivos.

En primer lugar, se realiza un estudio descriptivo las variables dependiente e independientes, cuyo resultado mostramos a continuación en el cuadro 7. En el caso de las variables productividad y endeudamiento, se presentan también los estadísticos de sus retardos de orden 1, por la razón que más adelante se explicará. Estos retardos se presentan en Stata como L1.productividad y L1.endeudamiento, donde L procede del término *lag*, retardo en inglés. La variable concentración de la propiedad, tal y como podemos ver, tiene un rango de números que, a efectos econométricos, va de 2 a 6, lo que se corresponde con una escala de un mayor a menor grado de concentración de la propiedad del capital de la empresa. O, visto desde la perspectiva inversa, de menor a mayor dilución del control ejercido por una persona o grupo empresarial.

**Cuadro 7: Principales estadísticos descriptivos de las variables de la muestra.**

Variable	Obs	Mean	Std. Dev.	Min	Max
numero	343	25.57143	14.58991	1	50
año	343	2007	2.002922	2004	2010
rdxemp	247	20989.12	17019.08	448.1072	98674.52
logemp	323	7.013277	1.502159	3.663562	10.85201
endeudamiento	323	2.300578	4.277543	-33.01419	50.94244
L1. endeudamiento	275	2.128258	3.545853	-33.01419	17.74938
productividad	297	140250.9	64997.66	9966.352	505463.6
L1. productividad	256	141775.5	65655.2	11612.93	505463.6
_Iprop_2	334	.0209581	.1434589	0	1
_Iprop_3	334	.1886228	.3917956	0	1
_Iprop_4	334	.0419162	.2006983	0	1
_Iprop_5	334	.242515	.4292471	0	1
_Iprop_6	334	.0209581	.1434589	0	1
_Inace_355	343	.0204082	.1415985	0	1
_Inace_642	343	.0204082	.1415985	0	1
_Inace_671	343	.1632653	.3701475	0	1
_Inace_672	343	.0204082	.1415985	0	1
_Inace_733	343	.0204082	.1415985	0	1
_Inace_737	343	.5714286	.4955946	0	1
_Inace_738	343	.0612245	.2400919	0	1
_Inace_874	343	.0204082	.1415985	0	1

## 2. Análisis de correlaciones.

Presentamos en el cuadro 8 la matriz de correlaciones de las variables manejadas en este estudio empírico. Aparece el valor en sí de la correspondiente correlación y, debajo de él, el p-valor (o valor de probabilidad) obtenido en el pertinente contraste en prueba bilateral para la hipótesis nula de correlación igual a cero.

Por razones de espacio, en el caso de las variables de productividad y endeudamiento, se presentan sólo las correlaciones en el caso de sus retardos de orden 1, que son los realmente utilizados más adelante como variables explicativas, según se explicará.

Se observan algunas correlaciones relativamente elevadas, las cuales habrán de tenerse muy presentes por si pudieran desembocar en problemas econométricos de multicolinealidad en estimaciones posteriores.

**Cuadro 8: Matriz de las correlaciones de las variables a estudiar**

	rdxemp	L.prod~d	L.ende~o	logemp	_Iprop_2	_Iprop_3	_Iprop_4	_Iprop_5	_Iprop_6
rdxemp	1.0000								
L.producti~d	0.6962	1.0000							
L.endeudam~o	-0.2857	-0.1152	1.0000						
logemp	-0.4564	-0.2309	0.0880	1.0000					
_Iprop_2	0.0430	.	-0.1691	-0.0067	1.0000				
_Iprop_3	0.1275	-0.0239	0.0319	0.0650	-0.0705	1.0000			
_Iprop_4	-0.0346	0.0027	-0.0212	-0.1327	-0.0306	-0.1008	1.0000		
_Iprop_5	0.0502	0.0366	0.1110	-0.0375	-0.0828	-0.2728	-0.1184	1.0000	
Iprop_6	0.4326	0.5598	0.0662	0.5030	0.1311	0.0000	0.0306		1.0000
	0.0740	0.3576	0.9080	0.8201	0.6967	0.1984	0.5773	0.1311	

## 3. Estimaciones con efectos aleatorios y efectos fijos. Test de Hausman.

Para alcanzar el objetivo del análisis, que persigue determinar qué factores son determinantes a la hora de impulsar una mayor intensidad de la inversión en I+D, ante todo se ha estimado una regresión lineal con datos de panel con el denominado modelo de efectos aleatorios y con el llamado modelo de efectos fijos. Además, se efectuó el test de Hausman, cuyo objetivo es determinar si las diferencias entre el estimador de efectos fijos y el estimador de efectos aleatorios son estadísticamente significativas. El referido test de Hausman se explicará con mayor detalle adelante.

A continuación, se presentan los resultados obtenidos de las estimaciones con efectos fijos y con efectos aleatorios, así como el test de Hausman. Véanse los cuadros 9, 10 y 11, respectivamente. Ahora bien, es obligado efectuar aquí una aclaración importante:

Como en la estimación de otros modelos similares, hay que tener en cuenta la eventual existencia de problemas de endogeneidad. En este caso concreto derivados de una posible simultaneidad que podría afectar a la relación entre la variable dependiente y las variables explicativas productividad y endeudamiento. Y así parece ocurrir. Para llegar a esta conclusión, ante todo y a través del comando `xtivreg` de Stata, en su opción de efectos fijos, realizamos una regresión entre la variable dependiente `rdxemp` y las variables `logemp` y productividad y endeudamiento, éstas dos últimas instrumentadas mediante sus respectivos retardos de orden 1. Y, a continuación, vía el comando `dmexogxt` de Stata, efectuamos el test Davidson-MacKinnon para exogeneidad, que arroja el siguiente resultado: 4.155841 F( 2,155) P-value = 0.0175. Por tanto, rechazamos la hipótesis nula de ausencia de endogeneidad.

En consecuencia, parece sensato admitir que puede existir un problema de endogeneidad. Para afrontarlo de la forma más sencilla posible, en todas las estimaciones que siguen, incluso las que se reflejan en los ya citados cuadros 9, 10 y 11, se utilizarán directamente como variables explicativas los retardos de primer orden de las variables productividad y endeudamiento. No se llega a aplicar métodos más complicados para afrontar el problema, por cuanto se entiende que éstos superarían los límites razonables de complejidad para el presente trabajo.

### **Cuadro 9: Estimación del modelo con efectos fijos**

Fixed-effects (within) regression	Number of obs	=	212		
Group variable: numero	Number of groups	=	46		
R-sq: within = 0.1676	Obs per group: min	=	1		
between = 0.2419	avg	=	4.6		
overall = 0.2467	max	=	6		
	F(3,163)	=	10.94		
corr(u_i, Xb) = -0.0680	Prob > F	=	0.0000		
rdxemp	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
productiv L1.	-.0035169	.0098248	-0.36	0.721	-.0229172 .0158834
endeudam L1.	-203.0842	150.0511	-1.35	0.178	-499.3788 93.21034
logemp	-6727.592	1253.976	-5.37	0.000	-9203.725 -4251.459
_cons	69939.78	9291.656	7.53	0.000	51592.25 88287.32
sigma_u	15030.101				
sigma_e	4211.541				
rho	.92719995 (fraction of variance due to u_i)				
F test that all u_i=0:	F(45, 163) =	26.98	Prob > F = 0.0000		

Al realizar el análisis de regresión mediante el modelo de efectos fijos, se observa que, a cualquiera de los habituales niveles de significación, el contraste del estadístico F rechaza la hipótesis nula de que el conjunto de los coeficientes estimados son iguales a cero. Sin embargo, considerando la misma hipótesis individualmente para cada uno de los coeficientes estimados, mediante los p valores de los estadísticos t se percibe que

sólo es significativa la variable del tamaño empresarial. En el presente caso, se omite en la estimación del modelo la variable propiedad, ya que es una variable indicador que, en la muestra utilizada, no varía a lo largo de los años dentro de cada empresa, de modo que no es adecuada para un modelo de efectos fijos.

Las variables explicativas mantienen en general el mismo signo respecto a la variable dependiente que en el análisis de la matriz de correlaciones, salvo en el caso de la variable explicativa L1.productividad, de manera que estaremos atentos a este signo en lo que sigue.

$R^2$  within corresponde al coeficiente de determinación de la regresión realizada con desviaciones respecto a la media.

$R^2$  between es el coeficiente de determinación de una regresión de las medias de la variable dependiente, que está en función de las medias de las variables independientes, calculado de la siguiente forma:

$$\text{corr}(x_{it}\hat{\beta}, y_{it})^2$$

El  $R^2$  es la medida de la bondad del ajuste del modelo. Para nuestro caso de análisis, no se van a interpretar estos coeficientes de correlación en ninguna estimación debido a que, como se explicará más adelante, se ha determinado la existencia de heterocedasticidad y autocorrelación, lo que reduce significativamente la fiabilidad del  $R^2$  como medida de la bondad del ajuste.

La última línea del cuadro 9 informa del estadístico F para el contraste de Breusch-Pagan, propuesto en Breusch y Pagan (1980). Ayuda a confirmar que no basta con realizar una simple regresión agrupada (*pooled*) mediante mínimos cuadrados ordinarios. Su hipótesis nula es que todos los errores específicos del individuo ( $u_i$ ) son cero. Y, si dichos errores son nulos, no tendría sentido utilizar técnicas de estimación específicas para panel (introducir efectos fijos o aleatorios, mínimos cuadrados generalizados factibles, etc.), sino que se podría aplicar una regresión por mínimos cuadrados ordinarios, sin tener en cuenta que los datos forman un panel. Se observa que se rechaza dicha hipótesis nula para cualquier nivel de significación, debido a que el p-valor es cero. En consecuencia, está adecuadamente justificado el uso de métodos específicos de panel.

Sigma-u, Sigma\_e y Rho, forman parte del contraste de Breusch-Pagan basado en los multiplicadores de Lagrange. Concretamente, Sigma-u es la desviación típica del error específico del individuo, es decir la raíz cuadrada de la varianza de  $u_i$ . Por su parte, Sigma\_e es la desviación típica de la perturbación aleatoria asociada a las observaciones ( $e_{it}$ ). Finalmente, Rho es la proporción que supone la varianza de  $u_i$  sobre la varianza total (varianza de  $u_i$  más varianza de  $e_{it}$ ) y, por lo tanto, se calcula dividiendo Sigma\_u entre la suma de Sigma\_u y Sigma\_e.

## Cuadro 10: Estimación del modelo con efectos aleatorios

```

i.prop          _Iprop_1-6          (_Iprop_1 for prop==A+ omitted)
note: _Iprop_2 omitted because of collinearity

Random-effects GLS regression              Number of obs   =       212
Group variable: numero                    Number of groups =        46

R-sq:  within = 0.1262                    Obs per group:  min =        1
        between = 0.4234                      avg =       4.6
        overall = 0.4365                      max =        6

Wald chi2(7) =       67.69
corr(u_i, X) = 0 (assumed)                Prob > chi2     =       0.0000

-----+-----
      rdxemp |          Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-----+-----
productiv L1. |   .0266591   .0102765     2.59  0.009   .0065176   .0468007
endeudam L1. |  -385.6452  165.2442    -2.33  0.020  -709.5179  -61.77244
      logemp |  -5848.323   865.082    -6.76  0.000  -7543.852  -4152.793
      _Iprop_2 | (omitted)
      _Iprop_3 |   5421.358   4147.489     1.31  0.191  -2707.571  13550.29
      _Iprop_4 |  -8347.649  7854.564    -1.06  0.288  -23742.31  7047.014
      _Iprop_5 |   1453.769   3860.754     0.38  0.707  -6113.171  9020.709
      _Iprop_6 |  -9929.462  10673.75    -0.93  0.352  -30849.63  10990.71
      _cons |   58814.23  6960.203     8.45  0.000   45172.48  72455.98
-----+-----
      sigma_u |   9029.4028
      sigma_e |   4211.541
      rho |   .82131984   (fraction of variance due to u_i)
-----+-----
Breusch and Pagan Lagrangian multiplier test for random effects
rdxemp[numero,t] = Xb + u[numero] + e[numero,t]
Estimated results:
      |          Var      sd = sqrt(Var)
-----+-----
      rdxemp |   3.06e+08   17485.77
      e |   1.77e+07   4211.541
      u |   8.15e+07   9029.403

Test:   Var(u) = 0
        chibar2(01) =    92.86
        Prob > chibar2 =    0.0000
-----+-----

```

El test de Wald efectúa un contraste de significación cuya hipótesis nula es que todos los coeficientes estimados son cero. Se observa que, para cualquiera de los niveles de significación habitualmente utilizados, se rechaza la hipótesis nula que establece que los coeficientes son todos iguales a cero.

Además, en este caso, respecto a la regresión con efectos fijos, mejoran los p valores de cada coeficiente individualmente considerado. Ahora, para cualquiera de los niveles de confianza habitualmente utilizados, no sólo es significativo el coeficiente estimado para la variable del tamaño empresarial, sino también el coeficiente correspondiente al desempeño (productividad). Y, al nivel de significación del 5%, es significativo el coeficiente estimado para la variable endeudamiento. A diferencia del modelo estimado con efectos fijos, las variables explicativas mantienen el mismo signo respecto a la variable dependiente que en el análisis de la matriz de correlaciones.

La última parte del cuadro 10 informa del contraste de Breusch-Pagan para un modelo de efectos aleatorios. Al igual que en el modelo de efectos fijos, se rechaza la hipótesis nula es que todos los errores específicos del individuo ( $u_i$ ) son cero. En consecuencia, está adecuadamente justificado el uso de métodos específicos de panel.

### Cuadro 11: Test de Hausman

	---- Coefficients ----			
	(b)	(B)	(b-B)	sqrt(diag(V_b-V_B))
	fixed	.	Difference	S.E.
L.producti~d	-.0035169	.0266591	-.030176	.004351
L.endeudam~o	-203.0842	-385.6452	182.5609	41.75312
logemp	-6727.592	-5848.323	-879.2693	1131.546

b = consistent under Ho and Ha; obtained from xtreg  
B = inconsistent under Ha, efficient under Ho; obtained from xtreg

Test: Ho: difference in coefficients not systematic

$$\begin{aligned} \text{chi2}(2) &= (b-B)' [(V_b-V_B)^{-1}] (b-B) \\ &= 19.60 \\ \text{Prob}>\text{chi2} &= 0.0001 \end{aligned}$$

Como podemos en el cuadro 11, que contiene el test de Hausman, se rechaza la hipótesis nula, por lo que se encuentran importantes diferencias entre ambos modelos y, en concreto, procede utilizar el modelo con efectos fijos.

Sin embargo, el test de Hausman se ha efectuado comparado un modelo con efectos fijos en que no figura la variable indicador i.prop y un modelo con efectos aleatorios que contiene dicha variable. Además, ese test tiene un problema de cómputo para llegar a resultados correctos cuando están por medio tal tipo de variables indicador, por intervenir entonces en el cálculo escalas muy diferentes. El recurso de utilizar el comando suest de Stata, aconsejado para situaciones de problemas de cómputo en el test de Hausman, no es viable en el presente caso, ya que estamos trabajando con datos de panel y dicho comando no es entonces aplicable.

Además, en cualquier caso, el test de Hausman no es robusto frente a posibles problemas de autocorrelación y heterocedasticidad, por lo cual llegados a este punto conviene analizar ante todo si esos problemas existen en nuestro caso. Es lo que se hace en el apartado siguiente.

#### **4. Análisis de autocorrelación y heterocedasticidad en efectos fijos y aleatorios.**

Tal y como se ha señalado, el contraste de Hausman realizado indica que es conveniente utilizar una estimación con efectos fijos. Sin embargo, también se acaba de apuntar que es conveniente realizar pruebas que permitan determinar la existencia o no de heterocedasticidad (diferentes varianzas) y autocorrelación entre las perturbaciones aleatorias o residuos.

#### **5. Contraste de autocorrelación**

El correspondiente análisis se efectúa mediante el denominado test de Wooldridge para autocorrelación. Este contraste fue desarrollado por Wooldridge (2002) y tiene como objetivo detectar autocorrelación en los errores de los modelos lineales de datos de

panel. El punto de partida es la hipótesis nula de que no existe autocorrelación de primer orden. A diferencia de lo que ocurrirá con la prueba de heterocedasticidad, bastará con realizar este contraste una sola vez y servirá tanto para efectos fijos como para efectos aleatorios, ya que no depende del método de estimación. El cuadro 12 contiene este test de Wooldridge para nuestro caso.

### **Cuadro 12: Contraste de autocorrelación (Wooldridge)**

```
Wooldridge test for autocorrelation in panel data
H0: no first order autocorrelation
      F( 1,      36) =      7.138
      Prob > F =      0.0113
```

El resultado nos permite rechazar la hipótesis nula para casi todos los niveles de significación habituales, pudiendo llegar a la conclusión de que existe autocorrelación.

## **6. Contrastes de heterocedasticidad (efectos fijos)**

Con el objetivo de profundizar en el análisis, se procedió a realizar pruebas de heterocedasticidad de los residuos en los modelos de regresión estimados. Estas pruebas son diferentes para efectos fijos y para efectos aleatorios.

En primer lugar, se realiza el contraste de heterocedasticidad en efectos fijos. Dicha prueba es un test de Wald para detectar heterocedasticidad en un modelo de efectos fijos y puede consultarse en Greene (2000: 598).

En este caso, la hipótesis nula es la homocedasticidad (iguales varianzas) de los residuos de la regresión de efectos fijos. Los contrastes de heterocedasticidad para efectos fijos se muestran a continuación.

### **Cuadro 13. Contrastes de heterocedasticidad (efectos fijos). Wald.**

```
Modified Wald test for groupwise heteroskedasticity
in fixed effect regression model

H0: sigma(i)^2 = sigma^2 for all i

      chi2 (46) =      84613.68
      Prob>chi2 =      0.0000
```

En este caso podemos ver que se rechaza la hipótesis nula de homocedasticidad (igualdad de varianzas) para cualquier nivel de significación habitual, dado que el p-valor es 0. Por tanto, existe heterocedasticidad.



## 7. Contrastes de heterocedasticidad (efectos aleatorios)

Tal como se ha mencionado anteriormente, los test de heterocedasticidad difieren según el modelo estimado, por lo que para efectos aleatorios se usa un test diferente al caso de efectos fijos.

El procedimiento aplicado consiste en realizar la regresión con efectos aleatorios y posteriormente estimar los residuos de la misma. Sobre los residuos resultantes se realiza una prueba de igualdad de varianzas (homocedasticidad) controlada por la variable que identifica a cada individuo en el panel (en nuestro caso, número). Ese contraste es un test robusto de Levene, en el que la hipótesis nula es la igualdad de varianzas entre grupos.

El paquete estadístico STATA proporciona el test de Levene y dos variantes del mismo que reemplazan la media por otros estadísticos: la variante de Brown, que reemplaza la media por la mediana y la variante de Forsythe, que reemplaza la media por la media truncada al 10% de los laterales. El cuadro 14 muestra el test robusto de Levene relativo a la homocedasticidad de los residuos en el modelo estimado con efectos aleatorios.

### **Cuadro 14. Contrastes de heterocedasticidad (efectos aleatorios)**

Levene W0 =	4.3792540	df(45, 166)	Pr > F = 0.00000000
Brow W50 =	2.5118664	df(45, 166)	Pr > F = 0.00001225
Forsythe W10 =	4.3792540	df(45, 166)	Pr > F = 0.00000000

Dicho cuadro 14 muestra que, tanto para el test de Levene como para sus variantes, la conclusión que se obtiene es el rechazo de la hipótesis nula para cualquier nivel de significación habitual, por lo que debe admitirse la existencia de heterocedasticidad en el modelo estimado con efectos aleatorios.

## 8. Reconsideración del test de Hausman y contraste de restricciones sobreidentificadas.

Tras realizar el análisis correspondiente, se determina la presencia de heterocedasticidad y autocorrelación, por lo que el test de Hausman no parece el más adecuado para este estudio. En su lugar, ha de emplearse el enfoque de regresión artificial recogido en Wooldridge (2002: 290-291), el cual resulta ser robusto a pesar de la existencia de heterocedasticidad y autocorrelación, garantizando así un estadístico de contraste positivo.

Para ello, se ejecuta en Stata el comando `xtoverid`, `robust cluster(numero)`, comando que implica la reestimación de la ecuación de efectos aleatorios incluyendo variables adicionales, que serían los regresores originales pero transformados en desviaciones respecto a la media. De este modo, el estadístico de contraste es un test de Wald sobre la significación de esas variables adicionales. Este test recibe el nombre de test de restricciones sobreidentificadas y se recoge en el cuadro 15. Su hipótesis nula es que se

cumplen las condiciones de ortogonalidad necesarias para que el estimador de efectos aleatorios sea consistente.

### **Cuadro 15. Test xtoverid para determinar la elección de modelo con efectos fijos o con efectos aleatorios**

```
Test of overidentifying restrictions: fixed vs random effects
Cross-section time-series model: xtreg re robust cluster(numero)
Sargan-Hansen statistic 26.048 Chi-sq(3) P-value = 0.0000
```

De acuerdo con el cuadro 15, se rechaza la hipótesis nula para cualquiera de los niveles de significación habituales, lo que es tanto como decir que procede optar por un modelo con efectos fijos.

Llegados a este punto, podría intentarse la estimación por mínimos cuadrados generalizados factibles (MCGF) mediante el comando `xtgls` de Stata, si bien introduciendo directamente en la especificación del modelo a estimar una variable ficticia representativa del efecto individual constante, pues, en principio, el citado comando `xtgls` está diseñado directamente para modelos con efectos aleatorios. No obstante, algunas pruebas hechas vía este método de estimación, y con diversas opciones para corregir los problemas de autocorrelación y heterocedasticidad detectados, no resultaron satisfactorias.

En consecuencia, por último se optó por estimar directamente un modelo final mediante el procedimiento de regresión Prais-Winsten, habida cuenta además de la recomendación de Beck y Katz (1995) citada páginas atrás. Recordemos que estos autores sostienen que los errores estándar corregidos para panel (PSCE) son más precisos que los estimadores de mínimos cuadrados generalizados factibles (MCGF).

## **9. Regresión de panel con errores estándar corregidos por autocorrelación y heterocedasticidad (Prais-Winsten).**

Se decide, pues, realizar una regresión para datos de panel con errores estándar corregidos por autocorrelación y heterocedasticidad (regresión Prais-Winsten).

Primero, dicho tipo de regresión se realiza sucesivamente con cada una de las variables independientes incluidas en las hipótesis planteadas al comienzo de la investigación empírica. Y, posteriormente, se efectúa una regresión general Prais-Winsten, que engloba todas las variables del modelo, calculando así los errores estándar del panel, corregidos por heterocedasticidad y por autocorrelación siguiendo un proceso AR(1) específico para panel, calculando la autocorrelación mediante el procedimiento Durbin-Watson.

Todo este proceso se recoge en los cuadros 16 a 20 y, al llegar a este último, permite comprobar que en los coeficientes estimados se mantienen los signos observados en la matriz de correlaciones y en esas regresiones univariantes.

Además, el cuadro 20 se aprovecha para presentar una estimación final que, precisamente por tener tal carácter y según anunciamos en su momento, introduce también como variables de control la variable tiempo (año), la variable subsector (nace) y el efecto de la coyuntura temporal sobre el correspondiente subsector (es decir, la interacción de variable año y variable nace), con lo cual se mejora la calidad de la estimación global, sin modificar por ello la conclusión que va a alcanzarse en el contraste de las hipótesis planteadas.

### **Cuadro 16. Regresión Prais-Winsten con la variable productividad.**

```

Group variable:  numero                Number of obs   =    214
Time variable:  año                   Number of groups =    46
Panels:         heteroskedastic (unbalanced)  Obs per group: min =    1
Autocorrelation: panel-specific AR(1)          avg = 4.652174
                                                max =    6
Estimated covariances =    46           R-squared       = 0.6059
Estimated autocorrelations =    46       Wald chi2(1)    =    6.60
Estimated coefficients =    2           Prob > chi2     = 0.0102

```

```

-----
          rdxemp |               Het-corrected
                |      Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
productividad |
  L1.          |   .1018074   .0396301    2.57   0.010    .0241338   .1794809
  |
  _cons        |  10463.47   5917.822    1.77   0.077   -1135.252  22062.19
-----+-----
          rhos = .7630161 .2442336 .4632197 .6443104 .9383788 ... .1094096
-----

```

### **Cuadro 17. Regresión Prais-Winsten con la variable endeudamiento.**

Prais-Winsten regression, heteroskedastic panels corrected standard errors

```

Group variable:  numero                Number of obs   =    230
Time variable:  año                   Number of groups =    49
Panels:         heteroskedastic (unbalanced)  Obs per group: min =    1
Autocorrelation: panel-specific AR(1)          avg = 4.693878
                                                max =    6
Estimated covariances =    49           R-squared       = 0.5708
Estimated autocorrelations =    49       Wald chi2(1)    =    0.83
Estimated coefficients =    2           Prob > chi2     = 0.3620

```

```

-----
          rdxemp |               Het-corrected
                |      Coef.   Std. Err.      z    P>|z|    [95% Conf. Interval]
-----+-----
endeudamiento |
  L1.          |  -343.1517   376.4573   -0.91   0.362   -1080.994  394.6912
  |
  _cons        |  23338.24   1157.663   20.16   0.000    21069.26  25607.22
-----+-----
          rhos = .6768154 .9775772 .8895632 .9675795 .6955383 ... .9427634
-----

```

### **Cuadro 18. Regresión Prais-Winsten con la variable tamaño.**

Prais-Winsten regression, heteroskedastic panels corrected standard errors

```

Group variable:  numero                Number of obs   =      247
Time variable:  año                    Number of groups =      49
Panels:         heteroskedastic (unbalanced)  Obs per group: min =      1
Autocorrelation: panel-specific AR(1)        avg = 5.040816
                                                max =      7
Estimated covariances =      49          R-squared       = 0.6208
Estimated autocorrelations =      49      Wald chi2(1)    = 39.62
Estimated coefficients =      2          Prob > chi2     = 0.0000
    
```

		Het-corrected				
rdxemp	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
logemp	-5516.614	876.4152	-6.29	0.000	-7234.356 -3798.872	
_cons	60476.67	6301.389	9.60	0.000	48126.18 72827.17	
rhos =		.6259332	.9648727	.9598055	.9796506 .9682807 ... .9330071	

### **Cuadro 19. Regresión Prais-Winsten con la variable propiedad.**

Prais-Winsten regression, heteroskedastic panels corrected standard errors

```

Group variable:  numero                Number of obs   =      247
Time variable:  año                    Number of groups =      49
Panels:         heteroskedastic (unbalanced)  Obs per group: min =      1
Autocorrelation: panel-specific AR(1)        avg = 5.040816
                                                max =      7
Estimated covariances =      49          R-squared       = 0.5903
Estimated autocorrelations =      49      Wald chi2(4)    = 60.81
Estimated coefficients =      5          Prob > chi2     = 0.0000
    
```

		Het-corrected				
rdxemp	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
_Iprop_2	0	(omitted)				
_Iprop_3	10703.02	5830.59	1.84	0.066	-724.7239 22130.77	
_Iprop_4	-1572.814	3369.341	-0.47	0.641	-8176.602 5030.973	
_Iprop_5	3745.06	2138.666	1.75	0.080	-446.6485 7936.768	
_Iprop_6	-12580.52	1860.56	-6.76	0.000	-16227.15 -8933.888	
_cons	21596.22	903.8758	23.89	0.000	19824.65 23367.78	
rhos =		.5656124	.9770859	.9592599	.9732197 .469229 ... .9565612	

Iprop\_2 - \_Iprop\_6 = 0  
chi2( 1) = 4.43  
Prob > chi2 = 0.0352

En la regresión univariante con la variable concentración de la propiedad, realizamos además un test de postestimación para ver si, en su conjunto, la variable indicador concentración de la propiedad es significativa respecto a la variable dependiente. Como podemos ver al final del cuadro 19, la hipótesis nula se rechaza, con lo que podemos afirmar que la concentración de la propiedad es una variable que tiene influencia en la intensidad de la inversión en I+D. Eso sí, el signo de tal influencia parece que varía según sea el grado de concentración de la propiedad.

**Cuadro 20. Regresión de panel con errores estándar corregidos por autocorrelación y heterocedasticidad (regresión Prais-Winsten) con todas las variables del estudio.**

Prais-Winsten regression, heteroskedastic panels corrected standard errors

Group variable:	numero	Number of obs	=	212	
Time variable:	año	Number of groups	=	46	
Panels:	heteroskedastic (unbalanced)	Obs per group: min	=	1	
Autocorrelation:	panel-specific AR(1)	avg	=	4.608696	
		max	=	6	
Estimated covariances	=	46	R-squared	=	0.7837
Estimated autocorrelations	=	46	Wald chi2(31)	=	1431.80
Estimated coefficients	=	55	Prob > chi2	=	0.0000

rdxemp	Het-corrected					
	Coef.	Std. Err.	z	P> z	[95% Conf. Interval]	
productividad						
L1.	.0959195	.0237725	4.03	0.000	.0493262	.1425128
endeudamiento						
L1.	-737.8882	359.6587	-2.05	0.040	-1442.806	-32.97008
logemp	-5679.496	487.5011	-11.65	0.000	-6634.981	-4724.012
_Iprop_2	0	(omitted)				
_Iprop_3	11838.84	4090.664	2.89	0.004	3821.283	19856.39
_Iprop_4	-4951.105	1888.117	-2.62	0.009	-8651.745	-1250.464
_Iprop_5	1794.989	1577.582	1.14	0.255	-1297.015	4886.994
_Iprop_6	-4339.198	2060.711	-2.11	0.035	-8378.117	-300.2796
_Iaño_2005	0	(omitted)				
_Iaño_2006	-2028.881	6778.12	-0.30	0.765	-15313.75	11255.99
_Iaño_2007	2131.195	9242.531	0.23	0.818	-15983.83	20246.22
_Iaño_2008	-3405.165	11156.41	-0.31	0.760	-25271.33	18461
_Iaño_2009	22133.44	11840.79	1.87	0.062	-1074.084	45340.96
_Iaño_2010	8732.367	11896.3	0.73	0.463	-14583.95	32048.68
_Inace_355	-14778.86	9222.597	-1.60	0.109	-32854.82	3297.097
_Inace_642	-25337.76	7254.542	-3.49	0.000	-39556.4	-11119.12
_Inace_671	1142.818	9800.523	0.12	0.907	-18065.85	20351.49
_Inace_672	-9232.521	9119.539	-1.01	0.311	-27106.49	8641.447
_Inace_733	-24723.53	7189.477	-3.44	0.001	-38814.65	-10632.42
_Inace_737	-2798.154	7032.113	-0.40	0.691	-16580.84	10984.53
_Inace_738	-9100.42	6914.333	-1.32	0.188	-22652.26	4451.423
_Inace_874	37280.76	9977.449	3.74	0.000	17725.32	56836.2
_IañoXnac_2005_355	0	(omitted)				
_IañoXnac_2005_642	31610.74	11853.43	2.67	0.008	8378.442	54843.04
_IañoXnac_2005_671	0	(omitted)				
_IañoXnac_2005_672	0	(omitted)				
_IañoXnac_2005_733	0	(omitted)				
_IañoXnac_2005_737	4417.744	12055.21	0.37	0.714	-19210.03	28045.52
_IañoXnac_2005_738	10327.33	12012.67	0.86	0.390	-13217.07	33871.74
_IañoXnac_2005_874	0	(omitted)				
_IañoXnac_2006_355	3831.515	6778.418	0.57	0.572	-9453.939	17116.97
_IañoXnac_2006_642	36960.33	11118.23	3.32	0.001	15169	58751.67
_IañoXnac_2006_671	-923.0892	6939.263	-0.13	0.894	-14523.8	12677.62
_IañoXnac_2006_672	0	(omitted)				
_IañoXnac_2006_733	20170.01	11113.77	1.81	0.070	-1612.574	41952.6
_IañoXnac_2006_737	9423.124	11387.2	0.83	0.408	-12895.38	31741.63
_IañoXnac_2006_738	12917.7	11368.92	1.14	0.256	-9364.964	35200.37
_IañoXnac_2006_874	0	(omitted)				
_IañoXnac_2007_355	-3380.447	9252.279	-0.37	0.715	-21514.58	14753.69
_IañoXnac_2007_642	28673.6	10629.82	2.70	0.007	7839.538	49507.66
_IañoXnac_2007_671	-5076.451	9471.414	-0.54	0.592	-23640.08	13487.18
_IañoXnac_2007_672	0	(omitted)				
_IañoXnac_2007_733	15316.58	10636.72	1.44	0.150	-5531.015	36164.18
_IañoXnac_2007_737	4901.913	11597.36	0.42	0.673	-17828.49	27632.31
_IañoXnac_2007_738	10351.12	11555.55	0.90	0.370	-12297.34	32999.57
_IañoXnac_2007_874	0	(omitted)				
_IañoXnac_2008_355	1655.916	11262.93	0.15	0.883	-20419.02	23730.85
_IañoXnac_2008_642	23586.01	10806.7	2.18	0.029	2405.261	44766.76
_IañoXnac_2008_671	1001.475	11310.38	0.09	0.929	-21166.45	23169.4
_IañoXnac_2008_672	0	(omitted)				
_IañoXnac_2008_733	20852.78	10762.36	1.94	0.053	-241.0495	41946.62

_IañoXnac_2008_737		11553.54	11629.26	0.99	0.320	-11239.39	34346.47
_IañoXnac_2008_738		11942.1	11651.27	1.02	0.305	-10893.96	34778.17
_IañoXnac_2008_874		-7690.791	11148.56	-0.69	0.490	-29541.56	14159.98
_IañoXnac_2009_355		-24431.24	11907.16	-2.05	0.040	-47768.85	-1093.631
_IañoXnac_2009_642		0	(omitted)				
_IañoXnac_2009_671		-22908.86	12013.16	-1.91	0.057	-46454.23	636.4954
_IañoXnac_2009_672		-24866.61	10781.83	-2.31	0.021	-45998.61	-3734.615
_IañoXnac_2009_733		0	(omitted)				
_IañoXnac_2009_737		-14513.72	9723.985	-1.49	0.136	-33572.38	4544.936
_IañoXnac_2009_738		-11424.41	9721.395	-1.18	0.240	-30478	7629.169
_IañoXnac_2009_874		-30245.6	11831.65	-2.56	0.011	-53435.21	-7055.996
_IañoXnac_2010_355		-8477.13	12046.57	-0.70	0.482	-32087.98	15133.72
_IañoXnac_2010_642		14612.42	9639.473	1.52	0.130	-4280.598	33505.44
_IañoXnac_2010_671		-8468.695	12079.19	-0.70	0.483	-32143.48	15206.09
_IañoXnac_2010_672		-6757.492	11469.27	-0.59	0.556	-29236.85	15721.87
_IañoXnac_2010_733		13674.76	9642.812	1.42	0.156	-5224.808	32574.32
_IañoXnac_2010_737		0	(omitted)				
_IañoXnac_2010_738		0	(omitted)				
_IañoXnac_2010_874		-14780.49	11894.91	-1.24	0.214	-38094.09	8533.11
_cons		44858.96	11248.47	3.99	0.000	22812.36	66905.57
-----							
rhos =		-.0383453	.7255867	.5267242	.7766572	.9310922	... .4838692
-----							

A la vista del cuadro 20, se deduce que todos los regresores son globalmente significativos, dado que el p-valor del test de Wald que aparece al final de dicha estimación es cero.

A su vez podemos ver que, en el modelo final, obtenemos un coeficiente  $R^2$  bastante elevado, con lo cual podemos afirmar que el modelo estimado explica bastante bien la variable dependiente. Del mismo modo, para las variables L1.productividad, L1.endeudamiento y logemp, estadísticos z de cada uno de los coeficientes estimados son significativos al nivel del 5%, de modo que cabe afirmar que son tres variables explicativas muy relacionadas con la variable dependiente.

Tras el análisis de las tres variables explicativas que acabamos de citar, pasamos a hacer una breve mención a las variables indicador utilizadas como control, dejando aparte la variable propiedad, pues es una variable explicativa principal incluida en una de las hipótesis planteadas y, además, ya fue analizada en el cuadro 19, a la vez que sus resultados en éste se mantienen ahora. El cuadro 21 recoge para esas variables indicador de control unos contrastes de significación de la hipótesis nula de que sus coeficientes estimados son iguales a cero.

### **Cuadro 21. Analisis de las variables indicador.**

_Iaño_2005-_Iaño_2010	Inace_355-Inace_874	IañoXnac_2005_355-IañoXnac_2010_8
chi2( 1) = 110.71	chi2( 1) = 0.54	chi2( 1) = 1.54
Prob > chi2 = 0.0000	Prob > chi2 = 0.4629	Prob > chi2 = 0.2140

A la vista del cuadro 21, se deduce que sólo la variable tiempo (año) tiene un conjunto de coeficientes estimados significativamente distintos de cero. En el caso de las otras variables indicador de control, se acepta la hipótesis nula y, por lo tanto, no son significativas como elementos de influencia en la explicación de la variable dependiente a que se refiere el modelo estimado.

## 10. Resultados globales del contraste de las hipótesis de partida.

En el presente epígrafe se procede a señalar los resultados alcanzados en el contraste de las hipótesis de partida, sobre la base del modelo final estimado. Para ello, se presenta el cuadro 22, que resume esos resultados del contraste de las hipótesis.

**Cuadro 22. Esquema de los resultados del contraste de hipótesis.**

Hipótesis	Variables Independientes	Indicador	Relación esperada	Relación obtenida
H1: Cuanto mayor sea el tamaño de la empresa, mayor será la intensidad de la inversión en I+D en las empresas de <i>software</i>	Logaritmo del número de empleados.	Logemp	positiva	Negativa
H2: Cuanto mayor sea el desempeño ( <i>performance</i> ), mayor será la intensidad en la inversión en I+D en empresas de <i>software</i>	Ratio de productividad.	Productividad	positiva	Positiva
H3: Cuanto mayor es la concentración de la propiedad, menor será la intensidad de la inversión de la empresa en I+D	Escala de concentración de la propiedad.	Prop	negativa	Positiva o negativa, según grados de concentración
H4: Cuanto mayor es el endeudamiento de una empresa, menor será la intensidad de inversión en I+D de la compañía	Inverso del ratio de solvencia.	Endeudamiento	negativa	Negativa

A continuación, se describen las hipótesis propuestas y los resultados obtenidos en su contraste.

*H1: Cuanto mayor sea el tamaño de la empresa, mayor será la intensidad de la inversión en I+D en las empresas de software.*

Contrariamente a lo que establecíamos como hipótesis inicial, las empresas con mayor tamaño tienen una menor inversión proporcional en Investigación y Desarrollo. No es un resultado empírico nuevo. En realidad, de esta manera nuestro trabajo se sitúa en

consonancia con los resultados obtenidos por Scherer y Ross (1990) y Veciana Verges (2007).

Probablemente, esto puede suceder debido a la singularidad de las empresas de *software*, cuya actividad productiva es muy singular: difiere de otros sectores, puesto que su base productiva se caracteriza por la creación de conocimiento e ideas, los cuales son activos intangibles. En ellas el factor humano es fundamental, además de que, en la presente investigación, se ha expresado la intensidad en la inversión en I+D precisamente en términos de ésta variable medida por empleado. Así, cabe explicar que empresas de tamaño relativamente reducido, es decir, que cuentan con un número de trabajadores no muy elevados, de hecho se caracterizan por ser más intensivas en inversión en I+D por empleado; y seguramente se trata de trabajadores muy especializados y de alta cualificación..

*H2: Cuanto mayor sea el desempeño (performance), mayor será la intensidad en la inversión en I+D en empresas de software.*

Como ya habíamos descrito, la variable desempeño es una de las más estudiadas en la determinación de la intensidad de la inversión en I+D. Tras analizar los datos de nuestra muestra, se acepta la hipótesis que establece que a mayor desempeño (*performance*) de la empresa se produce un aumento en la intensidad de la inversión en I+D.

Estos resultados apoyan la tesis de trabajos anteriormente mencionados, como González y Pargas (2010).

*H3: Cuanto mayor es la concentración de la propiedad, menor será la intensidad de la inversión de la empresa en I+D.*

En cuanto a la concentración de la propiedad de las acciones de la compañía, cabe destacar que en nuestros resultados empíricos muestran que la propiedad es una variable significativa en lo referente al nivel de inversión en I+D. Es decir, en principio encontramos una relación entre ambas variables.

Sin embargo, no podemos afirmar que los diferentes niveles en la escala de concentración afecten siempre de la misma manera, esto es, en idéntica dirección, a la variable dependiente analizada. Depende del grado de concentración al que nos estemos refiriendo en cada caso, dentro de una escala con varios niveles establecidos además de acuerdo con la metodología concreta aplicada al respecto por el Bureau Van Dijk al elaborar su base *Amadeus*.

De esta manera, nuestros resultados recuerdan en cierto modo los alcanzados por Beyer y Kraft (2007), con su relación en forma de U invertida. O, en el peor de los casos, tal vez sean concordantes con los resultados de Acs e Isberg (1991), quienes, como ya explicamos en el Capítulo I, no hallaron una relación direccional clara a la asociación entre concentración de la propiedad e inversión en I+D.



*H4: Cuanto mayor es el endeudamiento de una empresa, menor será la intensidad de la inversión en I+D de la compañía.*

Esta hipótesis cuarta se confirma, según los resultados obtenidos a partir de los datos considerados en este trabajo. Por lo tanto, podemos afirmar que el endeudamiento es una variable determinante negativa de la intensidad de la inversión en I+D, llegando así a las mismas conclusiones presentadas por Aghion et al. (2004).

Esto puede derivarse de que el mayor endeudamiento que pesa una empresa actúa como una restricción financiera, de modo que disminuye su capacidad de inversión futura, por ejemplo el año siguiente: por razones obvias, tendrá un peor acceso a los mercados financieros y su estructura de capital no la permitirá una mayor inversión en I+D. En el fondo, en asunto puede entenderse como la consecuencia de que las empresas con un substancial capital organizacional, es decir, importantes activos intangibles, afrontan muy relevantes costes de insolvencia en caso de verse abocadas a graves dificultades financieras. Por ello, es tradicional y general recomendarlas un ratio de endeudamiento bajo (Shapiro y Balbirer, 2000: 483-484). Es justamente una sugerencia muy apropiada para las empresas analizadas en el presente trabajo, caracterizadas por empleados de alta cualificación y con conocimientos específicos.

## CAPÍTULO IV. A MODO DE CONCLUSIÓN.

En el presente trabajo de investigación se pretendía conocer las variables determinantes de la intensidad de la inversión I+D, tal y como se anunció desde su inicio. Para ello, se ha realizado un estudio econométrico por medio del programa Stata, para una muestra de empresas europeas del sector del *software* y los servicios informáticos, durante los años 2004-2010. Las fuentes de información han sido las bases de datos *Amadeus* y *Economics of Industrial Research and Innovation*.

Con este fin, en el Capítulo I se ha presentado, entre otras cosas, una revisión de la literatura especializada existente al respecto, es decir, los resultados alcanzados en estudios empíricos previos ligados al objeto de investigación.

En segundo lugar, se ha propuesto un esquema metodológico a seguir en el trabajo. Así, se han formulado las hipótesis teóricas a contrastar y se han especificado tanto las variables utilizadas como las bases de datos y la muestra de empresas a analizar, a la vez que se ha presentado la estrategia econométrica prevista.

En tercer lugar, y como parte central de la investigación, se han mostrado los resultados derivados del estudio empírico efectuado.

En este sentido, a modo de conclusión, cabe resumir fácilmente los principales hallazgos empíricos contenidos en el Capítulo III de esta investigación, con la finalidad de intentar obtener un algún avance de conocimientos sobre los factores determinantes de la intensidad de la inversión en I+D en las empresas europeas de *software* y servicios informáticos durante los años considerados.

A la luz de los resultados empíricos aquí alcanzados, parece claro que existe una relación significativa entre las variables independientes analizadas y la variable dependiente estudiada, puesto que se han rechazado las hipótesis nulas formuladas en sentido contrario y se han aceptado las hipótesis alternativas de relación relevante.

Más en concreto, se ha confirmado que la variable explicativa desempeño (*performance*) está relacionada positivamente con la intensidad de la inversión en I+D. Mientras que, tal y como parecía evidente, la variable explicativa endeudamiento tiene una influencia negativa. En cuanto atañe a la variable explicativa que denominamos tamaño, se detecta una relación negativa, contraria a la hipótesis teórica tal como fue literalmente enunciada. Sin embargo, existe evidencia empírica previa en el mismo sentido que la hallada en el presente trabajo; y, además, cabe explicarla con suficiente lógica, como se ha señalado en su momento. Por su parte, los resultados obtenidos en el análisis de la variable explicativa concentración de propiedad conducen a pensar que la dirección de su impacto depende del nivel en que se sitúe dicho grado de concentración.

Claro está, estas conclusiones deben ser entendidas desde la peculiar perspectiva que representan las singulares características del sector del *software*. Es decir, “estos resultados no se pueden suponer generalizables a todo tipo de empresas, (...) sin

embargo, constituyen una información importante para tenerla en cuenta en estudios que indaguen sobre el tema con mayor amplitud” (Silva, 2007: 68).

Para finalizar este apartado de conclusiones, se proponen las siguientes líneas de la investigación futura.

En primer lugar, el estudio detallado de la posible relación bidireccional existente entre las variables aquí consideradas.

En segundo lugar, una apuesta por la recogida de datos relativos a otros sectores concretos, como se ha hecho ahora con la singular industria del *software* y de los servicios informáticos.

A la vez, habrá que profundizar en el conocimiento de las técnicas econométricas y en la posible estimación de modelos dinámicos para datos de panel.

En último término, y a la espera del transcurso del tiempo, se debería apostar por los trabajos de naturaleza comparativa y en diferentes ámbitos o niveles. En tal sentido, se sugiere el estudio a largo plazo de la zona europea en periodos económicos tanto de expansión como de recesión.

## Referencias bibliográficas

Acs, Z.J. y Audretsch, D.B. (1991): "Innovation and Size at the Firm Level", *Southern Economic Journal*, vol. 57, pp. 739-344.

Acs, Z.J. y Audretsch, D.B. (1993). *Small Firms and Entrepreneurship: An East-West Perspective*. Cambridge University Press, Cambridge.

Aghion, Philippe, Christopher Harris, and John Vickers (1997): "Competition and Growth with Step-by-Step Innovation: An Example", *European Economic Review, Papers and Proceedings*, vol. 41, pp. 771–782.

Amabile, M. (1982): "Social Psychology of Creativity: A Consensual Assessment Technique", *Journal of Personality and Social Psychology*, vol. 43, pp. 997-1013.

Alberts, C. y Dorofee, A. (2002): "Managing Information Security Risks: The OCTAVE (SM) Approach". Addison-Wesley, New York.

Aparicio, J. y Márquez, J. (2005): *Diagnóstico y especificación de modelos panel en STATA 8.0*, Curso de Métodos cuantitativos, CIDE. México.

Bah, R. y Dumontier, P. (2001): "R & D Intensity and corporate financial policy: some international evidence", *Journal of Business Finance and Accounting*, vol. 28, pp. 671-692.

Baumol W. (2002): "Entrepreneurship, Innovation and Growth: The David–Goliath Symbiosis", *Journal of Entrepreneurial Finance and Business Ventures*, vol. 7, pp. 1-10.

Bean, A. (1995): "Why some R&D organizations are more productive than others", *Research Technology Management*, vol. 1, pp.25-30.

Beck, Nathaniel and Jonathan N. Katz. (1995): "What To Do (and Not To Do) with TimeSeries Cross-Section Data", *American Political Science Review*, vol. 89, pp. 634–47.

Bhagat, S. y Welch, I. (1995): "Corporate research & development investments: International comparison". *Journal of Accounting and Economics*, vol. 19, pp. 443- 470.

Bhargava, A.; Franzini, L., y Narendranathan, W. (1982): "Serial Correlation and the Fixed Effects Model". *Review of Economic Studies*, vol. 49, p. 533-549.

Breusch, T. S. y Pagan, A. R. (1980): "The Lagrange multiplier test and its applications to model specification in econometrics". *Review of Economic Studies*, vol. 47, pp. 239-253.

Campbell-Kelly, M. (2003): *From airline reservations to Sonic the Hedgehog: a history of the software industry*. MIT Press, Cambridge (MA).

Cohen, W.M. y Keppeler, S. (1996): "A reprise of size and R&D", *Economic Journal*, vol. 106, pp. 925-951.

Cohen, W.M. y Levinthal, D.A. (1989): "Innovation and learning: the two faces of R&D", *The Economic Journal*, vol. 99, pp. 569-596.

Comisión Europea (2010): "*Europa 2020: Una estrategia para un crecimiento inteligente, sostenible e integrador*". Bruselas.

Costa, M. T.; Duch, N. y Lladó, J. (2000): "Determinantes de la innovación y efectos sobre la competitividad: el caso de las empresas textiles". Documento de trabajo 2000/4, Instituto de Economía de Barcelona (IEB), Barcelona. Disponible en: [www.pcb.ub.es/ieb](http://www.pcb.ub.es/ieb).

Díaz Martín, M. C. (1996): "Factores determinantes de la innovación tecnológica para las empresas pequeñas". *Cuadernos de Estudios Empresariales*, nº 6, pp. 145-154.

Dosi, G. (1988): "Sources, Procedures, and Microeconomic Effects of Innovation". *Journal of Economic Literature*. Vol. 26, pp: 1120-1171.

González Bravo, M. I. y Pargas Carmona, L. A. (2010): "Intensidad en I+D y desempeño empresarial en las PYMES: un enfoque multidimensional". *Revista Internacional de la Pequeña y Mediana Empresa*, vol. 1, pp. 40-58.

Gosh, S. (2012): "Does R&D intensity influence leverage? Evidence from Indian firm-level data". *Journal of International Entrepreneurship*, vol. 10, pp. 158-175.

Grabowski, H.G. (1968): "The determinants of Industrial Research and Development: A study of Chemical, Drug, and Petroleum Industries" *Journal of Political Economy*, vol. 76, pp. 293-306.

Grabowski, H.G. y Mueller D.G. (1978): "Industrial research and development, intangible capital stocks, and firm profit rates". *Bell Journal of Economics*, vol. 9, pp. 328-343.

Greene, W. (2000): *Econometric Analysis*, Prentice-Hall, New York.

Griliches, Z. (1979): "Issues in assessing the contribution of research and development to productivity growth", *Bell Journal of Economics*, vol. 10, pp. 92-116.

Griliches, Z. (1986): "Productivity, R&D and basic research at the firm level in the 1970s", *American Economic Review*, vol. 76, pp. 141-154.

Griliches, Z. (1995): "R&D and Productivity: Econometric results and Measurement issues", en: Stoneman, P. (ed.): *Handbook of the Economics of Innovation and Technological Change*, Blackwell, Oxford.

Griliches, Z. y Mairesse, J. (1984): "Productivity and R&D at the firm level", en Griliches, Z. (ed.): *Patents and productivity*, University of Chicago Press, Chicago, pp. 339-374.

Gumbau, M. (1997): "Análisis microeconómico de los determinantes de la innovación: aplicación a las empresas industriales españolas", *Revista Española de Economía*, vol. 14, pp. 41-66.

Hall, B.H. (2002): "The financing of research and development", *Oxford Review of Economic Policy*, vol. 18, pp. 35-51.

HILL, C.W.L. and Scott A. SNELL (1988). "External Control, Corporate Strategy, and Firm Performance in Research Intensive Industries," *Strategic Management Journal* (1986-1998), vol. 9, pp. 577-590

Joh, S.W. (2003): "Corporate governance and firm profitability: evidence from Korea before the economic crisis". *Journal of Financial Economics*, vol. 68, pp. 287-322.

Kraens et al., (2009): "Employment Growth and International Trade: a small open economy perspective". Working Papers. 09-9, University of Aarhus, Aarhus School of Business, Department of Economics.

Levene, H. (1960): "Robust tests for equality of variances", en Olkin, I. (ed.), *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, Stanford University Press, Menlo Park (CA), pp. 278-292.

Lhuillery, S. y Templé, P. (1995): "L'impact du C.I.R. sur l'innovation dans les entreprises", en Haudeville, B.; Héraud, J.A. y Humbert, M. (eds.), *Technologie et Performances Économiques*, Economica, París, pp. 155-182.

López Mielgo, N.; Montes Peón; J. M.; Prieto Rodríguez, J. y Vázquez Ordás, J. C. (2005): "Una Aplicación del Análisis Tobit a Datos de Panel: Factores Determinantes de la I+D en la Industria Española". Ponencia en sesión paralela dentro del VIII Encuentro de Economía Aplicada, Murcia.

Mairesse, J. y Hall, B. H. (1996): "Estimating the productivity of research and development: An exploration of GMM methods using data on French and United States manufacturing firms", Documento de trabajo nº 5501, NBER, Cambridge (MA).

Mansfield, E. (1962): "Entry, Gibrat's Law, Innovation, and the Growth of Firms", *American Economics Review*, vol. 52, pp. 1023-1051.

Marra, M.A. (2007): “Tamaño, restricciones financieras e inversión en I+D”. *Revista de Economía Aplicada*, vol. 15, pp. 99-123.

Martinsonn, G. (2009): “Finance and R&D Investments - is there a debt overhang effect on R&D investments?”, CESIS Electronic Working Paper Series nº 174, Estocolmo.

Mohnen, P. y Hall, B.H. (2013): “Innovation and productivity: an update”. UNU-MERIT Working Paper Series 021. United Nations University, Maastricht Economic and social Research and training Centre on Innovation and Technology, Maastricht.

Moreno, R. y Ortega-Argilés, R. (2007): “Firm competitive strategies and likelihood of survival: the spanish case”. *Papers on Entrepreneurship, Growth and Public Policy*, Max Planck Institute of Economics, Entrepreneurship, Growth and Public Policy Group.

Mowery, D.C. (1995): *The Boundaries of the US firm in R&D” coordination and information: historical perspectives on the organization structure*. Oxford University Press, Oxford.

Organización para la Cooperación y el Desarrollo Económico: (1997): *Medición de las Actividades Científicas y Tecnológicas. Directrices propuestas para recabar e interpretar datos de la innovación tecnológica: Manual Oslo*. OCDE. París.

Organización para la Cooperación y el Desarrollo Económico: (2002): *Manual de Frascati. Propuesta de Norma Práctica para Encuestas de Investigación y Desarrollo Experimental*. OCDE. París.

Organización para la Cooperación y el Desarrollo Económico: (2005): *Oslo Manual. Guidelines for collecting and interpreting innovation data*. OCDE. París.

Patel, . y Pavitt, K. (1994): “National innovation systems: why they are important, and how they might be measured and compared”, *Economics of Innovation and New Technology*, vol. 3, pp. 75-95.

Pindado, J. et al. (2011): “Family control and investment–cash flow sensitivity: Empirical evidence from the Euro zone”. *Journal of Corporate Finance*, vol. 17, pp. 1389–1409

Sánchez Bueno, M.J. (2008): *El Proceso Innovador y Tecnológico: Estrategias y Apoyo Público*. Netbiblo, Oleiros.

Scherer, F.M y Ross, D. (1990): *Industrial Market Structure and Economic Performance*. Houghton Mifflin. Boston.

Scherer, F.M. (1984): *Innovation and Growth: Schumpeterian Perspectives*. MIT Press, Cambridge (MA).

Schumpeter, J.A. (1934): *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle*. Harvard University Press, Cambridge (MA).

Shapiro, A.C. y Balbirer, S.D. (200): *Modern Corporate Finance*, Prentice Hall, Upper Saddle River (NJ).

Schumpeter, J.A. (1942): *Capitalism, Socialism and Democracy*, Harper, Nueva York.

Silva Aristeguieta, A. (2007): “Determinantes de la innovación en la empresa”. *Anales de la Universidad Metropolitana*, vol. 7, pp. 53-71.

Smith, A. (2001): *La riqueza de las naciones*. Alianza Editorial, Madrid.

Suriñach Caralt (dir.) (2005): “Innovación y estructura de la propiedad de las empresas españolas”, Ministerio de Industria, Turismo y Comercio, Dirección General de Política de la Pequeña y Mediana Empresa, Colección Pensamiento PYME nº 2.

Tsai, K.H. y Wang, J.C. (2005): “Does R&D *performance* decline with firm size? A reexamination in terms of elasticity”. *Research Policy*, vol. 34, pp. 966-976.

Tucker, G. et al. (2011): “A profile of *software* Industry Risk”. Disponible en: <http://www.protviti.com/en-US/Pages/A-Profile-of-software-Industry-Risk.aspx>.

Veciana Verges, J.M. (2007): “Las nuevas empresas en el proceso de innovación en la sociedad del conocimiento evidencia empírica y políticas públicas”. *Economía industrial*, nº 363, pp. 103-118.

Wennberg, K. (2009) “The roles of R&D in new firm growth. *Journal name Small Business Economics*, vol. 33, pp. 77-89.

Woolridge, J.M. (1988): “Competitive Decline and Corporate Restructuring: Is a Myopic Stock Market to Blame?”, *Journal of Applied Corporate Finance*, vol. 1, pp. 26-36.

Woolridge, J.M. (2002): *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge (MA).

Zacchiroli, S. (2011): “Debian: 18 years of free *software*, do-ocracy, and democracy”. En Workshop on Open Source and Design of Communication. ACM, New York.



## Índice de cuadros

CUADRO 1. ALGUNOS ESTUDIOS SOBRE LOS FACTORES DETERMINANTES DE LA INVERSIÓN EN I+D .....	13
CUADRO 2. ANÁLISIS COMPARATIVO DE DIMENSIÓN EMPRESARIAL / INNOVACIÓN TECNOLÓGICA .....	15
CUADRO 3. HIPÓTESIS A CONTRASTAR Y RESULTADOS ESPERADOS .....	21
CUADRO 4. VARIABLES INDEPENDIENTES Y DE CONTROL.....	22
CUADRO 5: RELACIÓN DE EMPRESAS SELECCIONADAS PARA EL ESTUDIO.....	25
CUADRO 6. BASES DE DATOS DE TODAS LAS VARIABLES.....	26
CUADRO 7: PRINCIPALES ESTADÍSTICOS DESCRIPTIVOS DE LAS VARIABLES DE LA MUESTRA .....	34
CUADRO 8: MATRIZ DE LAS CORRELACIONES DE LAS VARIABLES A ESTUDIAR.....	35
CUADRO 9: ESTIMACIÓN DEL MODELO CON EFECTOS FIJOS.....	36
CUADRO 10: ESTIMACIÓN DEL MODELO CON EFECTOS ALEATORIOS.....	38
CUADRO 11: TEST DE HAUSMAN .....	39
CUADRO 12: CONTRASTE DE AUTOCORRELACIÓN (WOOLDRIDGE) .....	40
CUADRO 13. CONTRASTES DE HETEROCEDASTICIDAD (EFECTOS FIJOS). WALD. ....	40
CUADRO 14. CONTRASTES DE HETEROCEDASTICIDAD (EFECTOS ALEATORIOS) .....	41
CUADRO 15. TEST XTOVERID PARA DETERMINAR LA ELECCIÓN DE MODELO CON EFECTOS FIJOS O CON EFECTOS ALEATORIOS .....	42
CUADRO 16. REGRESIÓN PRAIS-WINSTEN CON LA VARIABLE PRODUCTIVIDAD. ....	43
CUADRO 17. REGRESIÓN PRAIS-WINSTEN CON LA VARIABLE ENDEUDAMIENTO.....	43
CUADRO 18. REGRESIÓN PRAIS-WINSTEN CON LA VARIABLE TAMAÑO.....	44
CUADRO 19. REGRESIÓN PRAIS-WINSTEN CON LA VARIABLE PROPIEDAD.....	44
CUADRO 20. REGRESIÓN DE PANEL CON ERRORES ESTÁNDAR CORREGIDOS POR AUTOCORRELACIÓN Y HETEROCEDASTICIDAD (REGRESIÓN PRAIS-WINSTEN) CON TODAS LAS VARIABLES. ....	45
CUADRO 21. ANALISIS DE LAS VARIABLES IDENTIFICADOR. ....	46
CUADRO 22. ESQUEMA DE LOS RESULTADOS DEL CONTRASTE DE HIPÓTESIS. ....	47