

CAPÍTULO II

DATOS Y METODOLOGÍA

1. INTRODUCCIÓN

La consecución de los objetivos planteados en el proyecto origina la necesidad de la elección de la fuente de información adecuada y disponible, de manera que puedan obtenerse los datos necesarios y sea posible su manejo para su posterior análisis. En este capítulo se describe el proceso llevado a cabo en la investigación para disponer de los datos de forma que puedan ser analizados y se introducen las metodologías estadísticas y econométricas utilizadas.

El proceso comienza con la descripción de la fuente de información utilizada que permite la obtención de datos para las empresas objeto de estudio y la comparación entre los distintos grupos que se formen según los criterios de segmentación elegidos, de tal forma que puedan observarse comportamientos diferenciados y sus componentes explicativos. La identificación de los grupos de estudio da lugar a la necesidad del manejo de un gran volumen de datos, tanto por el número de empresas a considerar en las muestras como por la amplitud del esquema de ratios necesario para el análisis cuantitativo, y por la dimensión y el tratamiento estadístico de la encuesta para el análisis cualitativo.

Ambos análisis, cuantitativo y cualitativo, precisan de la obtención de muestras significativas de empresas que reproduzcan, según los criterios de segmentación requeridos, la estructura de la población de éstas en la economía española. Con el soporte de estas muestras, el análisis cuantitativo de la información parte del cálculo de un conjunto de ratios económico-financieros para cada una de las empresas presentes en la

muestra, que mediante su agregación en los grupos definidos por los criterios de segmentación permite el estudio comparado de éstos. La metodología del análisis cuantitativo efectuado comprende: *a)* la representación gráfica de la evolución durante el periodo de estudio de los valores de las medidas de posición central de los grupos de empresas que componen las muestras, *b)* el cálculo de los estadísticos descriptivos que caracterizan a esos grupos, *c)* la aplicación de *tests* y contrastes estadísticos que permitan confirmar o rechazar la existencia de comportamientos disímiles observados gráficamente, *d)* la aplicación de técnicas econométricas que ratifiquen o no las conclusiones de las pruebas estadísticas.

Por su parte, el análisis cualitativo efectuado, que también da respuesta al Convenio firmado entre SERVILAB y la Fundación Rafael del Pino en 2005, se lleva a cabo mediante la elaboración, envío y análisis de una encuesta. Esta tarea demoscópica está compuesta de dos partes: la primera se dirige a los mismos emprendedores encuestados en el *Primer Informe de Emprendedores* en el año 2002 y la segunda a nuevos emprendedores que han comenzado su actividad empresarial en los últimos tres años, entre 2003 y 2005.

2. FUENTE DE INFORMACIÓN UTILIZADA E IDENTIFICACIÓN DE LOS GRUPOS DE ESTUDIO

Para la realización del estudio se ha dispuesto de la base de datos del Sistema de Análisis de Balances Ibéricos (SABI). Se trata de una base financiera de las principales empresas españolas y portuguesas elaborada por Bureau van Dijk Electronic Publishing, S. A., y distribuida en España por Informa, Información Económica, S. A. Esta base de datos está a disposición del personal docente e investigador de la Universidad de Alcalá para los fines que les son propios. Incluye los estados financieros del periodo 1997-2004 para un alto porcentaje de las compañías mercantiles de las 17 Comunidades Autónomas españolas que depositan sus cuentas en los Registros Mercantiles.

En la versión de la base SABI utilizada (diciembre 2005) existen un total de 976.598 empresas, de ellas se han seleccionado aquellas que tienen formas jurídicas mercantiles de sociedad anónima, limitada y cooperativa. El conjunto de empresas con dichas formas jurídicas se ha segmentado según los criterios de tamaño y edad.

En función del criterio de tamaño se han diferenciado cuatro grupos según el número de trabajadores: microempresas, pequeñas, medianas y grandes empresas. Es decir, empresas con menos de diez trabajadores,

entre diez y 49, entre 50 y 249 y de más de 249 empleados, respectivamente.

Según el criterio de edad, se han considerado cuatro grupos, para los que se ha calculado su edad desde su año de constitución hasta el año 2005. Estos grupos son: empresas con más de diez años de funcionamiento (constituidas con anterioridad a 1995), empresas con más de seis años y menos de diez (creadas entre 1996 y 1999), empresas con más de tres años y menos de seis (fundadas entre el 2000 y 2002) y empresas con menos de tres años completos (en constitución a partir del año 2003).

Esta doble segmentación por los criterios de tamaño y edad organiza una disposición de dieciséis estratos en los que se divide la población de empresas de SABI. Población que se reduce respecto a la inicialmente considerada de 976.598 empresas a 588.031. Esta disminución se debe a que algunas de las empresas no adoptan una de las tres formas jurídicas mercantiles establecidas en el estudio y, principalmente, a la falta de información sobre el número de trabajadores o la fecha de constitución.

El número de empresas que componen cada uno de los dieciséis grupos con el porcentaje de cada uno en los que se divide la población para la que se dispone de información completa se muestra en el cuadro 2.1.

Esta población de empresas con información completa en la base SABI son sociedades mercantiles que han sobrevivido desde su fecha de constitución hasta el año 2004. Por tanto, los resultados que se obtengan de los correspondientes análisis cualitativo y cuantitativo deben ser interpretados como provenientes de empresas en las que la actividad del empresario, en su conjunto y hasta el momento, ha resultado exitosa. En caso de haber fracasado, la empresa no habría sobrevivido y no se dispondría de sus datos en la versión de la base SABI disponible en 2005.

3. MUESTRAS DE EMPRESAS

El análisis cuantitativo de la información económico-financiera de las empresas ha requerido diseñar tres muestras obtenidas a partir de la anterior población.

Una de ellas reproduce la estructura de la población de empresas con información completa en la Base SABI. En la medida de lo posible, el diseño de esta muestra tiene como objetivo tratar de obtener un esquema de la estructura empresarial española y de facilitar un grupo de control con que se comparen los resultados obtenidos en los grupos de empre-

CUADRO 2.1
POBLACIÓN DE EMPRESAS CON INFORMACIÓN COMPLETA EN LA BASE SABI DISTRIBUIDAS
POR TAMAÑO Y EDAD, NÚMERO Y PORCENTAJE

Empleo/Edad	> 10 años		De 6 a 10 años		De 3 a 6 años		< 3 años		Total	
	Empresas	%	Empresas	%	Empresas	%	Empresas	%	Empresas	%
Menor de 10	190.169	32,34	118.635	20,17	97.402	16,56	25.539	4,34	431.745	73,42
De 10 a 49	83.954	14,28	27.303	4,64	16.132	2,74	3.018	0,51	130.407	22,18
De 50 a 249	15.722	2,67	3.125	0,53	1.660	0,28	255	0,04	20.762	3,53
Mayor de 249	4.118	0,70	607	0,10	331	0,06	61	0,01	5.117	0,87
Total	293.963	49,99	149.670	25,45	115.525	19,65	28.873	4,91	588.031	100,00

sas que forman cada uno de los cuatro grupos de tamaño y de los cuatro de edad. Por esta razón se ha denominado como muestra *Control*.

Las otras dos muestras tienen por objetivo permitir el análisis de las diferencias y similitudes en los comportamientos de los grupos de tamaño y edad.

Para la determinación de los tamaños muestrales se utilizó el programa SOTAM (Sistema de Optimización para Tamaños de Muestra)¹, que permite obtenerlos imponiendo condiciones sobre la precisión en el muestreo y el nivel de significación² para la estimación, teniéndose la posibilidad de elegir entre los diversos resultados el que se considere más ajustado para cada caso.

Para la determinación del tamaño muestral del grupo *Control* se optó, entre varios niveles de significación y precisión adecuados, por los de 5 y de 2,5 por 100. El tamaño muestral obtenido es de 7.893 empresas, el cual se distribuye entre los diferentes estratos formados por el cruce de los criterios tamaño y edad, de tal forma que se reproduzcan los porcentajes relativos de dichos estratos en la población de empresas con información completa en la Base SABI, véase cuadro 2.2.

Para el cálculo de las muestras que incorporan los criterios de tamaño y edad se ha tomado para el nivel de significación y precisión el valor

CUADRO 2.2
MUESTRA DE EMPRESAS DE LA BASE SABI SEGÚN TAMAÑO
Y EDAD (*CONTROL*)

<i>Edad/Empleo</i>	<i>>10 años</i>	<i>De 6 a 10 años</i>	<i>De 3 a 6 años</i>	<i>< 3años</i>	<i>Total</i>
Menor de 10	2.553	1.592	1.307	343	5.795
De 10 a 49	1.127	366	217	41	1.750
De 50 a 249	211	42	22	3	279
Mayor de 249	55	8	4	1	69
Total	3.946	2.009	1.551	388	7.893

¹ Desarrollado por el profesor MANZANO ARRONDO de la Universidad de Sevilla y que se encuentra disponible a través de la red (<http://www.personal.us.es/vmanzano/distribuye/programas.htm>).

² El nivel de significación es la probabilidad de que la muestra seleccionada nos dé un intervalo de confianza que no contenga al verdadero valor del parámetro. Es habitual considerar un nivel de significación comprendido entre el 1 y el 5 por 100. El error de muestreo o precisión al estimar un parámetro es la semi-longitud del intervalo de confianza obtenido para dicho parámetro, es decir, $e_{\theta} = z_{\alpha/2} \sqrt{\text{Var}(\hat{\theta})}$ donde $z_{\alpha/2}$ es la medida de riesgo asociada al nivel de significación prefijado (PÉREZ, 1999).

del 5 por 100. De esta forma, se han obtenido, respectivamente, unos tamaños muestrales totales de 7.223 y 7.796 empresas. No obstante, al contrario que para la anterior muestra *Control*, dichos tamaños no se han calculado condicionados a los porcentajes que corresponden a cada estrato respecto al total de las empresas que figuran en el cuadro 2.1, sino que en esta ocasión el cálculo del número de elementos a incluir en cada muestra se ha realizado a partir del número de empresas que existen en cada estrato de tamaño o edad. Es decir, para el caso de la muestra estratificada por tamaño se ha calculado respecto a cada uno de los totales parciales que figuran en la última columna del cuadro 2.1, y para la segmentada por edad respecto a aquellos que aparecen en la última fila. Los porcentajes obtenidos figuran en el cuadro 2.3 y con ellos se ha condicionado la determinación de los cuatro tamaños muestrales para el criterio tamaño y de los cuatro para la edad.

Siguiendo el método indicado, los ocho tamaños muestrales obtenidos y su distribución entre las distintas dimensiones y antigüedades de las empresas son los que se presentan en el cuadro 2.4.

A partir de los distintos tamaños muestrales calculados, la selección de las empresas concretas que forman los estratos de cada muestra se realiza mediante un procedimiento aleatorio que otorga a cada elemento de la población total de 588.031 empresas (en el caso de la muestra *Control*) y a cada uno de las poblaciones parciales (en el caso de las muestras de tamaño y edad) la misma probabilidad de ser escogido, siendo el único condicionante el porcentaje que a su estrato le corresponde en el total.

CUADRO 2.3

POBLACIONES EN PORCENTAJES DE EMPRESAS DE LA BASE SABI SEGÚN TAMAÑO Y EDAD

<i>Población</i>	<i>Empleo/Edad</i>	<i>>10 años</i>	<i>De 6 a 10 años</i>	<i>De 3 a 6 años</i>	<i>< 3 años</i>	<i>Total</i>
Según tamaño	Menor de 10	44,05	27,48	22,56	5,915	100
	De 10 a 49	64,38	20,94	12,37	2,314	100
	De 50 a 249	75,72	15,05	8,00	1,228	100
	Mayor de 249	80,48	11,86	6,47	1,192	100
<i>Población</i>	<i>Edad/Empleo</i>	<i>Menor de 10</i>	<i>De 10 a 49</i>	<i>De 50 a 249</i>	<i>Mayor de 249</i>	<i>Total</i>
Según edad	>10 años	64,69	28,56	5,35	1,40	100
	De 6 a 10 años	79,26	18,24	2,09	0,41	100
	De 3 a 6 años	84,31	13,96	1,44	0,29	100
	< 3 años	88,45	10,45	0,88	0,21	100

CUADRO 2.4
MUESTRAS DE EMPRESAS DE LA BASE SABI SEGÚN
TAMAÑO Y EDAD

<i>Muestra</i>	<i>Empleo/Edad</i>	<i>>10 años</i>	<i>De 6 a 10 años</i>	<i>De 3 a 6 años</i>	<i>< 3 años</i>	<i>Total</i>
Según tamaño	Menor de 10	877	547	449	118	1.991
	De 10 a 49	1.268	412	244	46	1.970
	De 50 a 249	1.381	275	146	22	1.824
	Mayor de 249	1.157	171	93	17	1.438
<i>Muestra</i>	<i>Edad/Empleo</i>	<i>Menor de 10</i>	<i>De 10 a 49</i>	<i>De 50 a 249</i>	<i>Mayor de 249</i>	<i>Total</i>
Según edad	>10 años	1.285	567	106	28	1.986
	De 6 a 10 años	1.565	360	41	8	1.974
	De 3 a 6 años	1.658	275	28	6	1.966
	< 3 años	1.654	195	17	4	1.870

4. ESQUEMA DE RATIOS Y METODOLOGÍA DE ANÁLISIS CUANTITATIVO

Una vez diseñadas las muestras y seleccionadas las empresas que las componen se ha procedido a extraer los datos particulares de cada una de ellas para realizar los correspondientes análisis descriptivos, contrastes estadísticos y modelos econométricos.

En concreto, los datos de cada una de las empresas que se han considerado relevantes para este estudio son los ratios y el modelo de agrupación recogidos en el recuadro 2.1.

La elección del conjunto de ratios y sus agrupaciones depende del objetivo final perseguido. En el presente trabajo se pretende diferenciar entre los comportamientos de grupos de empresas en función de su tamaño y edad para ver cómo inciden estos aspectos en la actividad emprendedora. Para ello, se toma como base la clasificación de ratios del modelo propuesto por MAROTO (2005) en la serie de trabajos sobre las pyme con forma societaria³. En este trabajo, el modelo ha sido adaptado parcialmente a los datos y propuestas de definición de los ratios de

³ Esta serie viene realizándose desde el año 2002 por encargo de la Dirección General de Política de la Pequeña y Mediana Empresa (Ministerio de Industria, Turismo y Comercio). Los datos que utiliza son también extraídos de los estados financieros depositados por las empresas en los Registros Mercantiles, en este caso a partir de la base proporcionada por el Centro de Procesos de Estadísticos de los Colegios de Registradores de España. Las ediciones más recientes de estos trabajos se encuentran disponibles en <http://www.ipyme.org/IPYME/es-ES/EstadisticasPublicacionesEstudios/Publicaciones/>.

RECUADRO 2.1
RATIOS CALCULADAS PARA CADA EMPRESA Y AÑO

<i>Ratios de composición de la estructura del capital económico:</i>
Activo circulante ÷ Activo fijo
Existencias ÷ Activo total
<i>Ratios de composición de la estructura del capital financiero:</i>
Fondos propios ÷ Pasivo total
Exigible a corto plazo ÷ Financiación permanente
<i>Ratios de composición de la estructura del capital-coberturas:</i>
Financiación permanente ÷ Inmovilizado total
Acreedores a corto plazo ÷ Deudores
<i>Ratios indicadores de gestión de actividad:</i>
Valor añadido ÷ Cifra de negocios
Valor añadido ÷ Inmovilizado material
Fondo de maniobra ÷ Cifra de negocios
<i>Ratios indicadores de gestión de costes:</i>
Consumos de explotación ÷ Cifra de negocios
Gastos de personal ÷ Cifra de negocios
Gastos financieros ÷ Cifra de negocios
<i>Ratios indicadores de gestión-rotaciones:</i>
Acreedores comerciales ÷ Consumos de explotación
Deudores ÷ Cifra de negocios
<i>Ratios indicadores financieros de liquidez y solvencia:</i>
Tesorería ÷ Acreedores a corto plazo
Activo total ÷ Exigible total
Gastos financieros ÷ Resultado neto del ejercicio + Gastos financieros
<i>Ratios indicadores de rentabilidad:</i>
Rentabilidad financiera: (Rentabilidad económica · Apalancamiento financiero · Efecto Fiscal)
Rentabilidad económica: (Margen de explotación · Rotación del activo)
<i>Ratios indicadores de empleo y productividad:</i>
Productividad: valor añadido por trabajador

la base SABI. Asimismo, se han eliminado algunas ratios por estar correlacionadas con otras o por su menor relevancia para el estudio de la actividad emprendedora.

Estas veinticuatro ratios han sido calculadas con los registros de la base SABI, que contiene los estados financieros para el conjunto de las

22.912 [7.893 (muestra *Control*) + 7.223 (muestras de tamaño) + 7.796 (muestras de edad)] empresas que componen las muestras seleccionadas en cada uno de los ocho años del periodo 1997 a 2004. En conjunto se ha calculado el valor de 4.399.104 ratios (24 ratios · 8 años · 22.912 empresas).

En la obtención de los cálculos de las ratios se han tomado los valores de las partidas contables que figuran en la base SABI a fecha de cierre de cada ejercicio económico.

Este procedimiento de cálculo es adecuado conforme a la teoría y práctica financiera para todas las ratios que contrastan partidas de la misma naturaleza, es decir, variables fondo con variables fondo y variables flujo con variables flujo. En los casos de las ratios de rentabilidad, por razones operativas, se ha establecido este mismo procedimiento, aunque la relación esté formada por una variable flujo (numerador) y otra variable fondo (denominador). Ello, para el caso de las empresas supervivientes estudiadas significa, en términos generales, una leve infravaloración de los resultados de la rentabilidad económica y financiera para los distintos grupos de tamaño y edad, que no causa perjuicio al análisis. Ha de tenerse en cuenta, de un lado, que este análisis es de amplios grupos agregados de empresas, no de empresas individuales, y de otro, que el objeto del análisis es la observación de comportamientos y tendencias disímiles entre los grupos de edad y tamaño propuestos. Todos estos grupos se ven afectados por la mencionada infravaloración en igual medida. A modo de ejemplo, para la muestra *Control* (de 7.893 empresas) y el periodo 2000-2004, la diferencia media es del 0,26 por 100 en el caso de la rentabilidad financiera y del 0,48 por 100 en la rentabilidad económica.

A) Metodología del análisis estadístico

En los trabajos efectuados de revisión y depuración de datos se detectó cierta discontinuidad temporal en la obtención de valores de las ratios, existiendo casos para los que no en todos los años se obtenían resultados. Asimismo, se observaron casos en los que existía una amplia dispersión de los valores obtenidos. Lo primero no recomienda un tratamiento estadístico basado en datos de panel, y lo segundo aconseja que la información disponible sea tratada mediante el cálculo de medidas estadísticas de posición central que no se vean afectadas por la existencia de valores extremos. La medida que permite esto es la mediana de la distribución. Igualmente, es posible utilizar la media, pero sólo en el caso que para su cálculo se elimine de las muestras un suficiente número

de empresas que para cada ratio presentasen los valores superiores e inferiores.

Por ello, los estadísticos descriptivos (media, mediana y desviación típica) de las diversas ratios, que contienen las fichas que figuran en el cd adjunto, anexo 2, para cada uno de los grupos de tamaño y edad se han calculado tanto a partir de las medianas como de las medias de los valores tomados por las ratios para cada empresa durante el conjunto de años del periodo de estudio. En el caso de las medias, para evitar el problema expuesto de distorsión que provoca la existencia de valores extremos, se limpiaron las dos colas de la distribución de frecuencias de los valores medios tanto del 5 por 100 de los valores mayores como del mismo porcentaje de los menores.

De manera similar, en las representaciones gráficas de la evolución temporal de cada ratio se ha reflejado la mediana y media (sin valores extremos) para cada año de los valores de las ratios que para todas las empresas de las correspondientes muestras se han obtenido a partir de los datos disponibles en la base SABI. Estos gráficos se han realizado para reflejar conjuntamente la evolución de los distintos grupos de tamaño y edad y observar las diferencias y similitudes en su comportamiento temporal. En el estudio realizado en el capítulo V, principalmente se utiliza el estadístico mediana tanto para el análisis gráfico como estadístico. Los gráficos realizados para ver la evolución de los ratios en función de las medias únicamente se utilizan cuando existen dudas con el empleo de la mediana.

Las diversas pruebas y tests estadísticos utilizados para detectar el disímil comportamiento según tamaño o edad para cada ratio se han aplicado sobre las medianas de los valores de las ratios para cada empresa en los ocho años considerados. En concreto, los contrastes utilizados han sido el análisis de la varianza, el test de Scheffé y las pruebas no paramétricas de Kolmogorov-Smirnov y de Kuskal-Wallis. El resultado de estas pruebas estadísticas se ha complementado con el obtenido mediante la aplicación de técnicas econométricas de regresión con variables ficticias o *dummies*.

El análisis de la varianza es el procedimiento estadístico que se ha utilizado como método para contrastar la igualdad entre medias de varios grupos procedentes de poblaciones normales. Si con su aplicación se rechaza la hipótesis de que las medias sean iguales, entonces, se admite que al menos dos de las medias son distintas. En este caso hay que utilizar algún procedimiento que permita detectar qué medias son diferentes y cuáles no. Con este fin, el procedimiento que se utiliza en el presente trabajo es el test de Scheffé de comparaciones múltiples que consiste

en formular un contraste sobre una combinación lineal del número necesario de medias poblacionales.

El contraste de Kolmogorov-Smirnov se ha utilizado para comprobar si las distintas ratios estudiadas siguen o no una distribución normal. La finalidad perseguida es comprobar la hipótesis teórica de normalidad, que debe cumplirse para dar mayor fiabilidad al análisis de la varianza a la hora de verificar diferentes comportamientos entre las dimensiones o edades de las empresas contenidas en las muestras. Si como resultado de aplicar esta prueba no paramétrica no se puede afirmar la normalidad de la distribución de las ratios es necesario que los resultados obtenidos con el análisis de la varianza se traten de corroborar con el contraste de Kruskal-Wallis.

El contraste de Kruskal-Wallis es una versión no paramétrica del test de la F resultante del análisis de la varianza y sirve para investigar la hipótesis de que los grupos considerados por tamaño o por edad siguen la misma distribución o no. De esta forma, se trata de confirmar el diagnóstico de posibles disimilitudes en el comportamiento que proporcione el análisis de la varianza y la discriminación entre grupos debida al test de Scheffé.

B) Metodología de análisis econométrico

Como se ha indicado, objetivo esencial del trabajo es la identificación de comportamientos diferentes entre las distintas edades y tamaños de las empresas respecto a las tres variables de desempeño empresarial (rentabilidad financiera, rentabilidad económica y productividad). Por ello, además de calcular para cada ratio diversos contrastes estadísticos y las principales medidas descriptivas de cada tramo de edad y tamaño, se ha realizado un análisis de la varianza mediante una regresión con la inclusión de variables ficticias o *dummies*.

El modelo permite observar el valor de los coeficientes de las variables ficticias obtenidos de la regresión, así como su significación estadística (recuadro 2.2). De esta forma, aquellos coeficientes estadísticamente significativos de la variables *dummies* señalan la existencia de un comportamiento diferente con respecto al resto de categorías consideradas y, análogamente, la carencia de significación pone de manifiesto comportamientos comunes entre las diversas categorías.

Los resultados de este análisis econométrico permiten ratificar y cuantificar, o poner en cuestión, la existencia de comportamientos similares o distintos a los observados en los análisis estadísticos, de tal forma que las conclusiones a las que se llegue estén sustentadas técnicamente de la manera más robusta posible.

RECUADRO 2.2
MODELO DE REGRESIÓN ESPECIFICADO

$$Y_i = \alpha + \beta_1 \text{dummy_ta}_{1i} + \beta_2 \text{dummy_ta}_{2i} + \beta_3 \text{dummy_ta}_{3i} \\ [o, \text{alternativamente, } + \beta_1 \text{dummy_ed}_{1i} + \beta_2 \text{dummy_ed}_{2i} + \beta_3 \text{dummy_ed}_{3i}] + \varepsilon$$

Donde:

Y_i : Variables explicadas o endógenas, Rentabilidad financiera, Rentabilidad económica y Productividad, sucesivamente.

α : Constante del modelo.

Variables cualitativas o ficticias relativas al tamaño de la empresa o a su edad. En todos los casos, se introducen $h-1$ variables ficticias para evitar la colinealidad perfecta. Por ello, se introducen 3 variables ficticias para el tamaño e igualmente para la edad.

$$\text{Dummy_ta}_h = \begin{cases} 1 & \text{si la empresa seleccionada es de tamaño } h \\ 0 & \text{en caso contrario} \end{cases}$$

dummy_ta Variables de tamaño: Dummy_ta_{hi} identifica las observaciones de las empresas de diferente tamaño, según el número de empleados, donde h es un valor entre 1 y 4, cuyos valores son micro-empresas, empresas pequeñas, empresas medianas y empresas grandes.

o

dummy_ed: $\text{Dummy_ed}_h = \begin{cases} 1 & \text{si la empresa seleccionada pertenece a la edad } h \\ 0 & \text{en caso contrario} \end{cases}$

Variables de edad: Dummy_ed_{hi} identifica las observaciones de las empresas clasificadas según su edad o edad: empresas menores de tres años, entre tres y seis años, entre seis y diez años y de más de diez. Por tanto, h tomaría valores entre 1 y 4.

β : Coeficientes resultantes del modelo de regresión.

ε : Perturbación aleatoria del modelo.

Las estimaciones de los modelos de regresión se han hecho con el método de mínimos cuadrados ordinarios, manteniendo las mismas especificaciones en todos los casos. La selección de variables, descartando las no significativas, se ha realizado según el método de selección hacia atrás.

Tal y como se señala en el recuadro 2.2, en la especificación del modelo de regresión se utilizan $h-1$ variables ficticias para evitar la colinealidad perfecta. Por tanto, una vez realizadas las estimaciones con las $h-1$ variables ficticias, se deben recuperar los coeficientes estructurales a partir de los coeficientes *alfa* y *beta* estimados. Para esta recuperación se sigue el proceso que se expone en el recuadro 2.3, teniendo en cuenta

RECUADRO 2.3

PROCESO DE OBTENCIÓN DE LOS COEFICIENTES ESTRUCTURALES DE LAS VARIABLES FICTICIAS DE LA REGRESIÓN

Modelo estimado

$$Y_i = \alpha + \beta_1 \text{dummy_ta}_{1i} + \beta_2 \text{dummy_ta}_{2i} + \beta_3 \text{dummy_ta}_{3i} \\ [o, \text{alternativamente, } \beta_1 \text{dummy_ed}_{1i} + \beta_2 \text{dummy_ed}_{2i} + \beta_3 \text{dummy_ed}_{3i}] \\ + \varepsilon$$

Modelo con coeficientes estructurales

$$Y_i = d + a_1 \text{dummy_ta}_{1i} + a_2 \text{dummy_ta}_{2i} + a_3 \text{dummy_ta}_{3i} + a_4 \text{dummy_ta}_{4i} \\ [o, \text{alternativamente, } a_1 \text{dummy_ed}_{1i} + a_2 \text{dummy_ed}_{2i} + a_3 \text{dummy_ed}_{3i} + a_4 \\ \text{dummy_ed}_{4i}] \\ + \varepsilon$$

Por tanto, igualando:

$$\alpha + \beta_1 \text{dummy_ta}_{1i} + \beta_2 \text{dummy_ta}_{2i} + \beta_3 \text{dummy_ta}_{3i} \\ [o, \text{alternativamente, } \beta_1 \text{dummy_ed}_{1i} + \beta_2 \text{dummy_ed}_{2i} + \beta_3 \text{dummy_ed}_{3i}] \\ = d + a_1 \text{dummy_ta}_{1i} + a_2 \text{dummy_ta}_{2i} + a_3 \text{dummy_ta}_{3i} + a_4 \text{dummy_ta}_{4i} \\ [o, \text{alternativamente, } a_1 \text{dummy_ed}_{1i} + a_2 \text{dummy_ed}_{2i} + a_3 \text{dummy_ed}_{3i} + a_4 \\ \text{dummy_ed}_{4i}]$$

Donde debe verificarse que:

$$\text{dummy_ta}_{1i} + \text{dummy_ta}_{2i} + \text{dummy_ta}_{3i} + \text{dummy_ta}_{4i} = 1 \\ [o, \text{alternativamente, } \text{dummy_ed}_{1i} + \text{dummy_ed}_{2i} + \text{dummy_ed}_{3i} + \text{dummy_ed}_{4i} = 1]$$

Tomando como referencia la última variable en cada conjunto de dummy, se deduce:

$$\text{dummy_ta}_{4i} = 1 - \text{dummy_ta}_{1i} - \text{dummy_ta}_{2i} - \text{dummy_ta}_{3i} \\ [o, \text{alternativamente, } \text{dummy_ed}_{4i} = 1 - \text{dummy_ed}_{1i} - \text{dummy_ed}_{2i} - \text{dummy_ed}_{3i}]$$

Por tanto,

$$Y_i = (d + a_4) + (a_1 - a_4) \text{dummy_ta}_{1i} + (a_2 - a_4) \text{dummy_ta}_{2i} + (a_3 - a_4) \text{dummy_ta}_{3i} \\ [o, \text{alternativamente, } (d + a_4) + (a_1 - a_4) \text{dummy_ed}_{1i} + (a_2 - a_4) \text{dummy_ed}_{2i} + \\ (a_3 - a_4) \text{dummy_ed}_{3i}] \\ + \varepsilon$$

Dado que los modelos son equivalentes, se tiene que:

$$\alpha = (d + a_4) \quad B_1 = a_1 - a_4 \quad B_2 = a_2 - a_4 \quad B_3 = a_3 - a_4 \quad a_1 + a_2 + a_3 + a_4 = 0$$

que las sumas de coeficientes estructurales que afectan a las variables ficticias deben ser cero. Estas restricciones son necesarias para interpretar los coeficientes del modelo estructural —notado en el recuadro 2.3 con letras latinas— como expresivos de diferencias respecto a una media global (MELLE, 2001).

5. METODOLOGÍA DE ANÁLISIS CUALITATIVO: ENCUESTA

El trabajo de encuesta tiene dos condicionantes previos establecidos en el Convenio de investigación. De un lado, debían ser encuestadas las mismas empresas que lo fueron en el estudio del *Primer Informe de Emprendedores*; es decir, 458 encuestas dirigidas a ese mismo número de empresarios que respondieron entonces. De otro, con el objeto de mejorar el conocimiento de los nuevos emprendedores, debía también realizarse la misma encuesta dirigida a 150 (aprox.) nuevos empresarios de entre 0 y 3 años. En el cd adjunto, anexo 1, de este trabajo se presenta el cuestionario de la «Encuesta sobre emprendedores en España 2006».

Este diseño de la estrategia de encuesta permite ver los cambios cualitativos que se han experimentado por parte de los mismos empresarios entre 2002 y 2006, así como las diferencias respecto a los nuevos empresarios (de menos de 3 años).

Al mismo tiempo sería un anhelo de la investigación que este diseño se acomodase a la estructura del tejido empresarial español (en concreto, tomando como *proxy* a la estructura de la población de empresas con información completa en la Base SABI, 2005).

Sin embargo, este acomodo no es fácilmente alcanzable debido a los dos condicionantes indicados arriba, pero también a las modificaciones en el tiempo de la estructura empresarial y, principalmente, a la aleatoriedad que siempre acompaña al efectivo éxito en la obtención de respuestas a los cuestionarios enviados.

El trabajo de campo de la encuesta ha consistido en el envío postal (por parte de Dephimática, empresa especializada subcontratada por SERVILAB) de los cuestionarios destinados a las empresas encuestadas en 2002 y de los cuestionarios a empresas creadas entre 2003 y 2005. El objetivo de respuesta era, de un lado, obtener el mayor número de cuestionarios cumplimentados por parte de los empresarios repetidores y, de otro, conseguir, al menos, 150 respuestas completas de los nuevos empresarios. El trabajo de campo realizado por la empresa contratada se resume en el recuadro 2.4.

RECUADRO 2.4

TRABAJO DE CAMPO DE LA ENCUESTA A EMPRENDEDORES 2005

El trabajo de encuesta se realizó entre la última semana del mes de febrero y finales de junio; en este tiempo se realizaron todas las tareas habituales de campo, tales son:

1. Edición de los cuestionarios. Se definieron dos modelos de cuestionarios para la captura de información ajustándolos a la reiteración de la encuesta para las unidades ya encuestadas y para las nuevas empresas a encuestar.

2. Depuración de los datos de localización y búsqueda de teléfonos de contacto.
3. Preparación y realización del *mailing*, envío postal de la encuesta a la dirección de las empresas.
4. Emisión de llamadas de campo y localización de la persona responsable de la cumplimentación.
5. Atención de llamadas de las empresas para solucionar dudas o información aclaratoria.
6. Gestión de incidencias, como localización de empresas y de casos de cierres u otros.
7. Recepción de cuestionarios cumplimentados.

Se ejecutaron diversas acciones a los efectos de conseguir una respuesta adecuada de las empresas que integran la muestra a encuestar. Estas tareas se resumen a continuación:

- Contacto y confirmación de los datos de identificación, razón social, domicilio y NIF, de las empresas de la muestra.
- Identificación de la persona responsable de la cumplimentación del cuestionario, en su defecto de una persona de contacto, interlocutor válido.
- Confirmación de la recepción de la encuesta en la dirección de la empresa.
- Reenvío de cuestionarios por correo o por fax de las unidades que indicaron que no recibieron el cuestionario del primer envío, o en su caso porque los informantes prefieren recibirlo directamente a su nombre. Se estima que se han hecho reenvíos a casi el 50 por 100 de la muestra.
- Realización de 14.200 llamadas de recuperación hasta conseguir una respuesta favorable dentro del periodo de encuestación.
- Recoger las posibles incidencias que se detectan en las llamadas de campo, empresas cerradas, ilocalizables, etc.
- Atender las reclamaciones y dudas para apoyar la obtención de respuestas válidas.

Las principales argumentaciones de los informantes de las empresas que no respondieron a la encuesta, tras las reiteradas reclamaciones hechas por los agentes de campo, fueron:

- Empresas que se niegan rotundamente a cumplimentar encuestas que no sean «obligatorias».
- Respuesta negativa de hecho sin ningún tipo de argumentación; es el caso de los informantes que no atendieron a la llamada por diversos motivos, tales como:
 - La persona responsable de la cumplimentación de la encuesta en general se corresponde con una persona que ocupa un puesto de dirección y responsabilidad en la empresa, en un porcentaje alto de los casos sólo se ha podido hablar con una persona distinta que hizo de filtro de las llamadas, secretaria, recepcionista, sin conseguir una comunicación directa con el responsable. Ante esta situación se reenviaron faxes de reclamación dirigidos directamente a la persona responsable y nunca accedieron a nuestra demanda. En resumen sobre este punto el inconveniente encontrado es que la persona responsable y quien era la que debía completar el cuestionario era un alto cargo de la empresa y «no dispone de tiempo para atender a este tipo de demandas».
 - La falta de interés o poca importancia que las empresas han manifestado. Siempre se repiten las argumentaciones en trabajos de campo de estas características, falta de tiempo, de interés: «el empresario está para atender a otros trabajos y problemas de gestión de la empresa».
 - La persona responsable nunca puede atender nuestra llamada, está reunido, está fuera de la empresa, no se puede poner. Es un alto cargo de una empresa peque-

ña y tiene demasiadas ocupaciones. En resumen dicen: «no puede responder el empresario» y no hay persona que pueda responder en su nombre. Según las indicaciones especificadas en cuestionario de que la persona que tiene que responder al cuestionario es el «empresario principal o promotor» algunas empresas respondiendo a esta cuestión negaron toda posibilidad de rellenar la encuesta argumentando que esta persona no puede dedicarse a estos asuntos y no han querido delegar en otra la cumplimentación de la encuesta.

Los resultados alcanzados en el periodo de encuestación comprenden a un total de 507 encuestas con datos validados y contrastados con los informantes.

Tras la realización del trabajo de campo de encuesta se procedió a la grabación de cuestionarios, su codificación y validación.

Elaborado a partir del «Informe del trabajo de campo de la encuesta a emprendedores 2005» de Dephimática, empresa responsable de la ejecución del trabajo de campo.

A) Encuesta a los emprendedores ya encuestados en 2002

De los 458 emprendedores-empresas para los que se obtuvo respuesta en la encuesta en el año 2002, ahora, en 2006, se dispone de datos, en la base SABI, para 452. Pero no todas sus empresas se encuentran en funcionamiento. Cuatro de ellas se encuentran en fase de liquidación, ocho están en fase de extinción y doce no se encuentran en la actualidad registradas con el mismo nombre. De estas últimas, a través de información en la red se ha podido detectar que simplemente ha habido un cambio de forma jurídica, o denominación social, obteniendo evidencias de continuidad en seis de ellas. Por tanto, son 434 emprendedores los destinatarios de la encuesta diseñada al efecto para buscar la continuidad con la realizada en el año 2002.

Tras más de cuatro meses en campo, de los cuestionarios dirigidos a empresarios repetidores se han obtenido respuestas completas de 285 empresarios, lo que significa un 65,67 por 100 ($285/434 \cdot 100$) de las enviadas. Si se compara para cada uno de los 16 estratos las respuestas recibidas en 2006 con las empresas encuestadas en 2002, se observa que los totales para los distintos tamaños y edades en todos los casos superan el 50 por 100, y que en conjunto se ha obtenido respuesta de 62,2 por 100 de las empresas.

En el cuadro 2.5 se presenta la distribución del número de respuestas recibidas a los cuestionarios enviados en 2002 y en 2006, así como el porcentaje de repetición obtenido.

Los anteriores porcentajes permiten pensar que el índice de repetición en la contestación de los cuestionarios es suficiente para poder dar

CUADRO 2.5
DISTRIBUCIÓN DE RESPUESTAS DE EMPRESAS ENCUESTADAS EN 2002 Y 2006 Y PORCENTAJE DE REPETICIÓN

	< 3 años		De 3 a 6 años		De 6 a 10 años		> 10 años		Total		
	2002	2006	2002	2006	2002	2006	2002	2006	2002	2006	
Menor de 10	N.º respuestas	1	0	26	13	70	41	80	47	177	101
	% de repetición	0,0		50,0		58,6		58,7		57,1	
De 10 a 49	N.º respuestas	7	5	43	22	68	50	114	79	232	156
	% de repetición	71,4		51,2		73,5		69,3		67,2	
De 50 a 249	N.º respuestas	3	2	4	3	9	2	26	14	42	21
	% de repetición	66,7		75,0		22,2		53,8		50,0	
Mayor de 249	N.º respuestas	1	0	0	0	3	1	3	6	7	7
	% de repetición	0,0		100,0		33,3		200,0		100,0	
Total	N.º respuestas	12	7	73	38	150	94	223	146	458	285
	% de repetición	58,3		52,1		62,7		65,5		62,2	

validez a los resultados que se obtengan del estudio cualitativo de la continuidad de la actividad de los emprendedores. Esta apreciación se confirma si se observa que para la variable tamaño las contestaciones de la encuesta de 2006 para cada segmento de edad replica de forma muy cercana la estructura de contestaciones de 2002, tal como se puede comprobar en el cuadro 2.6.

Igualmente, como se ofrece en el cuadro 2.7, si se comparan las respuestas de los empresarios a los que ha vuelto a encuestarse en 2006, las obtenidas en 2002 y la estructura empresarial actual (SABI 2005) con relación a los respectivos totales de emprendedores que han contestado o al número de empresas de la Base SABI ahora seleccionadas, se desprende que las respuestas logradas se adecuan a las obtenidas cuatro años atrás para los distintos segmentos de tamaño de empresas propuestos. La única excepción es la del grupo más numeroso, las pequeñas empresas, que aumentan su presencia en las respuestas en un 5 por 100 más a costa del también numeroso grupo de las microempresas. Al mismo tiempo, estos dos mismos grupos de empresas que, de manera conjunta, representan el 90 por 100 de las empresas encuestadas en las dos sucesivas oleadas, son las que incluso en más de este porcentaje se hallan presentes en la estructura empresarial española (96 por 100). La circunstancia de que se hayan obtenido más respuestas de pequeñas empresas que de microempresas —a partir de que en la estructura del tejido español son mayoritarias ampliamente las segundas frente a las primeras— se compensa por la amplia presencia de respuestas de microempresas en el segmento de la encuesta referido a las empresas de nueva creación (menos de tres años) que se presenta posteriormente.

Por otra parte, por la relativa baja presencia de medianas y, principalmente, grandes empresas en la estructura empresarial española, el ligero mayor porcentaje de respuestas en estos tramos de tamaño favorece la validez de las conclusiones que induzca el análisis.

Al contrario de lo que ocurre desde la perspectiva del tamaño para los distintos grupos de empresas, desde la de la edad, las respuestas obtenidas de los empresarios a los que se ha repetido encuesta no se ajustan suficientemente a la estructura empresarial. Es más, esta situación empeora respecto a las respuestas obtenidas en el año 2002, tal como se refleja también en el cuadro 2.7. La explicación de este hecho estriba en la dificultad intrínseca que conlleva la tarea de encuestación. Para que se pudiera replicar la situación de 2002 y que al mismo tiempo se aproximase a la estructura empresarial de 2005 habría que haber conseguido un 100 por 100 de respuestas a los cuestionarios enviados y, además,

CUADRO 2.7

A: ESTRUCTURA EMPRESARIAL ESPAÑOLA (SABI, 2005)
 B: RESPUESTAS RECIBIDAS EN 2006
 C: RESPUESTAS RECIBIDAS EN 2002

	<i>Empleo/Edad</i>	<i>>10 años</i>	<i>De 6 a 10 años</i>	<i>De 3 a 6 años</i>	<i>Total</i>
A	Menor de 10	34%	21%	17%	73%
	De 10 a 49	15%	5%	3%	23%
	De 50 a 249	3%	1%	0%	4%
	Mayor de 249	1%	0%	0%	1%
	Total	53%	27%	21%	100%
B (2006)	Menor de 10	28%	7%	0%	35%
	De 10 a 49	42%	11%	2%	55%
	De 50 a 249	6%	1%	1%	7%
	Mayor de 249	2%	0%	0%	2%
	Total	78%	19%	2%	100%
C (2002)	Menor de 10	18%	16%	6%	40%
	De 10 a 49	26%	15%	10%	50%
	De 50 a 249	6%	2%	1%	9%
	Mayor de 249	1%	1%	0%	1%
	Total	50%	33%	16%	100%

aumentar la muestra en aquellos segmentos de edad y tamaño sub-representados en la encuesta de 2002. Todo ello conduce a que puedan existir algunos sesgos en las conclusiones que se desprendan de la encuesta referidos al conjunto empresarial.

No obstante, dado que los estratos más relevantes (microempresas y pequeñas) están suficientemente representados en las respuestas recibidas para empresas repetidas y, como se verá a continuación, estos tamaños son los más representados en la encuesta realizada a nuevas empresas, cabe admitir que las conclusiones del análisis cualitativo reflejen, globalmente, dicha realidad.

B) Encuesta a nuevos emprendedores

Los nuevos emprendedores encuestados son aquellos cuyas empresas han sido constituidas en los últimos tres años. Para diseñar la muestra de empresas a las que enviar el cuestionario diseñado se ha tomado de la población total de SABI (compuesta por 976.598 empresas) aquellas 53.370 sociedades que se han constituido entre el 1 de enero de 2003

y 31 de diciembre de 2005. Este número de compañías se estratificó por los criterios de tamaño y rama de actividad.

Para la segmentación por tamaño se diferenció entre las cuatro dimensiones ya definidas anteriormente. Para el criterio de rama de actividad se han definido siete Agrupaciones CNAE para las que se han agrupado los códigos de actividad CNAE a un nivel de desagregación de dos dígitos de la siguiente forma: agraria (códigos del 1 al 5), extractiva, energía y agua (códigos del 10 al 14 y del 40 al 41), manufactura tradicional (códigos del 15 al 23, del 26 al 28 y 36 y 37), manufactura intermedia y avanzada (códigos 24, 25 y del 29 al 35), construcción (código 45), comercio (códigos del 50 al 52) y resto de servicios (resto de códigos).

Tras la segmentación, el número de empresas para las que existe información válida se reduce a 28.873. Este conjunto de empresas se reparte en los 24 estratos resultantes de la forma que se muestra en el cuadro 2.8.

El objetivo de respuestas a obtener de la encuesta a las empresas constituidas entre el 1 de enero de 2003 y 31 de diciembre de 2005 es de 150 cuestionarios válidos, por lo que se estimó que, con base a las experiencias previas en este tipo de estudios, era necesario enviar el cuestionario a una cifra cercana al millar de empresas. Estos envíos se repartieron proporcionalmente al porcentaje que representa cada uno de los 24 estratos sobre el total de la población de empresas de la Base SABI de menos de tres años, segmentada por tamaño y rama de actividad según se refleja en el propio cuadro 2.8.

Si se sigue el reparto proporcional a las empresas presentes en la Base SABI, el número de cuestionarios a enviar a las empresas con más de 50 trabajadores sería muy reducido. Esta circunstancia podría provocar que el número de respuestas que se obtuvieran de los dos tamaños superiores de empresas fuera nulo o, aun no siendo nulo, podrían no reflejar las opiniones y características del conjunto de los emprendedores de esos estratos. Por esta razón, el número de cuestionarios enviados a las medianas empresas se aumentó a 50 y el de grandes empresas a 20, con lo que el número total de cuestionarios enviados pasa a ser de 1.057. La selección de las empresas concretas a las que se envía el cuestionario se realizó de forma aleatoria, condicionándose únicamente por el porcentaje que cada uno de los estratos representa sobre el total de la población de empresas (28.814) de menos de tres años con información completa en SABI.

Del envío de estos cuestionarios, tras los cuatro meses de trabajo de campo, se obtuvieron 222 respuestas completas, lo que significa un exce-

CUADRO 2.8
POBLACIÓN DE EMPRESAS DE LA BASE SABI DE MENOS DE TRES AÑOS SEGÚN TAMAÑO Y ACTIVIDAD
Y REPARTO DE CUESTIONARIOS A ENVIAR

Empleo Sector	Menor de 10		De 10 a 49		De 50 a 249		Mayor de 249		Total	
	SABI	Encuesta	SABI	Encuesta	SABI	Encuesta	SABI	Encuesta	SABI	Encuesta
Agrario	329	11	58	2	3	0	0	0	390	13
Extractivas	84	3	16	1	0	0	1	0	101	4
Manufactura tradicional	2.284	79	385	13	24	1	8	0	2.701	94
Manufactura avanzada	469	16	91	3	13	0	4	0	577	20
Construcción	4.799	166	1.119	39	75	3	0	0	5.993	208
Comercio	7.062	245	448	16	32	1	6	0	7.548	262
Servicios	10.461	361	894	32	107	4	42	2	11.504	400
Total	25.488	882	3.011	105	254	9	61	2	28.814	998

lente nivel de éxito para este tipo de empresarios sin experiencia previa en esta encuesta. Detalle del número de cuestionarios enviados y recibidos repartidos entre cada uno de los 24 estratos se muestra en el cuadro 2.9.

CUADRO 2.9
CUESTIONARIOS ENVIADOS Y RECIBIDOS A EMPRESAS
DE MENOS DE TRES AÑOS

Sector/Empleo	Cuestionarios a menores de 10		Cuestionarios a entre 10 y 49		Cuestionarios a entre 50 y 249		Cuestionarios a mayores de 249		Cuestionarios totales	
	Enviados	Recibidos	Enviados	Recibidos	Enviados	Recibidos	Enviados	Recibidos	Enviados	Recibidos
Agrario	11	6	2	0	0	0	0	1	13	7
Extractivas	3	1	1	0	0	0	0	0	4	1
Manufactura tradicional	79	19	13	4	6	2	0	0	98	25
Manufactura avanzada	16	4	3	3	0	0	0	0	19	7
Construcción	166	20	39	12	17	3	0	0	222	35
Comercio	245	43	16	10	6	0	0	2	267	55
Servicios	361	66	32	20	22	5	20	1	435	92
Total	882	159	105	49	50	10	20	4	1.057	222

Estos buenos resultados han hecho que con las respuestas obtenidas de los empresarios con menos de tres años de historia se reproduzca bastante fielmente la estructura empresarial que representa su población, tal como queda reflejado en el cuadro 2.10. Se observa que el mayor grado de respuesta de las empresas de más de 10 trabajadores y el menor de las microempresas no tiene por qué causar perjuicio a las conclusiones del análisis, dado que para este último segmento de edad empresarial las microempresas constituyen la inmensa mayoría del entramado empre-

CUADRO 2.10
ESTRUCTURA EMPRESARIAL DE EMPRESAS DE MENOS DE TRES
AÑOS EN SABI Y ESTRUCTURA DE RESPUESTAS RECIBIDAS

Empleo/Edad	< 3 años. Estructura SABI	< 3 años. Encuesta 2006
Menor de 10	88%	72%
De 10 a 49	10%	22%
De 50 a 249	1%	5%
Mayor de 249	0%	2%
Total	100%	100%

sarial (algo más del 88 por 100). Por tanto, en esta situación es necesario aumentar la presencia de respuestas para los segmentos sub-representados en la población, ya que el reducido número de respuestas puede introducir sesgos más graves a la hora de abordar el análisis.

C) Metodología de análisis de la encuesta

Una vez finalizado el trabajo de campo de la encuesta se procedió a la explotación de los datos obtenidos, distinguiendo entre los correspondientes a empresarios-empresas repetidores y empresarios-empresas nuevos. Los primeros permiten comparaciones con los que se obtuvieron en la encuesta realizada en 2002, mientras que los segundos facilitan el establecimiento de diferencias en función de la antigüedad del proyecto empresarial. Por otra parte, con el objeto de profundizar en algunos campos de la actividad emprendedora, la encuesta de 2006 incluía 17 preguntas que no se habían formulado en la de 2002, por lo que convenía dar un tratamiento específico a estos nuevos datos.

De esta manera, el análisis descriptivo de la encuesta se desarrolla en tres sucesivas secuencias. En la primera se ofrece de nuevo, cuatro años después del primer Informe de la actividad emprendedora en España, el panorama más reciente de la misma actividad, con la presentación del perfil y caracterización de los empresarios y su evolución respecto a 2002; en la segunda se presenta un estudio factorial y de *cluster* que identifica sus factores determinantes y las agrupaciones homogéneas de empresarios según el modo y compromiso con que desarrollan la actividad emprendedora; y en la tercera, se indaga en nuevas facetas de los emprendedores, particularmente en comportamientos empresariales a largo plazo asociados a actividades emergentes e industrias creativas.

En concreto, en el capítulo III se realiza el análisis de la primera secuencia, que recorre los epígrafes contenidos en la encuesta. Es decir, caracterización del empresario, datos de la empresa a 31 de diciembre de 2005, investigación y desarrollo (I+D), fuentes de financiación de la empresa, crecimiento y dimensión del proyecto empresarial, factores de éxito y problemas encontrados en la creación de la empresa y valoraciones sobre el espíritu emprendedor en España. Para una mejor comprensión de los resultados de la encuesta, en el análisis descriptivo estos epígrafes han sido reordenados en los siguientes apartados: 1) introducción; 2) caracterización del empresario español: principales rasgos; las motivaciones personales; la gran decisión; 3) las características del proyecto empresarial: dimensión, edad y crecimiento; decisiones empresariales; estructura laboral; 4) la valoración de los resultados empresa-

riales: las razones que explican el éxito y supervivencia; las causas del fracaso, y 5) conclusiones.

La segunda secuencia del análisis de la encuesta se basa en todas las preguntas-respuestas que contienen una valoración por parte del empresario mediante el procedimiento de «escalas Likert» o similares (muy importante, ..., nada importante, etc.). Con estos resultados valorativos y algunos otros de caracterización del empresario o de la empresa se ha creado un conjunto de 37 variables sobre las que se realiza el estudio de factores que caracterizan el espíritu empresarial y su relación con las *performances* de las empresas. Esta secuencia se recoge en la primera parte del capítulo IV.

Finalmente, la tercera secuencia del análisis utiliza la mayor parte de los resultados de las preguntas nuevas incorporadas a la encuesta de 2006, con la finalidad de profundizar en aspectos ligados a la creatividad competitiva de los emprendedores. Se recoge en la segunda parte del capítulo IV.

