



fedea

Fundación de
Estudios de
Economía Aplicada

**Profesionalización e internalización de
las decisiones económico-financieras**
Evidencia empírica sobre las Pymes
manufactureras en España

Marco Celentani

Universidad Carlos III y FEDEA

Miguel García-Posada

FEDEA

Teresa Molina

FEDEA

Colección Estudios Económicos
03-2010

ISSN 1988-785X
www.fedea.es

Resumen

En el presente trabajo estudiamos la profesionalización y la internalización de la toma de decisiones económico-financieras para las empresas manufactureras españolas utilizando la Encuesta de Estrategias Empresariales para el periodo 1990-2006. El objetivo del trabajo es hacer un estudio descriptivo de tales decisiones y encontrar posibles asociaciones con variables económicas y financieras.

Los patrones principales que se encuentran son los siguientes:

- *Empresas pequeñas.* Tienen los menores niveles de profesionalización. Tanto la profesionalización como la internalización aparecen positivamente relacionadas con una medida de tamaño que depende de número de trabajadores, ventas, pasivo y valor añadido.
- *Empresas medianas.* Tienen niveles de profesionalización intermedios. La profesionalización no tiene ningún determinante significativo. La internalización tiene una asociación positiva significativa con el tamaño.
- *Empresas grandes.* Tienen niveles de profesionalización e internalización elevados. Tanto la profesionalización como la internalización están asociados significativamente con el tamaño.
- A lo largo del periodo estudiado se observan las siguientes *tendencias*:
 - *Empresas pequeñas:* Un incremento en la profesionalización y una notable tendencia hacia externalización parcial y, sobre todo, total.
 - *Empresas medianas:* Un cierto aumento en la profesionalización y una ligera tendencia hacia la externalización total o parcial.
 - *Empresas grandes:* Ninguna variación apreciable en la profesionalización y una destacable tendencia hacia la externalización parcial pero no a la total.

Estos desarrollos son coherentes con la disminución en los tipos de interés reales y un aumento en la competencia entre las entidades de crédito, que a la vez que aumentan el retorno de la profesionalización reducen el retorno de la internalización, trasladando el coste de la reducción de la opacidad informacional de las Pymes a las entidades de crédito.

Tabla de contenidos

1. Introducción
 - 1.1 Estadísticas descriptivas de la muestra
 - 1.2 Clasificación de las empresas
 - 1.3 Medición de profesionalización e internalización.
2. Variables y metodología
3. Las Pymes
 - 3.1 Correlaciones parciales
 - 3.2 Tau b de Kendall
 - 3.3 Análisis
 - 3.4 Discusión
4. Comparación entre pequeñas y medianas empresas
5. Comparación entre Pymes y empresas grandes
6. La evolución temporal profesionalización e internalización
7. Conclusiones

Apéndice: Análisis de los determinantes de la profesionalización y de la internalización con una clasificación alternativa de las empresas

Referencias

Recuadros metodológicos

- A. El test chi-cuadrado de Pearson
- B. El Análisis de Componentes Principales
- C. El coeficiente V de Cramer
- D. Correlaciones parciales
- E. La Tau-b de Kendall
- F. Kernels y el “diagrama de caja y bigotes”
- G. El Análisis de Conglomerados y el algoritmo de k-medias

1. Introducción

El objetivo de este trabajo es estudiar la elección de profesionalizar y posiblemente internalizar las decisiones económico-financieras en las pequeñas y medianas empresas en España. Para ello empleamos la base de datos de la Encuesta Sobre Estrategias Empresariales (ESEE) que elabora la Fundación SEPI (antes Fundación Empresa Pública). Esta es una encuesta anual que se realiza a un panel de empresas representativo de las industrias manufactureras. El objetivo principal de la ESEE es recabar información sobre las estrategias de las empresas, es decir, sobre aquellas decisiones que adoptan las empresas sobre los instrumentos de competencia a su alcance en el territorio nacional.

La ESEE dispone de datos de las empresas españolas desde 1990 hasta 2006 con una periodicidad anual y su población de referencia son las empresas con 10 o más trabajadores. La ESEE recaba también información de algunas empresas con menos de 10 trabajadores, pero en este caso se pierde la representatividad de la industria manufacturera.

1.1 Estadísticas descriptivas de la muestra

Aunque la ESEE es anual, la pregunta de la encuesta clave para nuestro estudio sólo se plantea cada 4 años. Por lo tanto para llevar a cabo este análisis utilizamos datos de 1990, 1994, 1998, 2002 y 2006. Con ello contamos con aproximadamente 2000 empresas por año para un total de 9571 observaciones y 4128 empresas. La distribución de las observaciones se resume en las Tablas 1 y 2.

Tabla 1: Número de observaciones por año.

Año	Número	%
1990	2,188	22.9
1994	1,876	19.6
1998	1,776	18.6
2002	1,708	17.9
2006	2,023	21.1
Total	9,571	100.0

Tabla 2: Número de observaciones que las empresas aportan a la muestra.

Nº Observaciones	Empresas		Observaciones	
	Número	%	Número	%
1	1,846	44.7	1,846	19.3
2	750	18.2	1,500	15.7
3	504	12.2	1,512	15.8
4	427	10.3	1,708	17.9
5	601	14.6	3,005	31.4
Total	4128	100	9571	100

1.2 Clasificación de las empresas

En el resto del estudio utilizaremos la clasificación convencional de la Comisión Europea. Esta se basa en tres criterios: número de empleados, volumen de negocio y balance general, según se resume en la Tabla 3.

Tabla 3: Clasificación del tamaño de las empresas (Comisión Europea, 2003)

Categoría	Nº Empleados	Y (Volumen de Negocio	O Balance General)
Micro	< 10	< 2 millones €	< 2 millones €
Pequeña	< 50	< 10 millones €	< 10 millones €
Mediana	< 250	< 50 millones €	< 43 millones €

Las empresas que no cumplen ninguna de las tres condiciones expuestas en la Tabla 3 se denominan grandes.

Esta clasificación nos obliga a desechar aquellas observaciones para los que no tenemos alguna de estas tres variables. Para el análisis utilizamos por tanto un total de 8843 observaciones. La Tabla 4 resume la distribución de estas observaciones.

Tabla 4: distribución por tamaño de empresa y año de la muestra.

año	micro		pequeñas		medianas		grandes		Total
	número	%	número	%	número	%	número	%	número
1990	40	2.3	808	46.7	345	20.0	536	31.0	1,729
1994	108	6.1	768	43.5	383	21.7	508	28.8	1,767
1998	85	4.9	776	45.0	418	24.3	445	25.8	1,724
2002	70	4.2	719	43.3	393	23.7	480	28.9	1,662
2006	64	3.3	898	45.8	531	27.1	468	23.9	1,961
Total	367	4.2	3,969	44.9	2,070	23.4	2,437	27.6	8,843

Como se puede observar las empresas micro proporcionan solo 367 observaciones. Su escasa representatividad nos induce por tanto a excluirlas de nuestro análisis. Alrededor de un 45% de las empresas son pequeñas, mientras que las medianas constituyen un 24 % y las grandes un 28%.

Es importante destacar que la clasificación de la Comisión Europea agrupa empresas en cuatro conjuntos distintos. Esta clasificación es obviamente conveniente porque introduce una nomenclatura convencional. Pero la utilidad de la clasificación depende de que exista suficiente heterogeneidad entre empresas que pertenecen a distintos grupos y suficiente homogeneidad entre empresas que pertenecen a un mismo grupo. Como es obvio las nociones de homogeneidad y heterogeneidad no son unívocas y pueden depender del uso que se quiera hacer de la clasificación. Por esta razón hemos realizado un análisis de robustez, proponiendo una clasificación alternativa

que se basa en las mismas variables que las de la Comisión Europea y que sin embargo se realiza aplicando métodos puramente estadísticos (Análisis de Componentes Principales y Análisis de Conglomerados) al conjunto de observaciones. Los resultados de esta clasificación alternativa son muy parecidos a los de la clasificación de la Comisión Europea. En particular, una vez eliminadas las empresas con menos de 10 trabajadores, se distinguen tres grupos distintos, como en la clasificación de la Comisión Europea. Algo más de un 35 % de las empresas clasificadas como pequeñas según la Comisión Europea pasan a ser medianas con la clasificación alternativa. Asimismo, poco más de un 40% de las empresas clasificadas como medianas según la Comisión Europea pasan a ser grandes con la clasificación alternativa. Sin embargo, a pesar de estos cambios, las principales conclusiones que se obtienen con ambas clasificaciones son las mismas. Por esta razón hemos decidido mantener la denominación convencional de la Comisión Europea y presentar en el Apéndice los resultados con la clasificación alternativa.

1.3 Medición de profesionalización e internalización de las decisiones económico-financieras

En la sección 17 del apartado B, Clientes y Proveedores, del cuestionario de la ESEE se formula la siguiente pregunta:

Indique, para los siguientes servicios exteriores, cual ha sido la práctica de su empresa:

- *No utiliza*
- *Realizado por la empresa*
- *Contratado parcialmente*
- *Contratado en su totalidad.*

Entre los servicios para los que se plantea la pregunta se encuentran los "Servicios de asesoría económico-financiera". La práctica totalidad de las empresas (casi el 99%) responde a esta pregunta. La idea central de este trabajo es que estas observaciones permiten estudiar la decisión de profesionalizar la toma de decisiones económico-financieras, así como el grado de internalización de estas decisiones. Más concretamente nuestras conjeturas son las siguientes:

- Una respuesta de "No utiliza" indica una *escasa profesionalización* en la toma de decisiones económico-financieras, mientras que cualquiera de las otras tres respuestas indica un *mayor profesionalización*.
- Una respuesta de "Realizado por la empresa" indica una *internalización alta* en la toma de decisiones económico-financieras, una respuesta de "Contratado parcialmente" indica una

internalización intermedia y una respuesta de “Contratado en su totalidad indica una *internalización baja*.

En el resto del trabajo nos referiremos por brevedad simplemente a *profesionalización* y a *internalización* sin especificar que nos referimos a la toma de decisiones económico-financieras.

La Tabla 5 resume la distribución de respuestas total y por tamaño.

Tabla 5: distribución por tamaño de “Servicios de asesoría económico-financiera”

	Servicio de Asesoría Económico-financiera									
	No utiliza		Contratado en su totalidad		Contratado parcialmente		Realizado por la empresa		Total	
	número	%	número	%	número	%	número	%	número	% muestra
micro	86	24.0	150	41.8	69	19.2	54	15.0	359	97.8
pequeñas	748	19.3	1,370	35.4	876	22.6	879	22.7	3,873	97.6
medianas	260	12.7	345	16.8	476	23.2	969	47.3	2,050	99.0
grandes	215	8.9	190	7.8	534	22.0	1,490	61.3	2,429	99.7
Total	1,309	15.0	2,055	23.6	1,955	22.4	3,392	38.9	8,711	98.5

Aunque luego analizaremos esta relación de forma más rigurosa, la Tabla 5 ilustra que hay una relación positiva entre tamaño por un lado y profesionalización e internalización por el otro. El porcentaje de empresas que contesta “No utiliza” decrece a medida que el tamaño es mayor (desde el 24% de las micro hasta el 8,9% de las grandes). Para la respuesta “Realizado por la empresa” el porcentaje de empresas aumenta con el tamaño (desde el 15% de las micro hasta el 61,3% de las grandes).

La evolución temporal de las respuestas se resume en la Figura 1. En ella podemos ver cómo el porcentaje de empresas que responden “No utiliza” decrece a medida que pasan los años, y lo mismo ocurre con el porcentaje de empresas que responden “Realizado por la empresa”. Es decir, la tendencia en los años de la muestra ha sido hacia la profesionalización y la externalización.

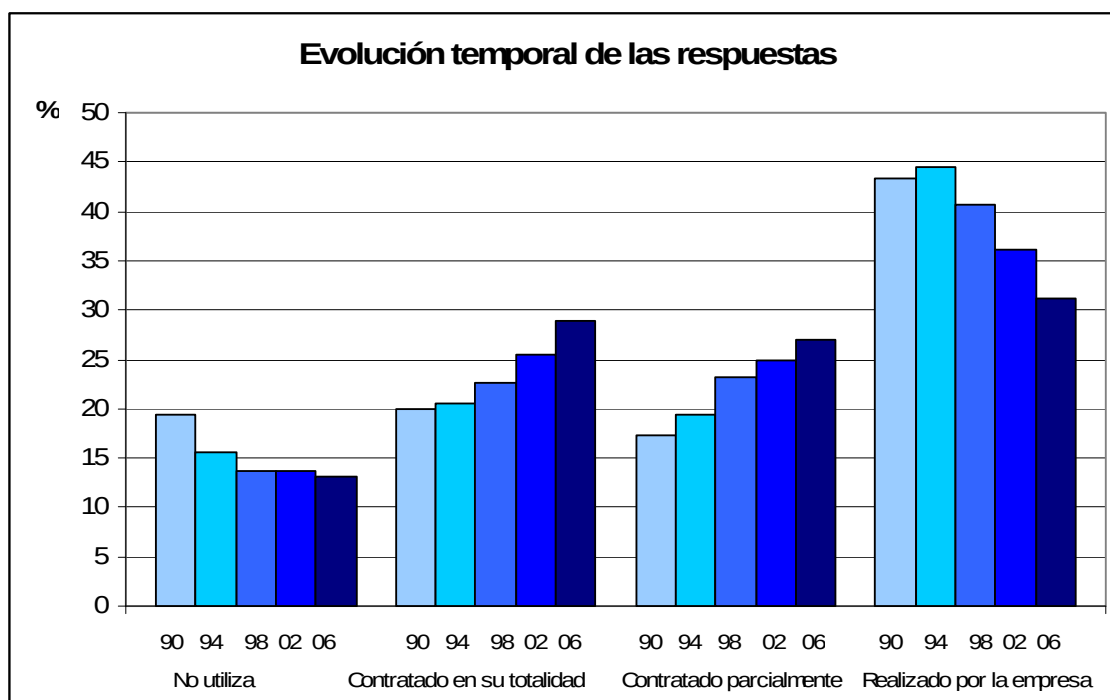


Figura 1

La Tabla 5 y la Figura 1, así como el análisis que expondremos a continuación describen un escenario coherente y por lo tanto sugieren que las respuestas al cuestionario de la ESEE arrojan luz sobre profesionalización e internalización en las Pymes manufactureras españolas. Pero es importante no olvidar que las variables que utilizamos para medir profesionalización e internalización no están exentas de errores de medida.

Para empezar existen tres fuentes principales de respuestas erróneas:

1. Respuestas erróneas inintencionadas;
2. Respuestas erróneas intencionadas;
3. Respuestas aleatorias.

Sobre la primera fuente obsérvese en la Tabla 5 que una proporción elevada de empresas, el 15%, contesta que no utiliza servicios de asesoría económico-financieras y que la proporción de empresas se reduce según aumenta el tamaño, con un 12,7% de empresas medianas y un 8,9% de empresas grandes. Si estos porcentajes parecen elevados es posible pensar que el planteamiento de la pregunta produce un sesgo a contestar “No utiliza”. La razón es que la pregunta se plantea en una sección dedicada a servicios exteriores y es posible que una empresa con un elevado grado de internalización conteste “No utiliza” en vez que “Realizados por la empresa” ya que entiende que no utiliza servicios *exteriores*. Si fuera así, entonces las estimaciones de empresas con un elevado

nivel de internalización (alrededor de un 40% en la muestra) podrían estar sesgadas a la baja debido al planteamiento de la pregunta en el cuestionario.

Sobre la segunda, no parece existir ningún incentivo para mentir intencionadamente en esta pregunta, ya que no hay ninguna respuesta que obligue a contestar a más preguntas y que por lo tanto aumente la carga de trabajo de la persona o personas encargadas de rellenar el cuestionario.

Sobre la tercera fuente de error, hemos testado dicha posibilidad mediante un test estadístico, el chi-cuadrado de Pearson. El resultado es rotundo: las respuestas no son aleatorias, para cualquier nivel de confianza.

Finalmente también es importante destacar que incluso respuestas no erróneas sólo proporcionan una aproximación a los niveles de profesionalización o de internalización, porque reflejan tan sólo las percepciones subjetivas de quien contesta al cuestionario.

Recuadro metodológico 1: el test chi-cuadrado de Pearson

El test chi-cuadrado de Pearson testea la hipótesis de que dos variables presentadas en una tabla de frecuencias (como la Tabla 5) son independientes.

Si disponemos de una tabla de frecuencias de r filas y c columnas con dos variables, y nuestras hipótesis son:

H_0 : las variables son independientes vs H_a : existe relación entre las variables.

El estadístico es:

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

donde $E_{ij} = \frac{R_i \cdot C_j}{n}$ es el número de observaciones de la celda (i,j) esperado bajo la hipótesis nula H_0 ; R_i es el total de observaciones de la fila i -ésima; C_j es el total de observaciones de la columna j -ésima; n es el total de observaciones; O_{ij} es el número de observaciones de la celda (i,j) .

Bajo H_0 , el estadístico sigue una distribución chi-cuadrado con $(r-1) \times (c-1)$ grados de libertad. Su intuición es sencilla: si el número de observaciones de cada celda (O_{ij}) diverge bastante del número esperado bajo la hipótesis nula de independencia (E_{ij}), el estadístico será mayor que el valor crítico de la distribución $\chi^2_{(r-1)(c-1)}$, por lo que rechazaremos dicha hipótesis y concluiremos que las dos variables están relacionadas.

En nuestro caso, hemos usado este test para la Tabla 5, para ver si existía alguna relación (aparte de la puramente aleatoria) entre los posibles tamaños de las empresas y las 4 posibles respuestas. El p -valor obtenido es 0, es decir, rechazamos la hipótesis nula de independencia entre las dos variables para cualquier nivel de confianza. Esto equivale a concluir que las respuestas no han sido dadas al azar, sino que dependen del tamaño de la empresa.

En el resto del presente informe analizaremos la relación entre profesionalización e internalización—medidas por la variable antes descrita— y los posibles determinantes de las mismas. En otras palabras, estudiaremos qué características se asocian a profesionalización e internalización así como la manera en la que estas asociaciones han evolucionado a lo largo de los años de la muestra. Inicialmente nuestro análisis se centrará en el conjunto de Pymes para luego diferenciar entre empresas pequeñas y medianas y finalmente contrastar los resultados obtenidos con los que se obtienen para empresas grandes.

2. Variables y metodología.

La ESEE proporciona mucha información sobre las empresas encuestadas y en particular sobre sus datos económicos y financieros. Esto permite estudiar los potenciales determinantes de profesionalización e internalización. En lo que sigue estudiaremos los siguientes determinantes potenciales:

- *Determinantes económicos:* valor añadido; gasto en I+D; sector; número de empleados; ventas; pasivo total; edad de la empresa;
- *Determinantes financieros:* fondos ajenos sobre pasivo; fondos ajenos a c/p de entidades de crédito sobre pasivo; fondos ajenos a l/p con entidades de crédito sobre pasivo; fondos de entidades de crédito sobre el total de los ajenos, coste medio de la deuda a l/p con entidades de crédito.

Las variables en valor han sido expresadas todas en euros del 2004. La última variable, el coste medio real de la deuda a l/p con entidades de crédito, requiere unas aclaraciones sobre su construcción. Las empresas reportan en el cuestionario el coste medio nominal de la deuda con entidades de crédito según el balance del año. Como es obvio, hemos eliminado las observaciones de dicha variable cuando la variable fondos ajenos a l/p con entidades de crédito sobre pasivo es igual a 0. Luego, hemos convertido el coste medio en real restándole la tasa de inflación de su año correspondiente, medida por la tasa de crecimiento del IPC.

El análisis de estos posibles determinantes se enfrenta con tres problemas principales:

- *Multicolinealidad:* Muchas de las variables señaladas anteriormente están fuertemente correlacionadas entre sí. Son bien sabidos los problemas econométricos que la multicolinealidad conlleva, como la pérdida de poder de los tests de significatividad, debida a errores estándar de los coeficientes anormalmente grandes.

- *Variables omitidas*: Realizar un análisis sencillo que se centra en pocos posibles determinantes puede producir conclusiones erróneas. Por ejemplo, en un análisis bivariante podríamos llegar a la conclusión de que existe una fuerte relación positiva entre gasto en I+D y profesionalización. Sin embargo, dicha relación puede ser espuria, si omitimos las variables que miden el tamaño de la empresa y dichas variables tienen una fuerte correlación tanto con el gasto en I+D como con la profesionalización.
- *Endogeneidad* (causalidad bidireccional): a priori, no está claro si la causalidad corre sólo en una dirección en la relación entre algunos de los determinantes de la profesionalización y ésta. Por ejemplo, ¿como debería interpretarse una asociación positiva entre peso de los fondos ajenos sobre el pasivo y profesionalización? ¿Es el mayor uso de fondos ajenos el que causa la profesionalización al hacerla más rentable? ¿Es la profesionalización la que causa un mayor uso de fondos ajenos haciendo más fácil y más ventajosa su captación? ¿O ambas?

Para afrontar el problema de la multicolinealidad hemos recurrido a una técnica estadística llamada Análisis de Componentes Principales (ACP). El ACP explota las correlaciones entre las diversas variables para agregarlas en un número menor de otras variables, llamadas componentes principales. Por construcción las diversas componentes son ortogonales entre sí, es decir, tienen correlación nula.

Para afrontar el problema de las variables omitidas teníamos que recurrir a alguna técnica de análisis multivariante con la que se pudiera medir la relación entre profesionalización e internalización y un posible determinante, una vez descontado el efecto de los otros determinantes. Una solución obvia es realizar una regresión múltiple. Sin embargo, bien sabidas son las consecuencias perversas que la endogeneidad de los regresores —uno de los problemas de nuestros datos— tienen sobre este tipo de estimación. En el caso en cuestión es muy difícil encontrar variables instrumentales adecuadas, que es el remedio típico para este problema.

Por todo lo anterior parece prudente realizar análisis descriptivos que no requieren hacer hipótesis de exogeneidad ni asumir relaciones de causalidad. Esto quiere decir que en vez de intentar validar empíricamente una teoría de profesionalización e internalización de decisiones económico-financieras, nos proponemos sencillamente realizar un análisis descriptivo que identifique los principales determinantes de profesionalización e internalización y que por lo tanto facilite su predicción. Las dos técnicas descriptivas que hemos utilizado y que se describen brevemente a continuación son las correlaciones parciales y la Tau b de Kendall.

- La correlación parcial entre dos variables no es más que su correlación directa, una vez que se descuenta el efecto de todas las otras variables sobre ellas.

- La Tau b de Kendall es una medida de correlación para variables ordinales. Una variable ordinal es aquella variable cuyos valores se pueden ordenar de menor a mayor, pero entre los cuales no se puede determinar una distancia exacta. Por ejemplo, una variable altura que toma valores “bajo”, “alto” y “muy alto” es una variable ordinal, pues sus valores expresan una jerarquía, pero no permiten determinar si la distancia en centímetros entre “bajo” y “alto” es la misma que entre “alto” y “muy alto”.

La estrategia de análisis que seguiremos se puede resumir de la siguiente forma:

1. Reducir el número de potenciales determinantes de profesionalización e internalización mediante el ACP.
2. Estudiar la relación entre profesionalización e internalización y los determinantes mediante:
 - (a) Correlaciones parciales entre cada posible respuesta (p.ej. “No utiliza”, convertida en una variable dummy que toma el valor 1 si la empresa opta por ella y 0 en caso contrario) y las componentes principales obtenidas en 1.
 - (b) Correlaciones ordinales Tau b de Kendall entre la variable “internalización” (que toma el valor 1 para “Contratado en su totalidad”, 2 para “Contratado parcialmente”, 3 para “Realizado por la empresa”) y las mismas componentes principales.

Recuadro metodológico 2: el Análisis de Componentes Principales (ACP)

Introducción

El análisis de componentes principales (ACP) es un método que permite reducir un conjunto de variables originales en otro conjunto de variables, llamado conjunto de componentes principales, de menor dimensión. Su idea básica es que, si las variables originales están muy correlacionadas entre sí, existe información redundante que puede ser desechada. Una de sus ventajas respecto a otras medidas de agregación, tales como sumas ponderadas, es que permite agregar variables que se encuentran expresadas en diferentes unidades de medida (como es nuestro caso). Las componentes principales se caracterizan por ser combinaciones lineales de las variables originales y por estar incorreladas entre sí. Es lógico que exijamos a las componentes principales estar incorreladas entre sí, para que así cada nueva componente informe de aspectos diferentes de la realidad.

Implementación

Aunque el ACP busca la reducción del número de variables de un determinado conjunto, es necesario reseñar que, en general, ello no puede ser realizado sin la pérdida de cierta cantidad de información. Por ello, en el terreno aplicado, siempre debemos tener en cuenta que esta pérdida de información sea tolerable. Existen varios criterios en la literatura, ninguno de ellos concluyente, para la selección del número de componentes. Comentaremos dos; ambos toman a la varianza como indicador de la información que aporta cada variable.

1. Criterio de Kaiser: este criterio utiliza un razonamiento intuitivo: deberían ser seleccionadas aquellas componentes que aportan más información (varianza) que las variables originales. Dado que las variables originales son tipificadas en el ACP, su varianza es 1. Por tanto, se deberían tomar aquellas componentes con autovalor (que por construcción es igual a su varianza) mayor que 1. Sin embargo, este razonamiento requiere una matización. Mientras que las componentes principales están incorreladas (toda la información que aportan es única), las variables originales están correlacionadas (parte de la información que aportan es redundante). Ello hace que el umbral de comparación suela relajarse a 0.8.

2. Criterio de la proporción de la varianza explicada y de las comunalidades: por construcción, la varianza total de las variables originales es la suma de las varianzas de las componentes principales, donde la varianza de la primera componente es mayor que la varianza de la segunda, la de la segunda que la de la tercera, etc.

Por tanto, un criterio de selección consiste en quedarse con un número de componentes suficiente como para que su varianza sea una proporción importante de la varianza total. Generalmente el umbral es un 70%. De forma complementaria, se considera que el ACP propuesto es válido si explica al menos una determinada proporción de la varianza de *cada* variable (también llamada comunalidad). En este caso el umbral que se considera es de un 50%.

Interpretación: matriz de componentes, cargas y rotación VARIMAX.

Una vez extraídas las componentes principales, es necesario saber cuáles son las variables originales que forman en mayor medida cada una de las componentes, para así poder interpretar qué hay detrás de cada una de las componentes. Esto se realiza mediante la matriz de componentes. Esta no es más que una matriz que muestra los coeficientes de correlación (también llamados cargas) entre las variables originales y las diversas componentes. Para cada componente buscaremos cuáles son las variables originales que están más correlacionadas con ella en valor absoluto, puesto que éstas serán las que más contribuyan a la explicación de dicha componente.

Tras esto, la rotación de componentes puede ser utilizada para facilitar la interpretación de los resultados. La rotación de componentes busca simplificar al máximo la matriz de componentes. La idea es obtener unas cargas más extremas (lo más cercanas posible al -1 o al 1) y otras cargas cercanas al 0. El método de rotación más utilizado en la literatura es el VARIMAX, el cual respeta la ortogonalidad de dichas componentes.

Medidas de adecuación muestral

Uno de los pocos supuestos que debe satisfacer la muestra para que el uso del ACP sea adecuado es que las variables presenten suficiente multicolinealidad. La medida de adecuación muestral más utilizada es el estadístico Kaiser-Meyer-Olkin (KMO). El KMO para la variable

$$i \text{ es: } KMO = \frac{\sum_{i \neq j}^J r_{ij}^2}{\sum_{i \neq j}^J r_{ij}^2 + \sum_{i \neq j}^J a_{ij}^2} \text{ donde } r_{ij} \text{ es el coeficiente de correlación de Pearson entre}$$

las variables i y j y a_{ij} es el coeficiente de correlación parcial entre las mismas variables. Su intuición es sencilla: si dos variables comparten un factor común, es decir, están altamente correlacionadas con otra variable, su coeficiente de correlación parcial estará cercano a cero ($a_{ij} \approx 0$) a pesar de que su coeficiente de correlación lineal sea alto, por lo que $KMO \approx 1$

Existe un KMO para cada variable, y un KMO global, el cual es:

$$KMO = \frac{\sum_{i \neq j}^J \sum_{i \neq j}^I r_{ij}^2}{\sum_{i \neq j}^J \sum_{i \neq j}^I r_{ij}^2 + \sum_{i \neq j}^J \sum_{i \neq j}^I a_{ij}^2}$$

La interpretación de los valores del estadístico es la siguiente:

Valor del KMO	Adecuación muestral al análisis de componentes principales
(0.9,1.0)	Espléndida
(0.8,0.89)	Meritoria
(0.7,0.79)	Mediana
(0.6,0.69)	Mediocre
(0.5,0.59)	Mala
(0,0.49)	Inaceptable

El umbral a partir del cual un ACP no es deseable depende en última medida del investigador; la literatura a veces se decanta por 0.6 y a veces por 0.5. En todo caso, valores por encima de 0.7 son siempre suficientes, mientras que valores por debajo de 0.5 no son nunca aceptables. Si el KMO global no alcanza el umbral deseado, lo que se hace es eliminar las variables con KMO individual más bajo, hasta que el global llegue a dicho umbral.

3. El caso de las Pymes

La tendencia temporal hacia una creciente profesionalización y una mayor externalización observada para el total de la muestra se confirma para el conjunto de las Pymes, como se puede observar en la Figura 2 a continuación.

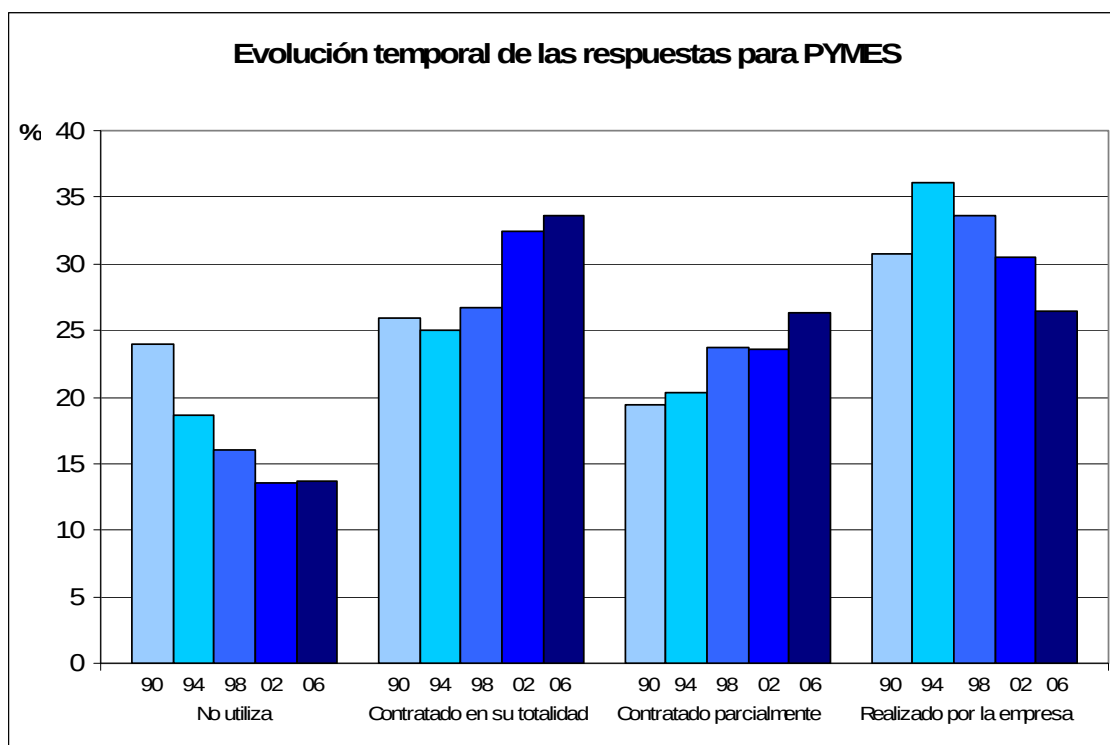


Figura 2

El primer paso de nuestro análisis es realizar un ACP para todos los posibles determinantes salvo el sector, dado que esta es una variable cualitativa que no puede ser estudiada mediante el ACP. Sin embargo, este ACP genera un KMO global (medida de adecuación muestral al ACP; para más información véase el recuadro metodológico 2) muy mediocre (0.5), e inaceptable para las variables financieras (alrededor de 0.2).

Por lo tanto, optamos por realizar dos ACPs en vez de uno sólo:

- Un ACP para las variables económicas: valor añadido, gasto en I+D, número de trabajadores, volumen de ventas, pasivo total, edad de la empresa.
- Un ACP para las variables financieras: estructura del pasivo (fondos ajenos sobre pasivo; fondos ajenos a c/p de entidades de crédito sobre pasivo; fondos ajenos a l/p de entidades de crédito sobre pasivo; fondos de entidades de crédito sobre el total de los ajenos); coste medio real de la deuda a l/p con entidades de crédito.

El ACP para las variables económicas genera un KMO “meritorio” (0.85), lo que indica que las variables en cuestión se prestan muy bien al ACP. Tanto los criterios de Kaiser como el de la proporción de varianza explicada y las comunales nos indican que hemos de retener tres componentes. Tras esto rotamos las tres componentes extraídas mediante el método VARIMAX, para mejorar su interpretación. La matriz de componentes resultante nos muestra que la primera componente está asociada a tres medidas de tamaño (número de trabajadores, volumen de ventas, pasivo total) y al valor añadido, con cargas para todas ellas alrededor de 0.5. La segunda está asociada únicamente al gasto en I+D, con una carga de 0.99. La tercera está asociada únicamente a la edad de la empresa, con una carga de casi 1. Esto quiere decir que es posible resumir las seis variables económicas originales en tres componentes, una que capta el tamaño y el valor añadido, otra que capta el gasto en I+D y otra que capta la edad de la empresa.

El ACP para el conjunto de variables financieras genera un KMO global inaceptable (alrededor de 0.17). La solución en este caso es eliminar del análisis “fondos ajenos a c/p de entidades de crédito sobre pasivo”, es decir la variable con KMO individual más bajo. Aún así el KMO global sigue quedando muy debajo de 0.5. Reiterando el proceso de eliminación de las variables con KMO individuales más bajos no produce mejora sustanciales. Esto quiere decir que las variables financieras no se prestan a ser resumidas mediante un ACP y será necesario estudiar cada una de ellas por separado.

Antes de proceder a las correlaciones parciales entre los posibles determinantes (algunos resumidos por el ACP) y profesionalización e internalización es importante analizar el posible impacto del sector al que pertenecen las empresas. Al ser una variable cualitativa el sector no puede ser incluido en el ACP, pero tampoco va a poder ser incluido en un análisis de correlaciones parciales. Para ver si hay una relación significativa entre sector y profesionalización e internalización recurrimos a dos estadísticos, el ya explicado test chi-cuadrado de Pearson y el coeficiente V de Cramer. El V de Cramer es una medida de asociación para variables cualitativas. Está acotado entre 0 y 1, donde 0 indica nula relación entre las variables y 1 perfecta relación entre ellas. Los valores de estos estadísticos para la relación entre sector y las respuestas a la pregunta sobre asesoría económico-financiera son:

$$Pearson\ chi^2(57) = 172.25 \quad Pr = 0.00$$

$$Cramér's\ V = 0.09$$

Aunque el test chi-cuadrado de Pearson rechaza la hipótesis nula de que las dos variables sean independientes, el valor del coeficiente V de Cramer presenta un valor muy pequeño que sugiere una relación muy débil. Esta conclusión es reforzada por el hecho que no hemos tenido en cuenta

otros factores que pueden ser comunes a ambas variables, como tamaño de la empresa, y que pueden estar provocando esta débil relación. En concreto, el V de Cramer para la relación entre tamaño y sector (0.22), muestra una relación mediana entre estas variables y corrobora esta intuición. Por lo tanto, no consideramos al sector un determinante significativo ni de la profesionalización ni de la internalización. Por esa razón, en los análisis posteriores no se llevará a cabo una desagregación por sectores.

Recuadro metodológico 3: el coeficiente V de Cramer.

El V de Cramer se define como:

$$V = \sqrt{\frac{\chi^2}{n(k-1)}}$$

donde $\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$ (es decir, el estadístico chi-cuadrado del test de Pearson, véase el recuadro metodológico 1) ; $K=\min(I,J)$ (es decir, el mínimo entre el número de filas y el número de columnas de la tabla de frecuencias); n es el número total de observaciones.

El V de Cramer está acotado entre 0 y 1, donde 0 indica nula relación entre las variables y 1 perfecta relación entre ellas.

3.1 Correlaciones parciales

El análisis de correlaciones parciales entre las respuestas a la pregunta sobre profesionalización e internalización se resume en las Tablas 6-9, que reportan la correlación parcial entre cada respuesta a la pregunta sobre asesoría económica-financiera y los posibles determinantes. Cada respuesta es una variable dummy donde 1=SI, 0=NO. La Tabla 6 muestra la relación entre la respuesta “No utiliza” (que interpretamos como escasa profesionalización) y los determinantes potenciales. Las Tablas 7-9 muestran la relación entre los determinantes potenciales y las otras posibles respuestas ordenadas de “Contratado en su totalidad” (internalización baja), a “Contratado parcialmente” (internalización intermedia) hasta “Realizado por la empresa” (internalización alta).

Es necesario resaltar que los potenciales errores de medida explicados con anterioridad nos hace tomar un 10% como el nivel de significación mínimo, dado que los errores estándar de los coeficientes serán mayores y también por tanto los p-valores. Además, es útil recordar que el

planteamiento de la pregunta puede concentrar la mayoría de los errores de medida en las respuestas “No utiliza” y “Realizado por la empresa”, ya que empresas que realizan internamente responden que no utilizan servicios externos.

En las Tablas 6-9, *1ª componente ACP* resume tamaño (número de trabajadores, volumen de ventas, pasivo total) y valor añadido, *2ª componente ACP* resume el gasto en I+D, y *3ª componente ACP* resume la edad, *fondos c/p e.c. sobre pasivo* son los “fondos a corto plazo con entidades de crédito sobre pasivo”; *fondos e.c. sobre ajenos* son los “fondos de entidades de crédito sobre el total de los ajenos”; *coste medio deuda l/p* es el “coste medio de la deuda a largo plazo con entidades de crédito”. Los coeficientes que son significativos como mínimo al 10% se resaltan en negrita.

Recuadro metodológico 4: las correlaciones parciales.

La correlación parcial entre dos variables no es más que su correlación directa, una vez que se descuenta el efecto de todas las otras variables sobre ellas. Así como la correlación simple es la versión de estadística descriptiva del coeficiente de regresión simple, las correlaciones parciales son la versión descriptiva de los coeficientes de regresión múltiple. La diferencia entre los dos es que en la primera no se necesitan supuestos sobre exogeneidad, mientras que en la segunda sí, así como las conclusiones a las que se llega. En el primer caso, hablaríamos de una *relación* entre x e y, teniendo en cuenta z, w, etc, mientras que en el segundo hablaríamos de que x *causa* y, teniendo en cuenta z, w. etc. Eso sí, el signo de los coeficientes de correlación parcial es el mismo que el de los coeficientes de una regresión múltiple, aunque por supuesto no su valor absoluto.

Para las variables x, y, z, el coeficiente de correlación parcial a_{xy} entre las variables x e y dado el factor común z es:

$$a_{xy} = \frac{r_{xy} - r_{xz} \cdot r_{yz}}{\sqrt{(1 - r_{xz}^2)(1 - r_{yz}^2)}}$$

donde r_{xy} es el coeficiente de correlación simple de Pearson.

De la anterior fórmula podemos ver que $a_{xy} = r_{xy}$ si y sólo si $r_{xz} = r_{yz} = 0$, es decir, si z no es un factor común para ambas variables. También de la misma fórmula se puede apreciar que a_{xy} y r_{xy} pueden tener signos diferentes.

En el caso de más 3 variables, la forma más sencilla de calcularlo es:

- regresar x sobre los factores comunes.
- regresar y sobre los factores comunes.
- calcular el coeficiente de correlación simple entre los residuos resultantes de ambas regresiones.

Su intuición es sencilla: por construcción, los residuos son aquella parte de la variable dependiente que no puede ser explicada por los regresores (en este caso, los factores comunes). Por lo tanto, mediante las dos regresiones “limpiamos” las variables de posibles factores comunes para luego poder calcular su correlación directa.

Tabla 6: Correlación parcial entre “No utiliza” y los posibles determinantes

Variable	Corr.	p-valor
<i>1ª componente ACP</i>	-0.06	0.00
<i>2ª componente ACP</i>	0.00	0.90
<i>3ª componente ACP</i>	0.01	0.44
<i>fondos ajenos sobre pasivo</i>	0.00	0.99
<i>fondos c/p e.c. sobre pasivo</i>	-0.02	0.34
<i>fondos l/p e.c. sobre pasivo</i>	0.00	0.96
<i>fondos e.c. sobre ajenos</i>	0.02	0.30
<i>coste medio deuda l/p</i>	0.04	0.01

(obs=3369)

Tabla 7: Correlación parcial entre “Contratado en su totalidad” y posibles determinantes

Variable	Corr.	p-valor
<i>1ª componente ACP</i>	-0.19	0.00
<i>2ª componente ACP</i>	-0.03	0.15
<i>3ª componente ACP</i>	0.00	0.84
<i>fondos ajenos sobre pasivo</i>	0.01	0.69
<i>fondos c/p e.c. sobre pasivo</i>	-0.01	0.73
<i>fondos l/p e.c. sobre pasivo</i>	0.00	0.99
<i>fondos e.c. sobre ajenos</i>	0.00	0.95
<i>coste medio deuda l/p</i>	-0.05	0.00

(obs=3369)

Tabla 8: Correlación parcial entre “Contratado parcialmente” y posibles determinantes

Variable	Corr.	p-valor
<i>1ª componente ACP</i>	0.05	0.01
<i>2ª componente ACP</i>	-0.01	0.64
<i>3ª componente ACP</i>	0.04	0.01
<i>fondos ajenos sobre pasivo</i>	0.03	0.09
<i>fondos c/p e.c. sobre pasivo</i>	-0.01	0.40
<i>fondos l/p e.c. sobre pasivo</i>	-0.01	0.67
<i>fondos e.c. sobre ajenos</i>	0.02	0.22
<i>coste medio deuda l/p</i>	-0.04	0.03

(obs=3369)

Tabla 9: Correlación parcial entre “Realizado por la empresa” y posibles determinantes

Variable	Corr.	p-valor
1ª componente ACP	0.21	0.00
2ª componente ACP	0.03	0.05
3ª componente ACP	-0.05	0.01
fondos ajenos sobre pasivo	-0.03	0.05
fondos c/p e.c. sobre pasivo	0.04	0.04
fondos l/p e.c. sobre pasivo	0.01	0.43
fondos e.c. sobre ajenos	-0.03	0.04
coste medio deuda l/p	0.05	0.00

(obs=3369)

3.2 Tau-b de Kendall.

Introducimos ahora una nueva variable que toma los siguientes valores dependiendo de las respuestas a la pregunta sobre asesoría económico-financiera:

Valor de la variable	Respuesta
No definido	No utiliza
1	Contratado en su totalidad
2	Contratado parcialmente
3	Realizado internamente

El resultado es una variable que mide el nivel de internalización en las empresas que profesionalizan la toma de decisiones económico-financieras, es decir en las empresas que no contestan “No utiliza”. Si no imponemos una métrica sobre esta variable que permita cuantificar y por lo tanto comparar las distancias entre 1 y 2 y entre 2 y 3, esta variable es una variable *ordinal* y su relación con los determinantes potenciales debe ser estudiada a través de la Tau b de Kendall. La ventaja de este procedimiento frente al anterior es que es más sintético, dado que resume la relación global entre cada determinante e internalización en un sólo número. Sin embargo, este procedimiento tiene una desventaja frente al anterior porque produce correlaciones (ordinales) *simples*, no parciales. Por ello, ciertas correlaciones entre un determinante y la *internalización* pueden verse distorsionadas por el efecto de un factor común en ambas. Los resultados se resumen en la Tabla 10.

Recuadro metodológico 4: la Tau b de Kendall.

La Tau b de Kendall es una medida de correlación ordinal entre dos variables X e Y, es decir, mide cuán parecido es el orden (*ranking*) de las observaciones realizado a partir de la variable X del orden realizado a partir de la variable Y. Por ejemplo, si tenemos datos para un grupo de individuos sobre su peso y altura, una forma de medir la correlación entre estas dos variables sería hacernos preguntas del tipo: ¿Es el individuo más alto también el más pesado? ¿Cuál es el ranking en cuanto a altura del individuo más pesado?

Supongamos que ordenamos a 8 personas por altura, donde el individuo A es el más alto (valor=1) y el tercero más pesado (valor=3):

Persona	A	B	C	D	E	F	G	H
Ranking por altura	1	2	3	4	5	6	7	8
Ranking por peso	3	4	1	2	5	7	8	6

Ahora cogemos como referencia el número 3 de cada uno de los dos rankings y nos preguntamos: ¿Qué valores están a la derecha del 3 en *ambos* rankings? Por ejemplo, el valor 4: esto es lo que llamamos un par concordante. ¿Qué valores están uno a la derecha y otro a la izquierda del 3? Por ejemplo, el valor 1: este es un par discordante.

Sin embargo, hay un tercer caso, no expresado en la anterior tabla: ¿qué ocurre si dos individuos tienen la misma altura, el mismo peso o ambas medidas iguales? En este caso, dichas observaciones deberían ser tratadas de forma especial, pues distorsionan los rankings. Estos casos son conocidos con el nombre de empates.

La Tau b de Kendall para las variables X e Y es:

$$\tau_b = \frac{C - D}{\sqrt{(C + D + E_x)(C + D + E_y)}}$$

donde C es el número de pares concordantes; D es el número de pares discordantes; E_x y E_y son el número de empates en las variables X e Y, respectivamente.

El rango de la Tau b de Kendall y la interpretación de sus valores es análogo al del coeficiente de correlación simple: [-1, 1], donde -1 indica perfecto desacuerdo entre los rankings de las dos variables, 1 perfecto acuerdo y 0 ninguna relación.

Tabla 10: Correlaciones Tau b de Kendall entre internalización y los potenciales determinantes

Variable	Corr.	p-valor
1ª componente ACP	0.26	0.00
2ª componente ACP	0.06	0.00
3ª componente ACP	-0.03	0.00
fondos ajenos sobre pasivo	-0.03	0.00
fondos c/p e.c. sobre pasivo	0.08	0.00
fondos l/p e.c. sobre pasivo	-0.05	0.00
fondos e.c. sobre ajenos	0.01	0.42
coste medio deuda l/p	0.05	0.00

3.3 Análisis

La Tabla 6 que analiza los determinantes de la profesionalización evidencia que sólo dos variables son significativas, *1ª componente ACP* y *coste medio deuda l/p*.

- La correlación parcial negativa con *1ª componente ACP* nos lleva a concluir que *a mayor tamaño y mayor valor añadido, mayor profesionalización*.
- La correlación parcial positiva con *coste medio deuda l/p* hace pensar en una relación negativa entre profesionalización y dicha variable. Sin embargo, dicha relación puede ser espuria, generada por la coincidencia a lo largo del periodo muestral de un progresivo aumento en la profesionalización y del descenso en los tipos de interés reales. Este aspecto se discutirá en la sección 3.4.

Para analizar los determinantes de la internalización hay que hacer referencia a las Tablas 7-10. De este análisis emergen solo dos determinantes claros de la internalización, la *1ª componente ACP* y el coste medio de la deuda a largo plazo con entidades de crédito.

- La *1ª componente ACP* tiene un coeficiente negativo significativo en la Tabla 7, un coeficiente positivo significativo en la Tabla 8 y un coeficiente positivo significativo en la Tabla 9, pero con mayor valor absoluto que el anterior. Estos coeficientes sugieren que, al aumentar el tamaño y el valor añadido, las empresas tienden a abandonar la internalización baja a favor de la intermedia o de la alta. El hecho que el coeficiente de la Tabla 9 sea mayor en valor absoluto que el de la Tabla 8, además sugiere que al aumentar el tamaño y el valor añadido las empresas también tienden a abandonar la internalización intermedia a favor de la alta. Estos patrones son confirmados por el coeficiente positivo significativo de la Tabla 10.
- El *coste medio deuda l/p* tiene un coeficiente negativo significativo en la Tabla 7, un coeficiente negativo significativo en la Tabla 8, pero menor en valor absoluto que el anterior, y un coeficiente positivo significativo en la Tabla 9. Estos coeficientes sugieren que al aumentar el coste medio de la deuda a largo plazo con entidades de crédito, las empresas tienden a abandonar la internalización baja o intermedia a favor de la alta. El hecho que el coeficiente de la Tabla 8 sea menor en valor absoluto que el de la Tabla 7 se puede explicar por el hecho que las transiciones de internalización intermedia a alta determinadas por el aumento en coste medio de la deuda a largo plazo con entidades de crédito, son en parte compensadas por algunas transiciones de internalización baja a intermedia. Todo lo anterior sugiere una relación positiva entre coste medio de la deuda a largo plazo con entidades de crédito e internalización. Pero, como en el caso de la

profesionalización, dicha relación puede ser espuria, consecuencia del descenso simultáneo de la internalización y de los tipos de interés reales. Este aspecto se discutirá en la sección 3.4.

- Un examen de las Tablas 7-10 no evidencia ningún patrón claro entre internalización y ninguno de los otros determinantes potenciales que se han estudiado, ya que para cada uno de estos determinantes aparecen coeficientes no significativos.

3.4 Discusión

El análisis llevado a cabo hasta ahora se ha realizado para el conjunto de todas las observaciones de las que disponíamos. Esto tiene una ventaja importante porque el mayor número de observaciones y la mayor variabilidad a lo largo del intervalo muestral (17 años) permite una mejor estimación de las relaciones entre las variables que se estudian. Pero ignorar que las observaciones proceden de años distintos tiene la desventaja de hacer posible que se encuentren relaciones espurias entre las variables. Esto puede darse cuando alguna variable omitida del análisis tiene un impacto sobre dos o más variables incluidas en el análisis.

Por esta razón hemos repetido el análisis para cada uno de los años de la muestra con el objetivo de comprobar si los resultados cualitativos son robustos. El resultado de este ejercicio es que la relación entre profesionalización e internalización se mantiene pero la relación entre profesionalización e internalización y el coste medio real de la deuda a largo plazo con entidades desaparece.

Para entender por qué es útil hacer referencia a las Figuras 2, 3 y 4. Como se recordará, la Figura 2, presentada anteriormente, representa la evolución de profesionalización e internalización de las Pymes a lo largo de los años muestrales y evidencia que se ha producido un aumento en la profesionalización y un descenso en la internalización. Las Figuras 3 y 4 a continuación representan la distribución de los costes medios de la deuda a largo plazo con entidades de crédito en el primer y el último año de la muestra, 1990 y 2006. En las Figuras 3 y 4 se evidencian también el tipo de inflación nominal del año, medido por la tasa de crecimiento del IPC, y la media de los costes medios de la deuda a largo plazo con entidades de crédito declarados por las empresas encuestadas.

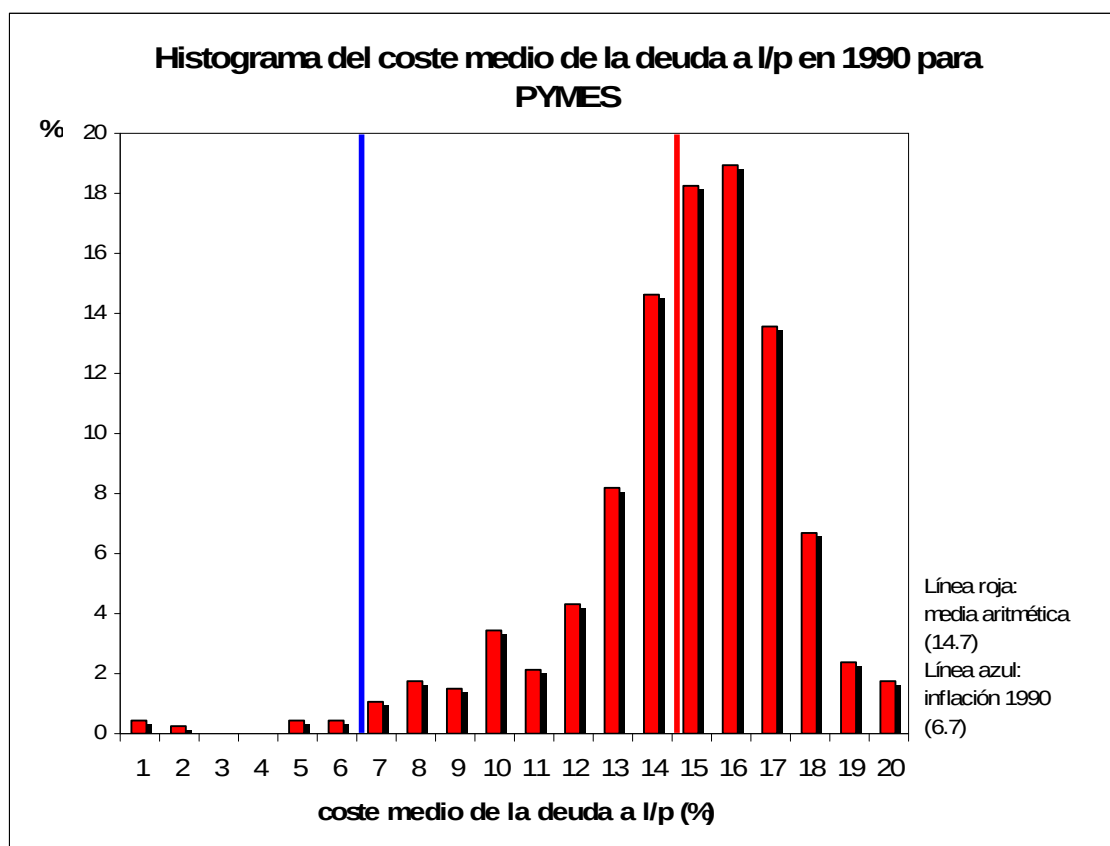


Figura 3

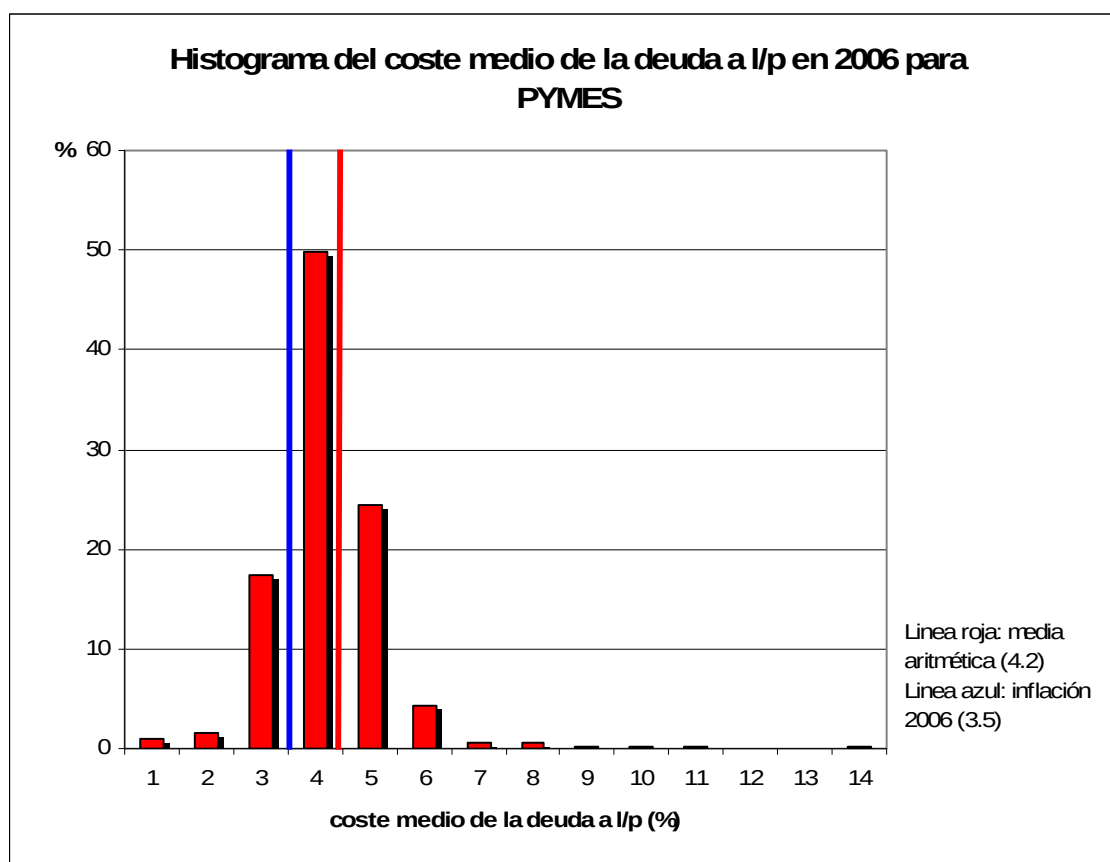


Figura 4

Las diferencias más llamativas evidenciadas por las Figuras 3 y 4 son las siguientes:

- La dispersión en 1990 es mucho mayor que en 2006;
- La diferencia entre los costes medios (nominales) y la inflación en 1990 es mucho mayor que en 2006. En promedio la diferencia, y por lo tanto la media de coste medio real de la deuda en 1990 es 7,98%, mientras que en 2006 es tan sólo 0,67%.

La comparación entre 1990 y 2006 sintetiza los desarrollos que se han producido en los años intermedios en los que se han producido bajadas progresivas y graduales en los costes medios de la deuda a largo plazo con entidades de crédito. En 1994, 1998, y 2002, los costes medios reales fueron en promedio, 6,72%, 3,77% y 1,20%, respectivamente, y se ha observado también una progresiva reducción en la variabilidad de los costes medios de la deuda reportados por las empresas encuestadas.

El examen de las Figuras 2, 3 y 4 permite comprender por qué el coste medio de la deuda a largo plazo con entidades de crédito aparece como determinante significativo de profesionalización e internalización para el conjunto de la muestra pero no para cada año de la muestra. Como habíamos apuntado con anterioridad, las relaciones entre profesionalización e internalización por un lado y el coste medio de la deuda a largo plazo con entidades de crédito por el otro son espurias y dependen del hecho de que a lo largo del periodo muestral se producen progresivos y graduales

- ascensos en la profesionalización;
- descensos en la internalización;
- descensos en los costes medios de la deuda a largo plazo con entidades de crédito.

Tratar de interpretar estos desarrollos temporales va mas allá de los objetivos de este trabajo, ya que ellos dependen de muchas variables que no se analizan en el presente estudio como los tipos de interés de mercado, las condiciones de competitividad en el sector crediticio, la disponibilidad de trabajadores cualificados para realizar internamente o externamente los servicios de asesoría económico-financieras, o los desarrollos tecnológicos que pueden haber favorecido la externalización de servicios.

Sin embargo el cuadro ofrecido por las Figuras 2, 3 y 4 es coherente con un escenario que se basa en las siguientes premisas:

- El retorno de profesionalizar la toma de decisiones económico-financieras es mayor cuando una empresa planea recurrir a más financiación ajena;

- Las Pymes son empresas *informacionalmente opacas* (véanse por ejemplo Berger, Klapper y Udell (2001), Berger, Rosen y Udell (2007), Berger y Udell (2006), Scott y Dunkelberg (2003). La opacidad informacional dificulta el acceso a fondos ajenos, en particular a deuda con entidades de crédito y su reducción lo facilita.
- Internalizar los servicios de asesoría económico-financiera ayuda a reducir la opacidad informacional de una empresa (en las negociaciones con potenciales proveedores de fondos) pero la obliga a soportar el coste de esta reducción.
- Entre 1990 y 2006 se ha producido una bajada en los tipos de interés reales.
- Entre 1990 y 2006 se ha producido un incremento en la competencia entre entidades de crédito; esto se ha reflejado tanto en condiciones de préstamos más ventajosa como en una creciente disponibilidad a soportar los costes de reducción de la opacidad informacional.

Bajo estas premisas el escenario que emerge es uno en el que la reducción en los tipos reales y el incremento en la competencia entre las entidades de crédito han aumentado el retorno de la profesionalización y por lo tanto han producido su aumento. Pero a la vez el incremento en la competencia entre las entidades de crédito ha producido un equilibrio en el que el coste de reducción de la opacidad informacional de las Pymes ha sido trasladado de las Pymes a las entidades de crédito, produciendo por lo tanto una reducción en la internalización.

4. Comparación entre pequeñas y medianas empresas

En esta sección se analizan las empresas pequeñas y medianas por separado, para ver si comparten los mismos determinantes de la profesionalización e internalización. El análisis que se realiza es idéntico al explicado en la sección anterior, por lo que nos centraremos en resultados y conclusiones. Nos basaremos en Tablas análogas a las Tablas 6-10 pero no las reportaremos por brevedad¹.

- *Profesionalización.*
 - Las empresas pequeñas tienen los mismos determinantes de la profesionalización que las Pymes en su conjunto. Es decir, *a mayor tamaño y mayor valor añadido, mayor profesionalización*. Para el conjunto de la muestra emerge también una relación negativa con el coste medio real de la deuda a largo plazo con entidades de crédito. Pero como para el conjunto de las Pymes, el análisis por año demuestra que se trata de una relación espuria.

¹Dichas Tablas están a disposición de los interesados.

- En el caso de las empresas medianas no se encuentra ningún determinante significativo de la profesionalización.
- *Internalización.* En cuanto a los determinantes de la internalización, no existen diferencias claras entre empresas pequeñas y grandes. Al igual que para el conjunto de las Pymes, existe una relación positiva entre tamaño, valor añadido e internalización para ambos tipos de empresas. Al aumentar el tamaño y el valor añadido, las empresas tienden a abandonar la internalización baja a favor de la intermedia o de la alta. Además, al aumentar el tamaño y el valor añadido las empresas también tienden a abandonar la internalización intermedia a favor de la alta. Como en el caso de la profesionalización, para el conjunto de la muestra emerge una relación positiva con el coste medio de la deuda a largo plazo con entidades de crédito. Pero, como en los casos anteriores, la relación es espuria.

Las Figuras 5 y 6 permiten comprender que como en los casos anteriores la relación del coste medio de la deuda a l/p con profesionalización e internalización es espuria, ya que para los subconjuntos de empresas pequeñas y medianas se observan los mismos patrones temporales que para el conjunto de las Pymes.

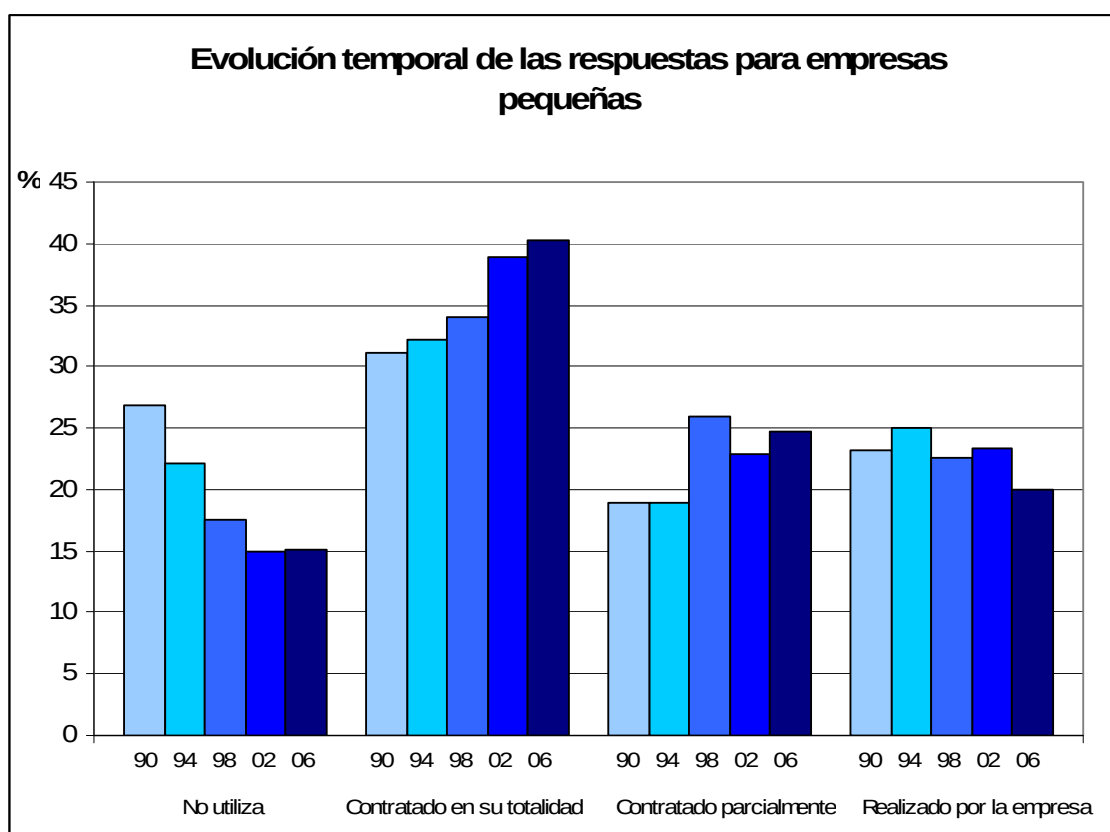


Figura 5

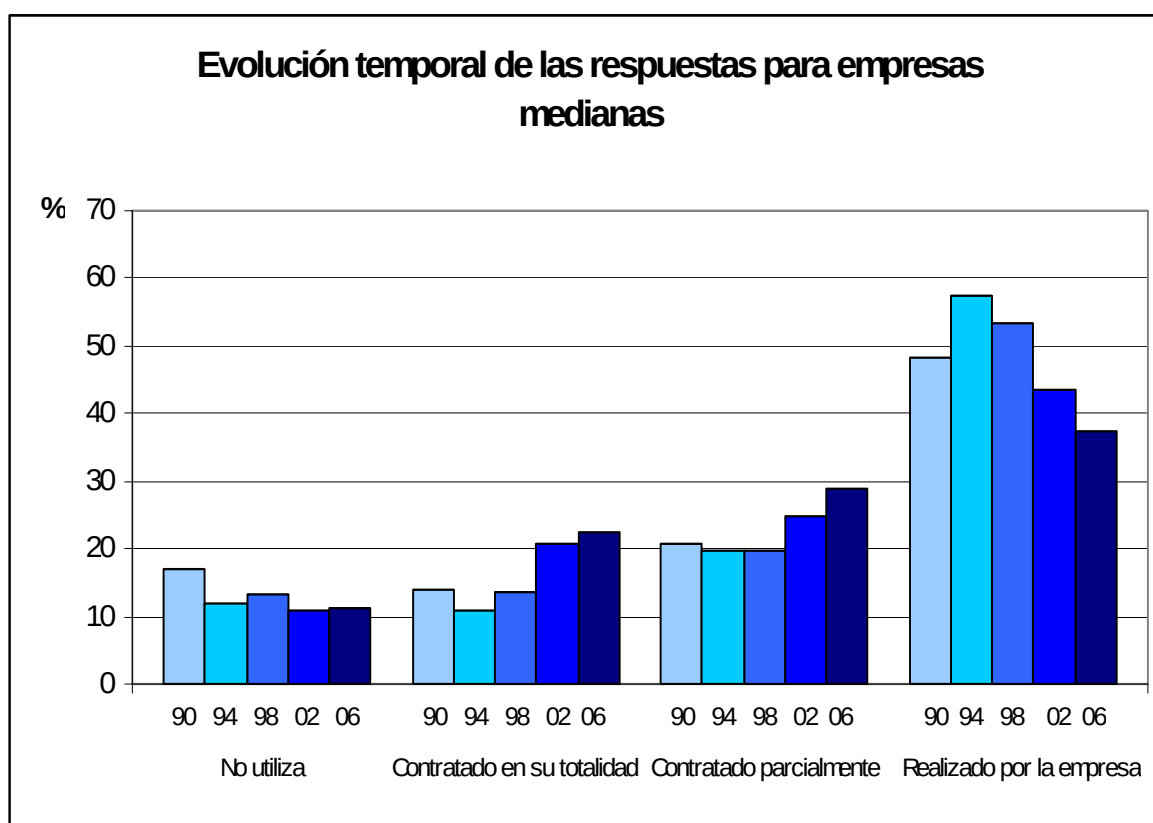


Figura 6

Es interesante observar cómo el nivel de profesionalización de las empresas medianas apenas varía a lo largo del periodo muestral, para comprender por qué el coste medio de la deuda a l/p no aparece como determinante de la profesionalización.

5. Comparación entre Pymes y empresas grandes.

En esta sección se contrastan los resultados obtenidos para el total de las Pymes con aquellos obtenidos para las empresas grandes. Con el propósito de hacer la exposición más concisa, no reportamos las Tablas de correlaciones parciales y de la Tau-b de Kendall.

- *Profesionalización.* Existe una diferencia importante en cuanto a los determinantes de la profesionalización entre Pymes y empresas grandes, el coste medio real de la deuda a largo plazo con entidades de crédito. En el caso de las grandes, no existe una relación clara entre coste medio y profesionalización, aunque sí que existe una relación positiva entre tamaño, valor añadido y profesionalización.
- *Internalización.* No existe una relación clara entre coste medio e internalización, aunque sí que existe una relación positiva entre tamaño, valor añadido e internalización.

La evolución temporal de profesionalización e internalización para empresas grandes se representa en la Figura 7.

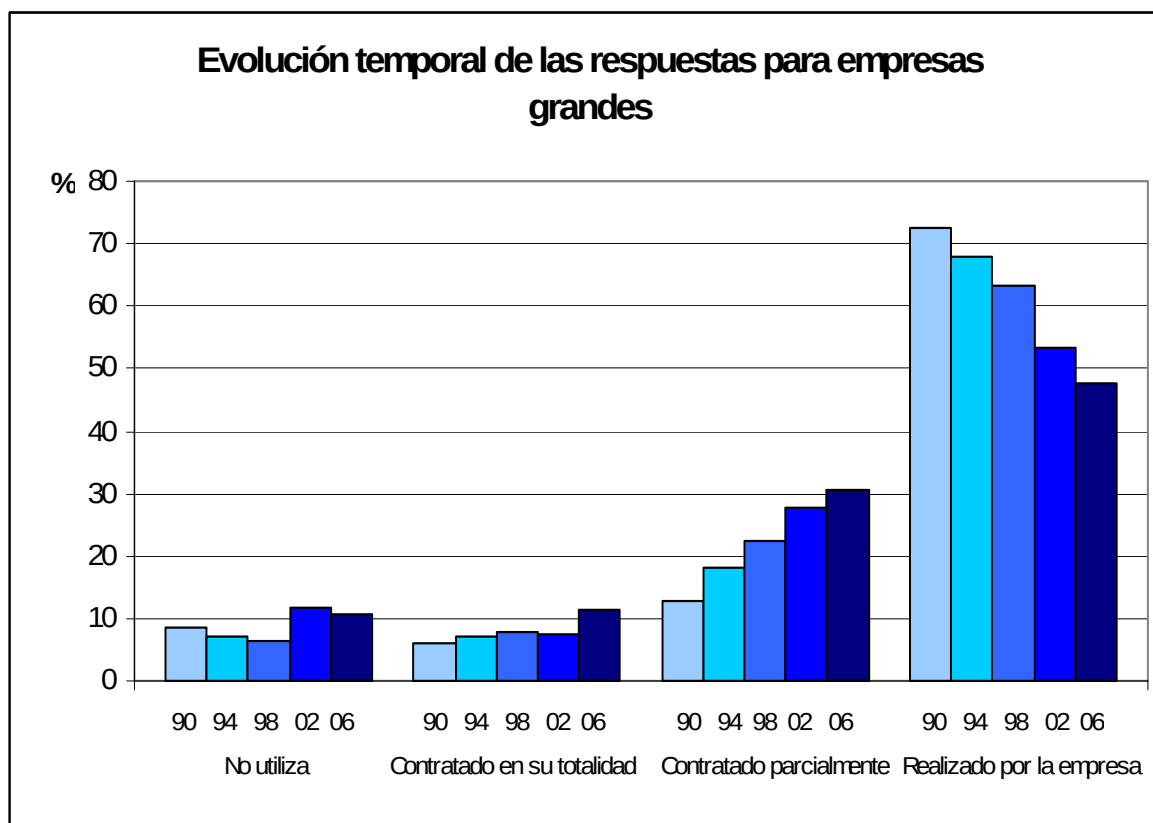


Figura 7

Es interesante destacar que a diferencia de las Pymes y en particular a diferencia de las pequeñas, el nivel de profesionalización para las empresas grandes apenas varía a lo largo del periodo muestral (si se exceptúa una pequeña disminución en 2002 y 2006). Pero como para las Pymes, el nivel de internalización disminuye a lo largo de los años. También es interesante destacar que para las empresas grandes como para las Pymes se observa un descenso progresivo y gradual del coste medio real de la deuda a l/p que en los 5 años muestrales tiene los siguientes valores promedio: 6,36%, 4,84%, 2,84%, 0,65% y 0,51%.

6. La evolución temporal de profesionalización e internalización

El análisis anterior ha destacado los principales determinantes de profesionalización e internalización. El objetivo de esta sección es cuantificar el impacto de estos determinantes. Para ello a continuación se presentan unas Figuras que resumen la manera en la que las decisiones de profesionalización e internalización varían a lo largo de la distribución de las variables económicas

que presentan una asociación más clara con profesionalización e internalización, ventas y pasivo. La razón por la que excluimos del análisis el número de trabajadores es por su elevada correlación con las ventas, mientras que el valor añadido se ignora por tener un impacto cuantitativo menor.

Para cada una de estas variables se analiza la distribución de las respuestas a la pregunta sobre profesionalización e internalización para el conjunto de todos los años de la muestra, así como para el primero y el último año de la muestra para evidenciar los cambios a lo largo del periodo muestral. Las distribuciones se distinguen para empresas pequeñas, medianas y grandes.

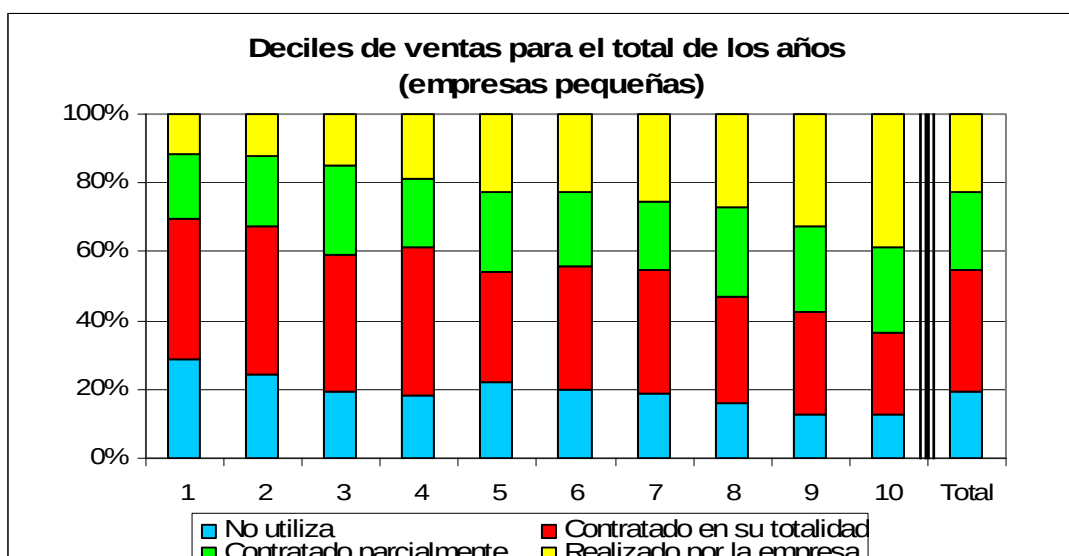


Figura 8

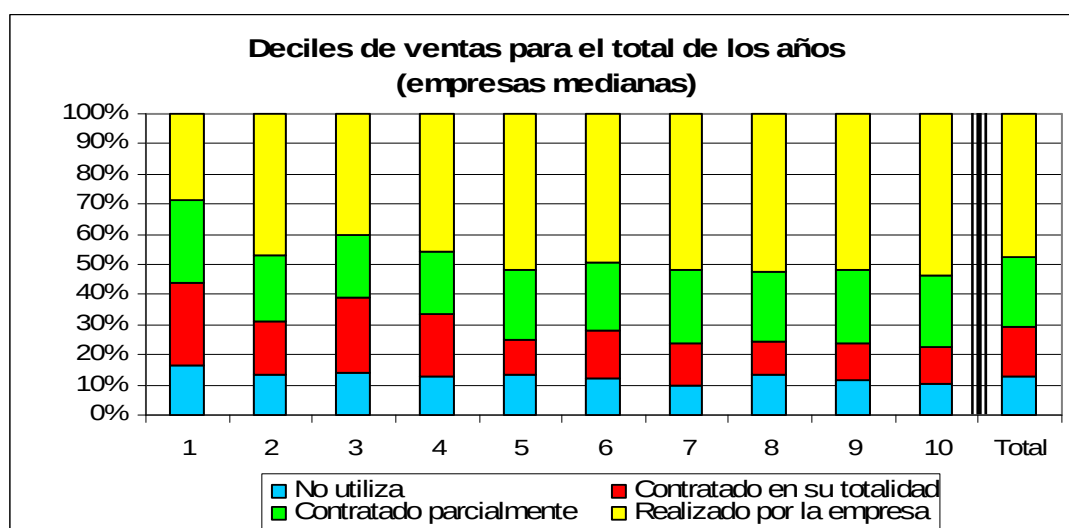


Figura 9

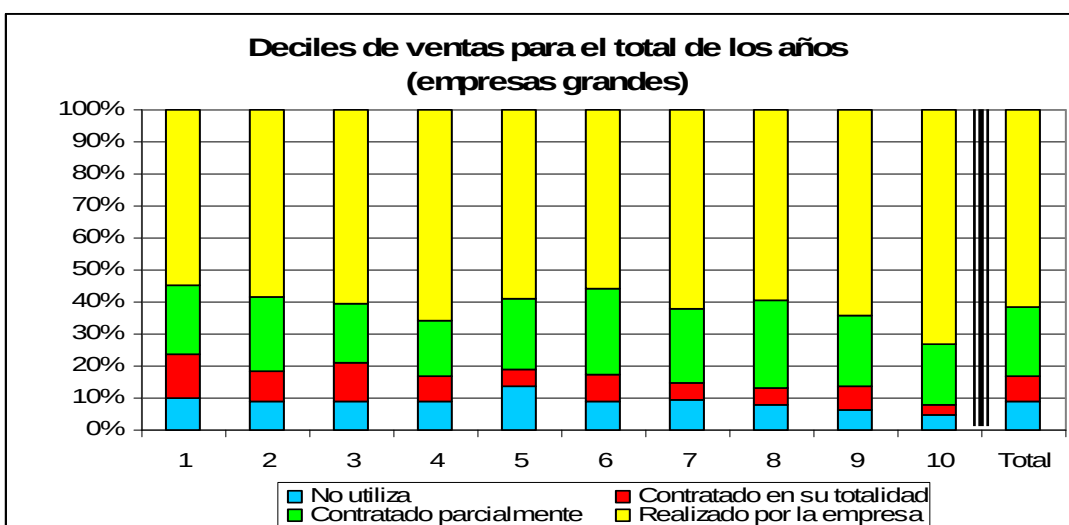


Figura 10

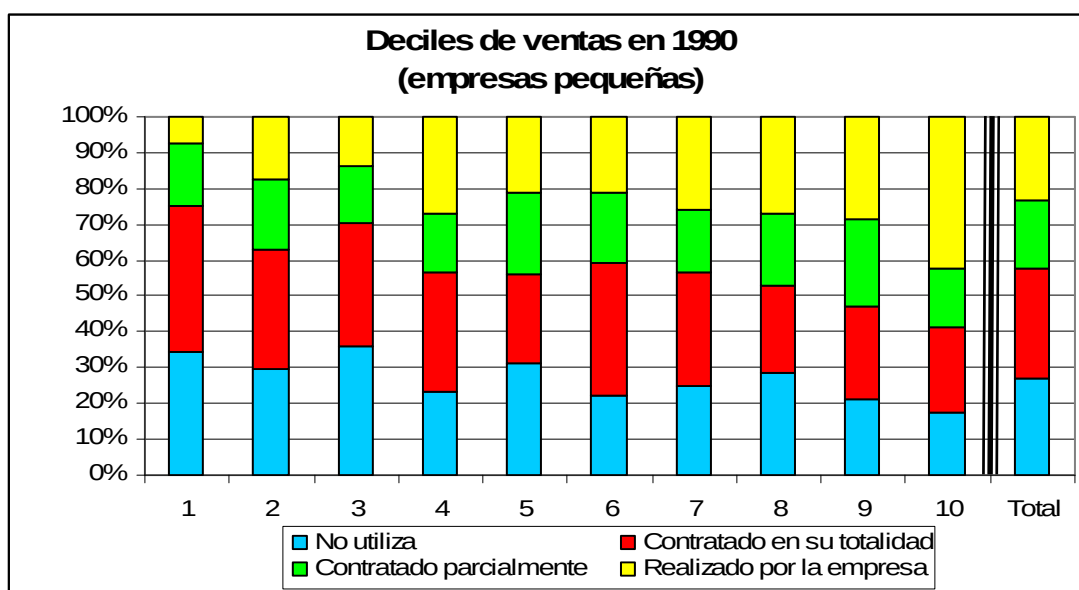


Figura 11

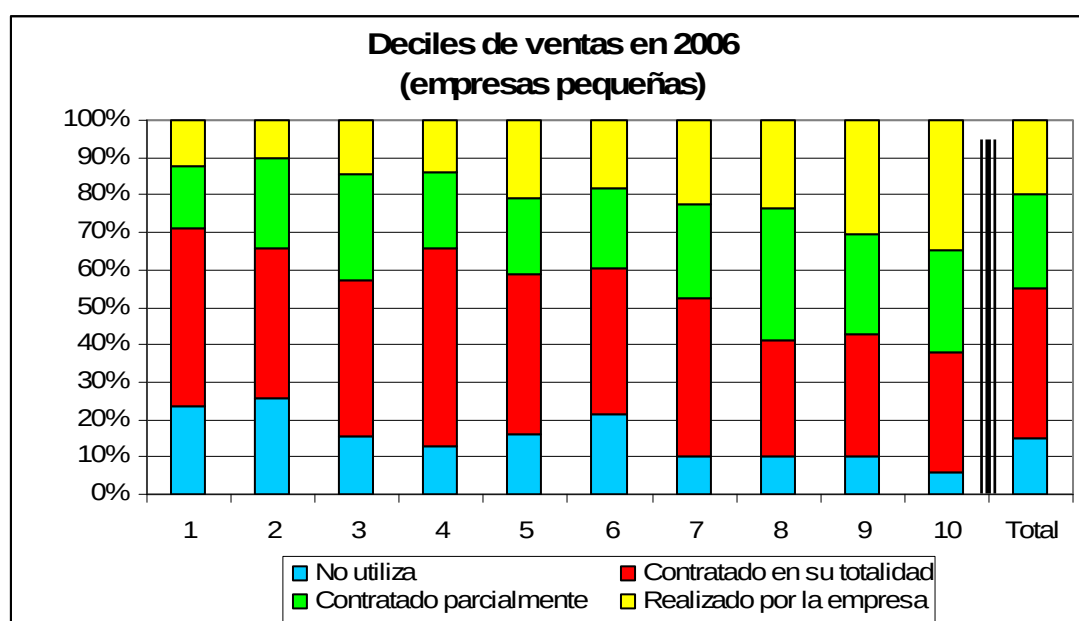


Figura 12

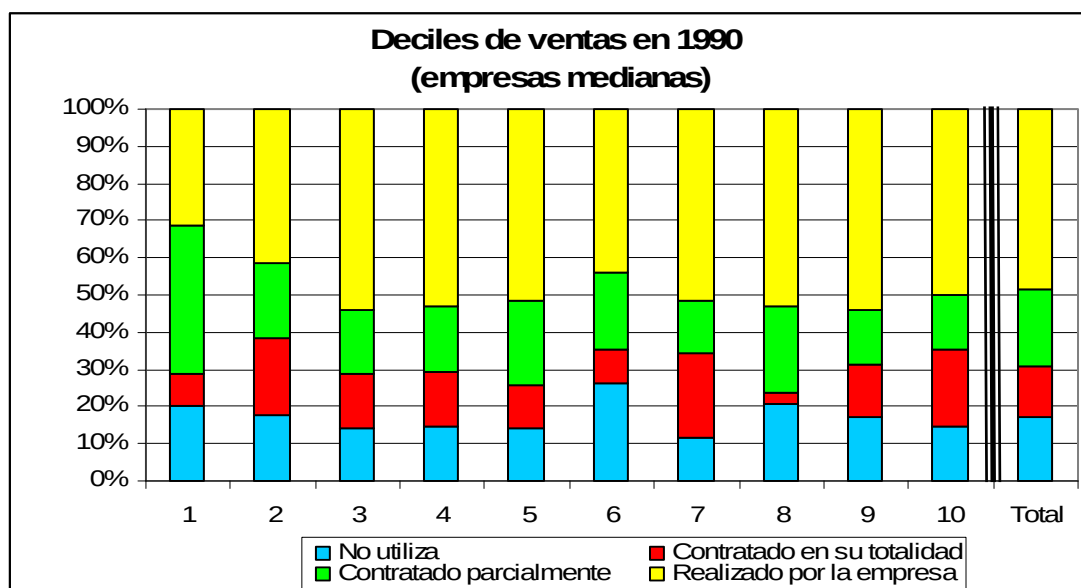


Figura 13

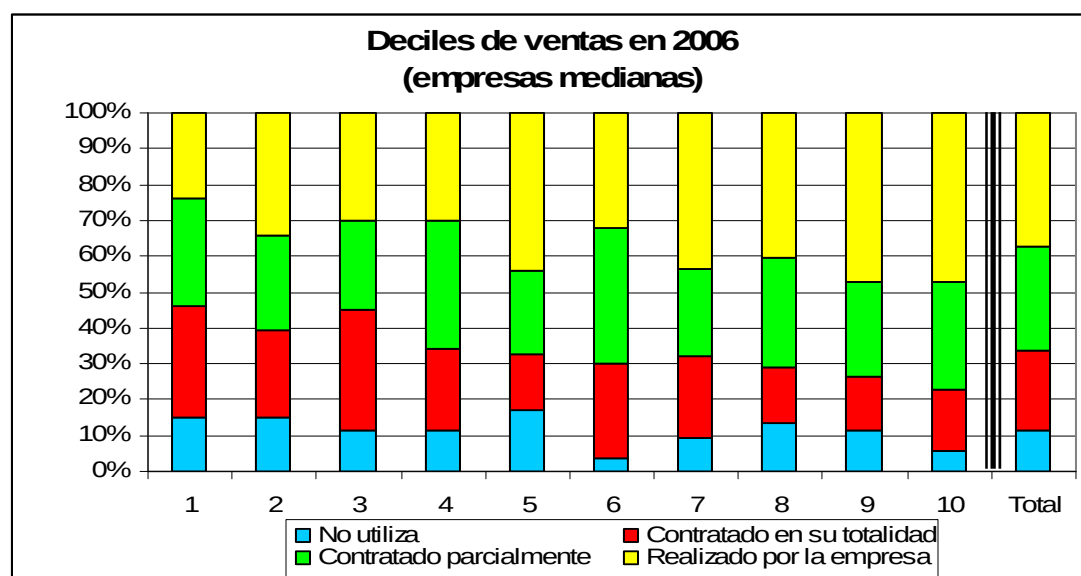


Figura 14

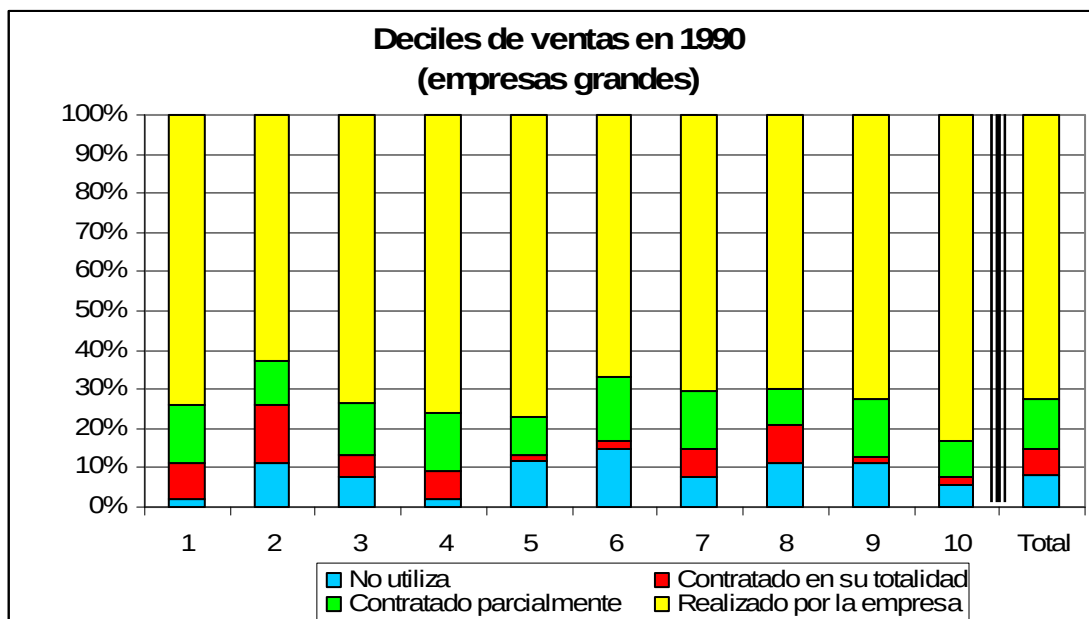


Figura 15

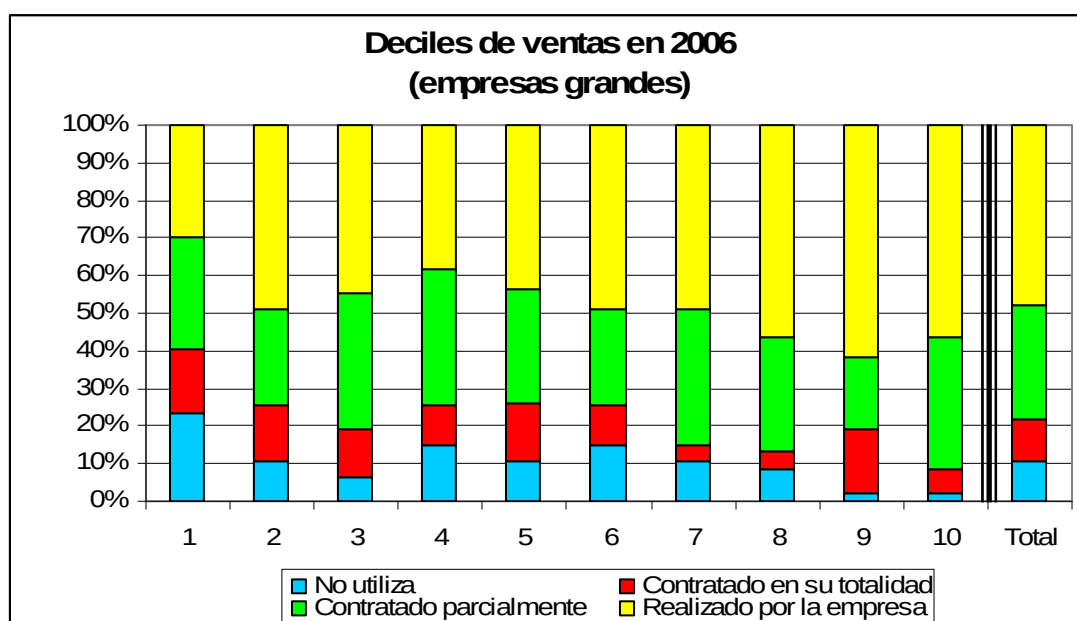


Figura 16

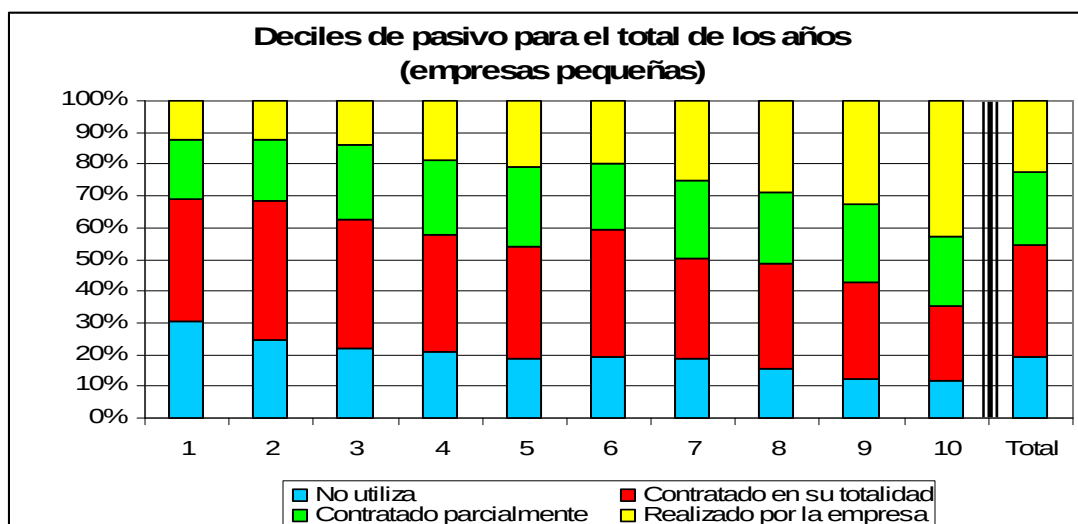


Figura 17

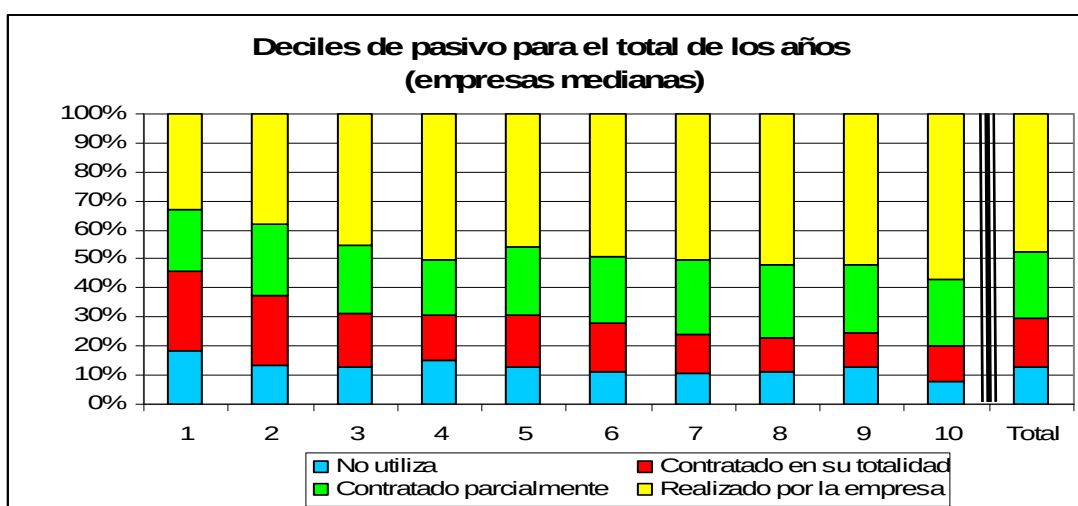


Figura 18

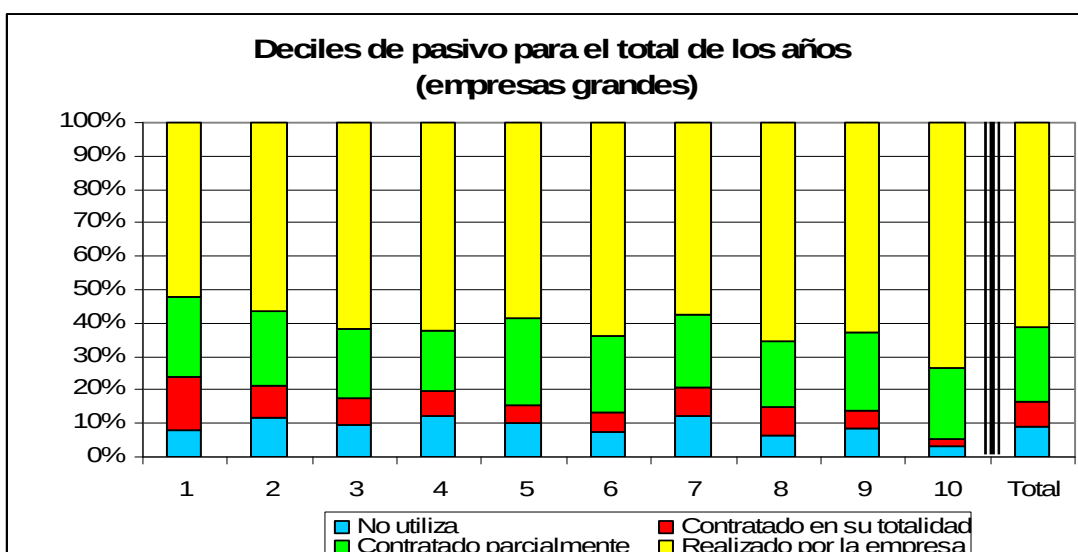


Figura 19

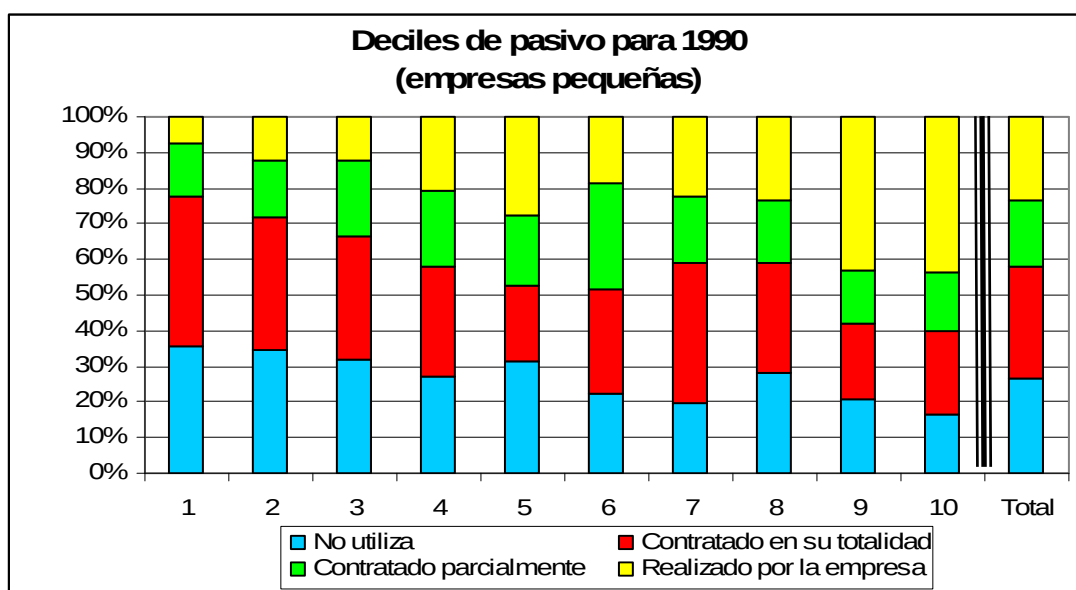


Figura 20

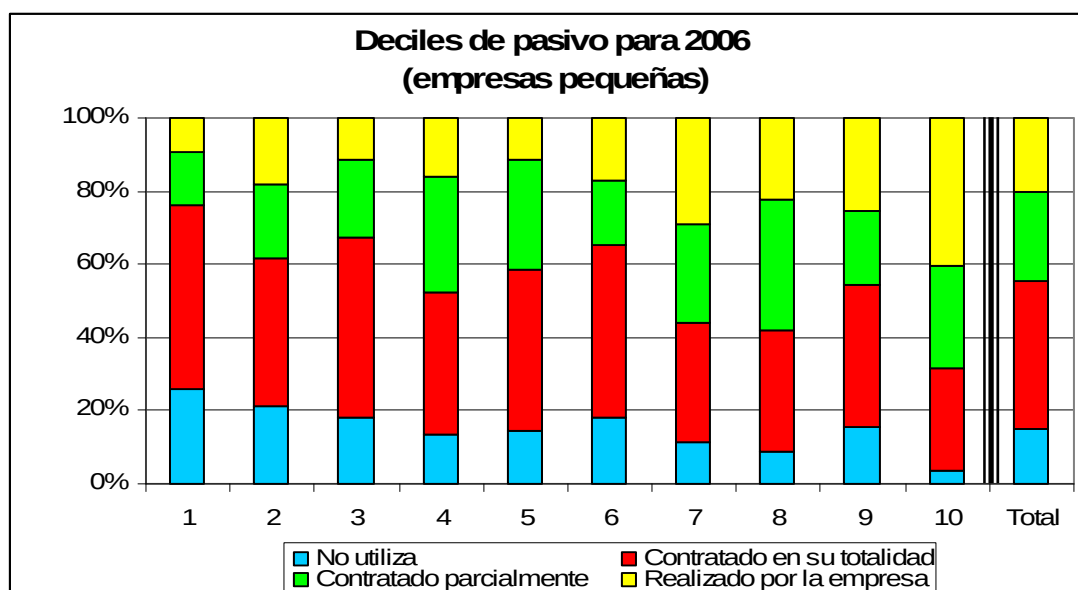


Figura 21

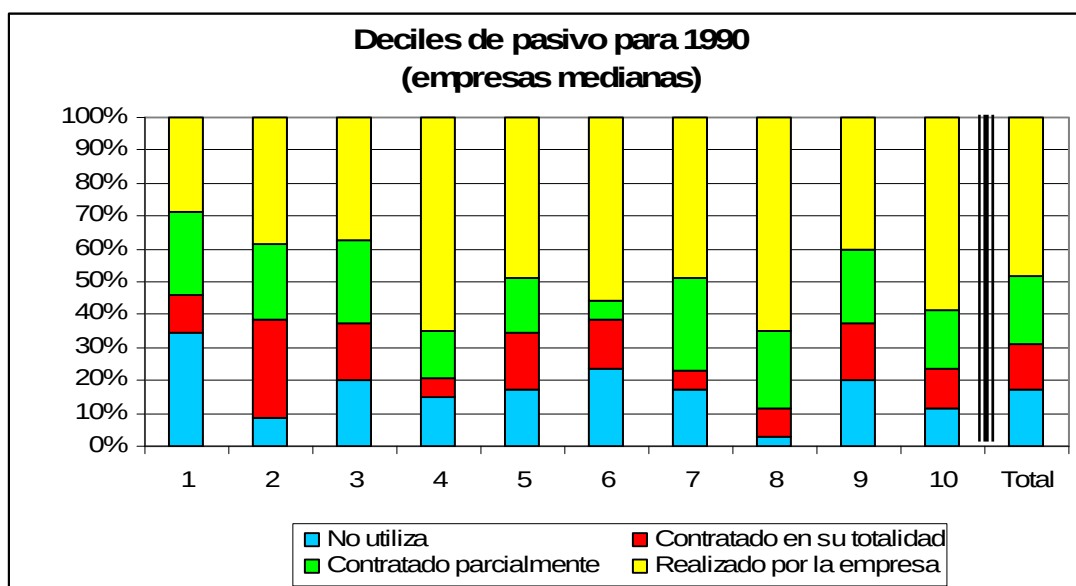


Figura 22

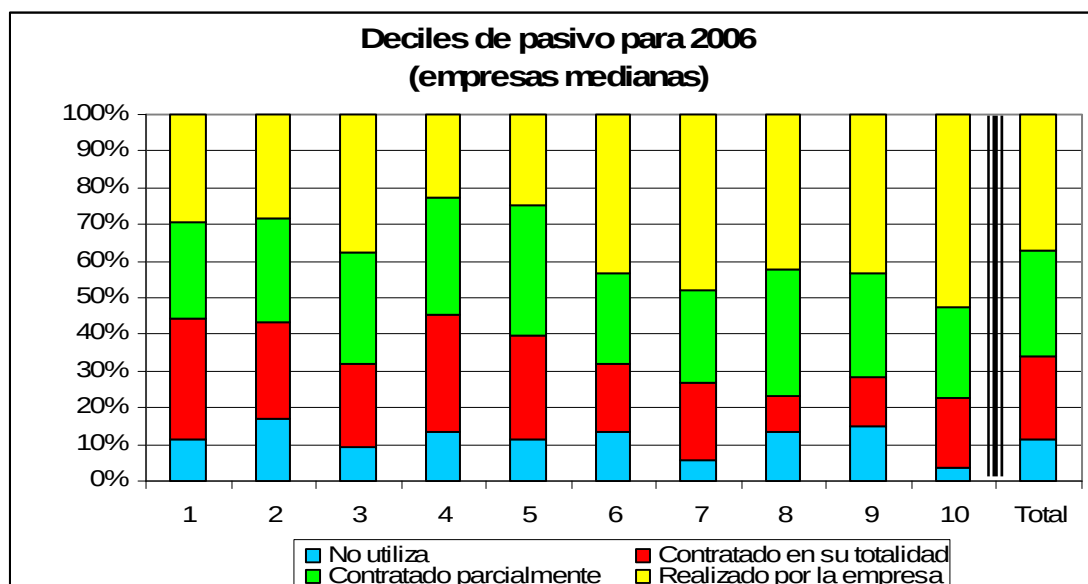


Figura 23

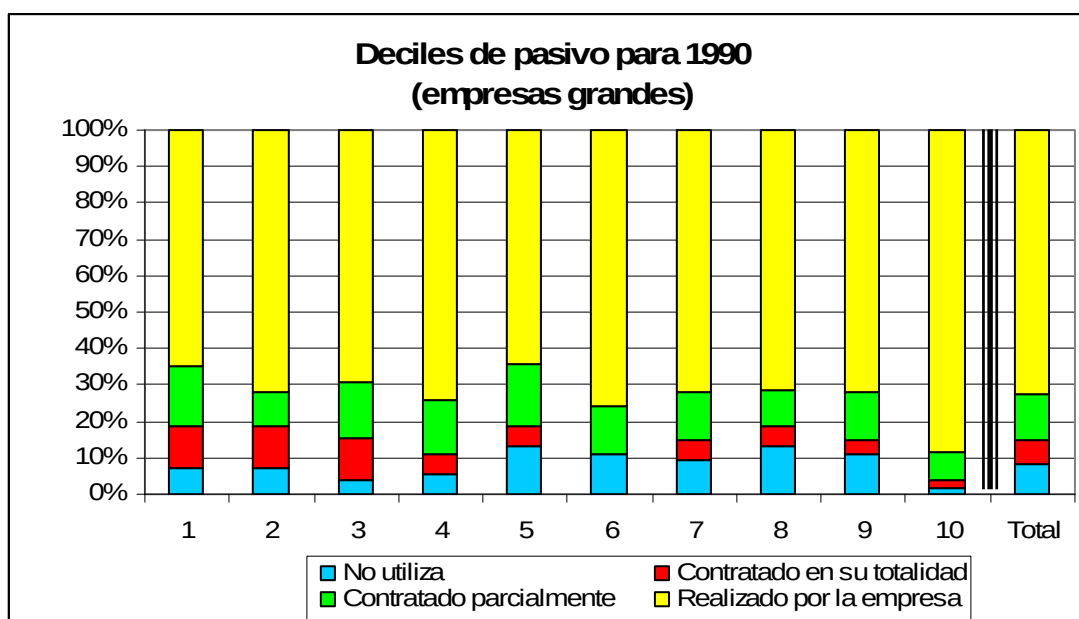


Figura 24

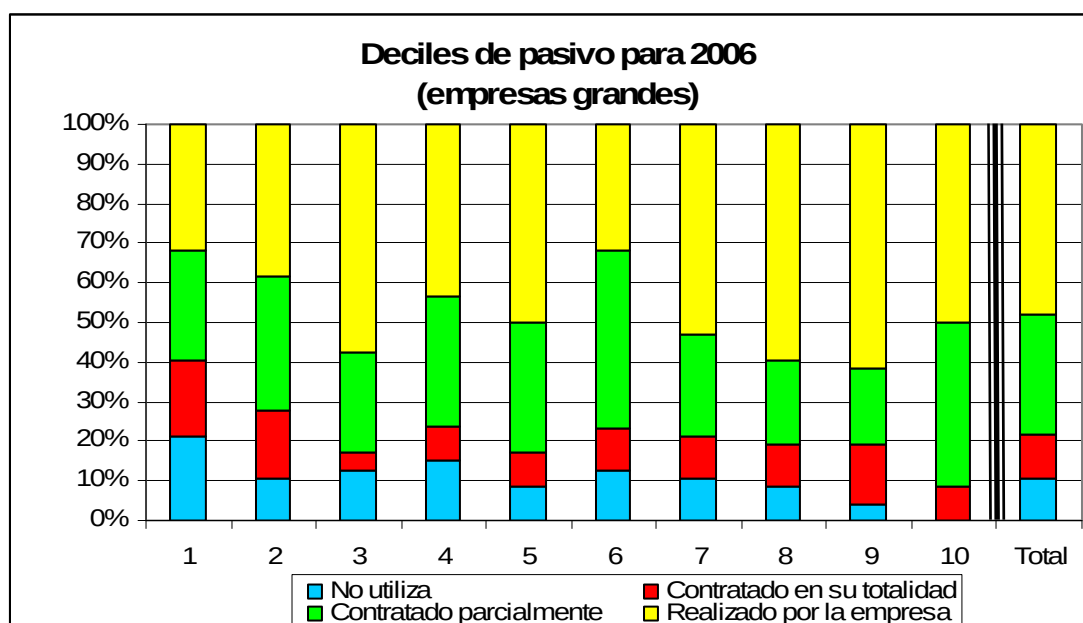


Figura 25

7. Conclusiones

El estudio de profesionalización e internalización de la tomas de decisiones económico-financieras para las empresas manufactureras española en el periodo 1990-2006 evidencia los siguientes patrones principales:

- *Empresas pequeñas.* Tienen los menores niveles de profesionalización. Tanto la profesionalización como la internalización aparecen positivamente relacionadas con el tamaño, medido en términos de número de trabajadores, ventas, pasivo y valor añadido.
- *Empresas medianas.* Tienen niveles de profesionalización intermedios. La profesionalización no tiene ningún determinante significativo ya que su nivel es relativamente estable a lo largo de la distribución de las variables que miden el tamaño (compárense por ejemplo los niveles de profesionalización de empresas pequeñas y medianas en las Figuras 8 y 9 o en las 17 y 18 donde se representan en función de ventas o pasivo). La internalización tiene una asociación positiva significativa con el tamaño.
- *Empresas grandes.* Tienen niveles de profesionalización e internalización elevados. Tanto la profesionalización como la internalización están asociados significativamente con el tamaño.
- A lo largo del periodo estudiado se observan las siguientes *tendencias*:
 - *Empresas pequeñas:* Un incremento en la profesionalización y una notable tendencia hacia la externalización parcial y, sobre todo, total.
 - *Empresas medianas:* Un cierto aumento en la profesionalización y una ligera tendencia hacia la externalización total o parcial.
 - *Empresas grandes:* Ninguna variación apreciable en la profesionalización y una destacable tendencia hacia la externalización parcial.

Estos desarrollos son coherentes con la disminución en los tipos de interés reales y un aumento en la competencia entre las entidades de crédito, que a la vez que aumentan el retorno de la profesionalización reducen el retorno de la internalización, trasladando el coste de la reducción de la opacidad informacionales de las Pymes a las entidades de crédito.

Apéndice: Análisis de los determinantes de profesionalización e internalización con una clasificación alternativa de las empresas

Como destacamos en la sección 1.2, los resultados que hemos obtenido no dependen de la clasificación del tamaño de empresas empleada. Esto lo hemos probado realizando un análisis de robustez en el que usamos una clasificación alternativa basada en las mismas variables que las de la Comisión Europea, pero empleando métodos puramente estadísticos (Análisis de Componentes Principales y Análisis de Conglomerados) para definir las categorías. Si bien la proporción de empresas en cada categoría cambia ligeramente, las principales conclusiones que obtenemos con ambas clasificaciones son las mismas.

Primero, mediante el Análisis de Componentes Principales (ACP) conseguimos sintetizar nuestras 3 variables de interés (número de empleados, volumen de ventas y pasivo total) en una sola, que es la primera componente principal². Esto se puede conseguir porque dicha componente explica un 84% de la varianza total de las variables originales, las comunalidades son muy superiores al 50% y los autovalores del resto de las componentes son inferiores a 0,8. Además, el estadístico KMO nos da un valor de 0,72, por lo que estas variables se prestan adecuadamente a un ACP.

Con esta componente queremos generar un índice de tamaño que vaya de 0 a 100, donde mayores valores expresen mayor tamaño. Sin embargo, primero observamos la distribución estadística de nuestra componente (concretamente de la puntuación de dicha componente) mediante un kernel. Dicho kernel nos muestra una distribución muy asimétrica a la derecha, es decir, la mayor parte de sus valores se encuentran a la izquierda y presenta una larga cola de pocos valores muy grandes a la derecha (Gráfico 1). Esto dificulta la realización del Análisis de Conglomerados, puesto que uno de sus supuestos básicos es la ausencia de datos atípicos. Por ello realizaremos primero una transformación logarítmica (en concreto, el logaritmo neperiano) para corregir por dicha asimetría, tras haber reescalado la variable para que sus valores sean mayores que 0 (dado que el logaritmo neperiano no está definido para valores negativos). El resultado lo podemos apreciar en el Gráfico 2. Una vez hecho esto, usamos un diagrama de “caja y bigotes” para ver si quedan datos atípicos. El Gráfico 3 nos muestra un dato atípico, el cual eliminamos. Por último, volvemos a reescalar la variable, esta vez para que el rango de sus valores vaya de 0 a 100. Con esto tenemos ya hecho nuestro Índice de tamaño, donde 0 corresponde a la observación con menor tamaño y 100 a la observación con mayor tamaño.

² Véase el recuadro metodológico 2 para una descripción del ACP.

Gráfico 1: Kernel de la primera componente principal

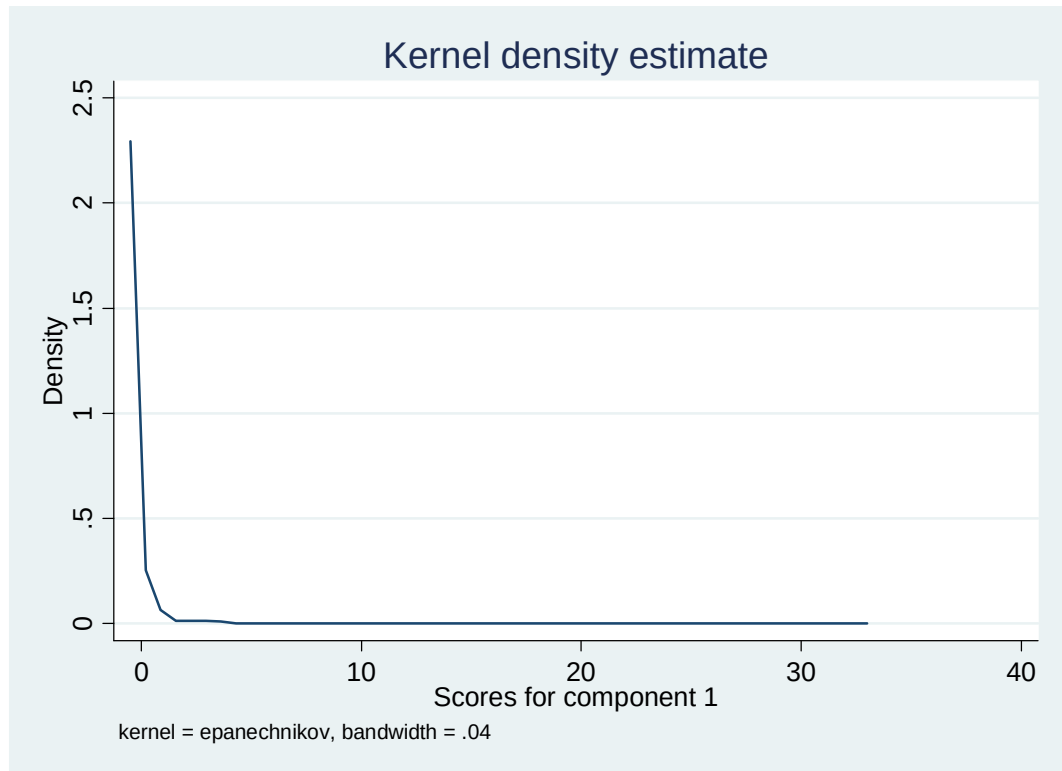


Gráfico 2: Kernel de la transformación logarítmica de la primera componente principal.

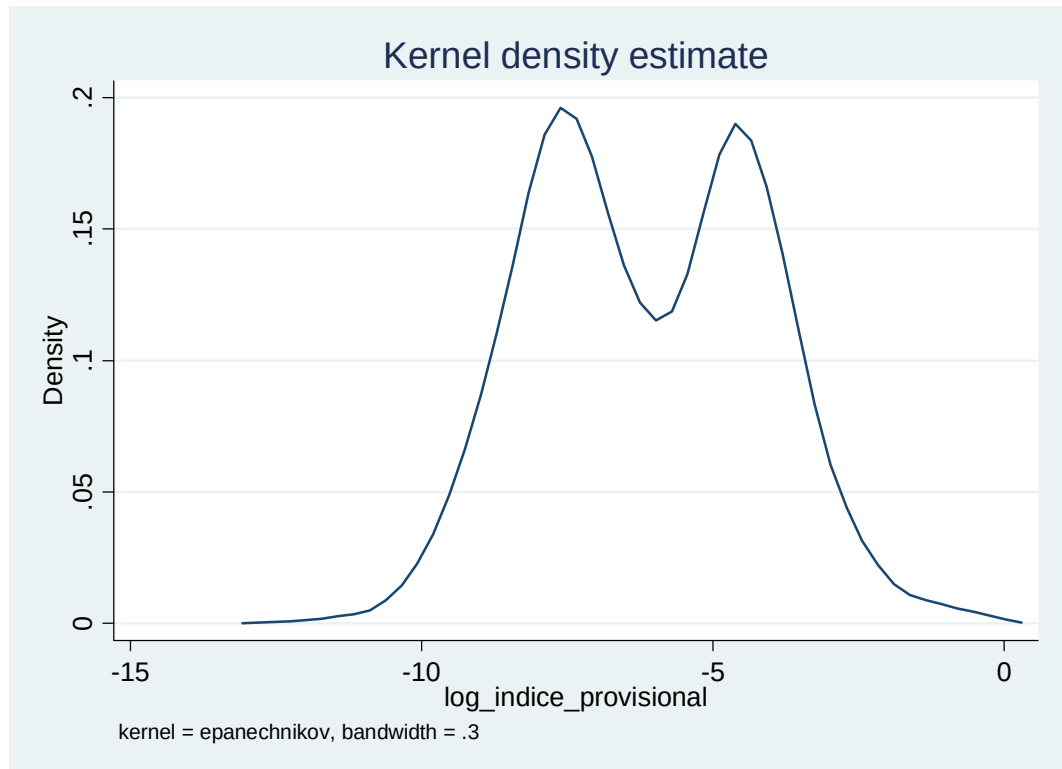
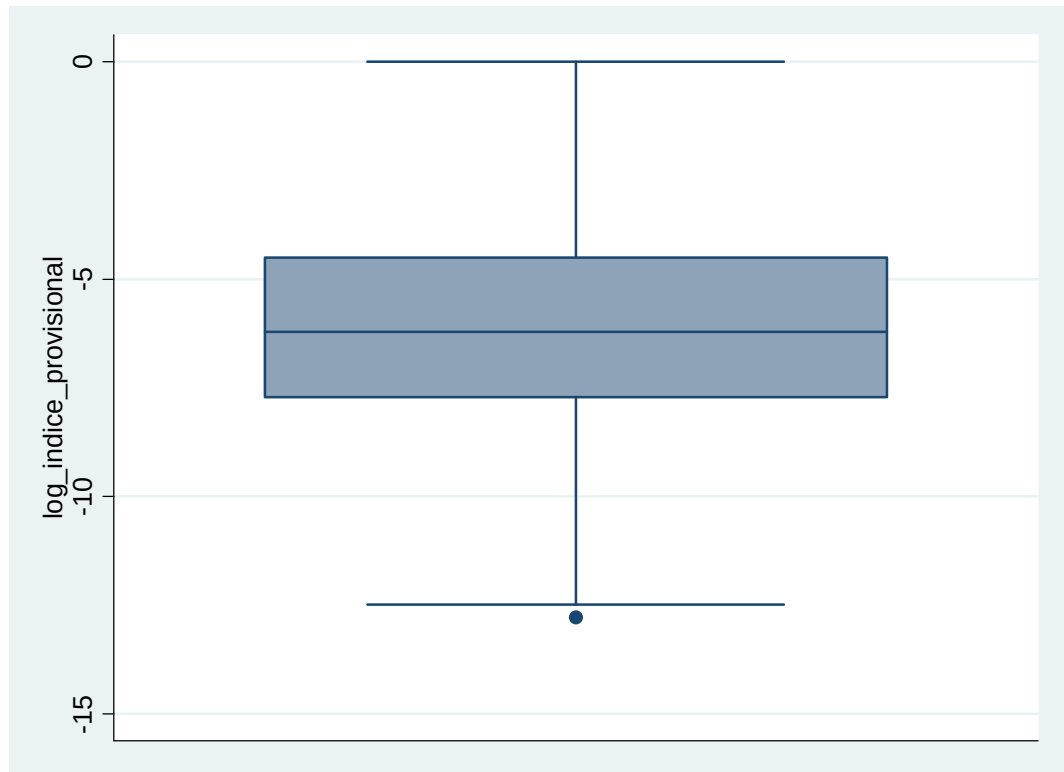


Gráfico 3: Diagrama de caja y bigotes de la transformación logarítmica de la primera componente principal.



Recuadro metodológico 6: Kernels.

- Kernels: los kernels son estimaciones no paramétricas de la distribución de una variable. Este procedimiento consiste en estimar dicha distribución de forma parecida a la que resultaría con un histograma, pero de una forma más continua, sin la limitación impuesta por el agrupamiento en columnas. Adicionalmente, la existencia de un parámetro de ancho de banda permite elegir el grado de suavización presente en la curva, aunque existen varios métodos para calcular su valor óptimo. Matemáticamente, la fórmula del kernel con un ancho de banda “h” de la variable X toma la forma:

$$f_h(x) = \frac{1}{Nh} \cdot \sum_{i=1}^N K\left[\frac{x - x_i}{h}\right]$$

donde $K(\cdot)$ corresponde a una función de densidad y “h” es el ancho de banda.

La función de densidad que se toma más habitualmente es la Epanechnikov:

$$K(x) = \frac{3}{4} \cdot (1 - x^2) \cdot I(|x| \leq 1)$$

donde $I(\cdot)$ es la función indicador, la cual toma el valor 1 si el valor absoluto de x es menor o igual a 1.

Recuadro metodológico 7: “Diagrama de caja y bigotes”

- “Diagrama de caja y bigotes”: este gráfico sirve para identificar las principales características de la distribución de una variable. Los bordes horizontales de la caja muestran el primer y tercer cuartil, mientras que la línea horizontal que la cruza muestra la mediana. Las líneas en forma de T invertida y T muestran el límite inferior y superior, respectivamente (los “bigotes”). Más allá de dichos límites los datos se consideran atípicos.
En concreto, los límites son calculados de la siguiente manera:
-límite inferior= $Q1-1.5 \times R.I.$
-límite superior= $Q3+1.5 \times R.I.$
donde Q1 y Q3 son el primer y el tercer cuartil, respectivamente, y R.I. es el rango intercuartílico, es decir, $R.I.=Q3-Q1$, la altura de la caja.

Tras esto, el Análisis de Conglomerados, mediante uno de sus algoritmos, el k-medias, clasifica las empresas en un número de categorías prefijadas. En nuestro caso agrupamos las empresas en tres categorías prefijadas, pequeñas, medianas y grandes. En este punto es necesario explicar por qué no incluimos la categoría de empresas micro. Otro de los supuestos básicos del Análisis de Conglomerados es la representatividad de la muestra. Nuestra muestra, la ESEE, tiene como población de referencia las empresas con 10 o más trabajadores, pero incluye también información sobre una pequeña proporción de empresas con menos de 10 trabajadores (de las cuales la mayoría serían catalogadas como micro-empresas, según los criterios de la Comisión Europea). Ello resta representatividad la muestra, a cambio de sólo añadir un número muy pequeño de observaciones. Por lo tanto en nuestro análisis decidimos excluir toda empresa con menos de 10 trabajadores, dejando así de lado la categoría micro.

El algoritmo de k-medias nos da tres categorías con los siguientes valores:

- Empresas pequeñas: de 0 a 38.25
- Empresas medianas: de 38.26 a 56.49
- Empresas grandes: de 56.50 a 100

Una de las críticas más habituales al algoritmo de k-medias es su falta de robustez, es decir, la posibilidad de generar clasificaciones distintas siendo implementado en la misma muestra, dado que ha de empezar con k centroides aleatorios. Dicho problema suele desaparecer cuando la muestra es suficientemente grande –como es nuestro caso– pues las clasificaciones convergen a unos valores comunes. Para comprobar que éste fuera el caso, el algoritmo se ha repetido 50 veces. Los resultados no dejan lugar a dudas: en todas las ocasiones los valores que separan a las 3 categorías de empresas son exactamente los mismos, es decir, la convergencia es total.

Recuadro metodológico 7: el Análisis de Conglomerados y el algoritmo de k-medias

El Análisis de Conglomerados permite clasificar los individuos pertenecientes a una determinada población en un conjunto de grupos o clusters, a partir de ciertas características observables, es decir, variables.

Existen dos tipologías básicas dentro del análisis de clusters, los métodos jerárquicos y los métodos partitivos. Los métodos jerárquicos configuran grupos con estructura arborescente, de forma que los clusters de niveles más bajos van siendo englobados en otros de niveles superiores. Por el contrario, los métodos no jerárquicos o partitivos asignan los casos a clusters diferenciados que el propio análisis configura, sin que unos dependan de otros. En la investigación aplicada, los métodos jerárquicos se suelen utilizar para muestras pequeñas, mientras que los métodos partitivos permiten el análisis de gran cantidad de observaciones.

Dentro de los métodos partitivos, el análisis más utilizado es el K-medias. El algoritmo K-medias asigna cada observación al cluster cuyo centro (también llamado centroide) está más cerca de dicha observación. El centroide no es más que la media de todas las observaciones del cluster (en el caso de un análisis multivariante, el centroide será el vector de las medias de cada una de las variables). La distancia se puede medir de diversas maneras, siendo la más usual la distancia Euclídea. La distancia Euclídea no es más que la distancia geométrica entre dos puntos. En el caso univariante estos dos puntos se sitúan en una recta, mientras que en el caso de dos variables se sitúan en el habitual mapa de coordenadas X/Y. El principal objetivo de este algoritmo es minimizar la variabilidad *dentro* de cada grupo y maximizar la variabilidad *entre* los diversos grupos, es decir, generar grupos tan homogéneos y al mismo tiempo tan diferentes de los otros grupos como sea posible.

Los pasos del algoritmo son los siguientes:

Elegir el número de clusters, k.

Generar k centroides iniciales.

Asignar cada una de las observaciones al cluster más cercano.

Calcular los k nuevos centroides.

Repetir los pasos previos hasta que se alcance cierto criterio de parada.

Si el criterio de parada es alcanzado, el algoritmo se da por concluido.

Respecto al paso 2.), cabe reseñar que los k centroides iniciales pueden ser:

-aleatorios: k observaciones de la muestra, tomadas al azar, k valores sacados de una distribución uniforme o los centroides de k clusters generados al azar.

-no aleatorios: las primeras k observaciones o las últimas k observaciones de la muestra.

En la literatura se suele aconsejar usar los aleatorios para que los centroides iniciales no se vean influidos por el orden de los datos, sobre todo si la muestra es suficientemente grande. Cuando la muestra no lo es, éstos tienen una desventaja fundamental: no permiten conseguir la misma clasificación cada vez que el análisis es implementado sobre los mismos datos, dado que los k centroides iniciales son cada vez distintos.

Respecto al paso 5), ejemplos de criterios de parada son “que ninguna iteración más pueda ya disminuir la variabilidad dentro de cada grupo”; “un número limitado de pasos de iteración”; o “que la diferencia obtenida entre los centroides de dos iteraciones consecutivas sea menor que un valor prefijado.”

En la Tabla 11 podemos observar la distribución de las empresas según su tamaño con esta clasificación, mientras que en la Tabla 12 podemos ver la distribución de las empresas según el criterio de tamaño de la Comisión Europea, una vez que eliminamos la categoría de “micro”.

Tabla 11: distribución de las empresas según su tamaño con la clasificación alternativa.

	Número	%
pequeñas	2,554	30.20
medianas	2,625	31.04
grandes	3,279	38.77
Total	8,458	100

Tabla 12: distribución de las empresas según su tamaño con la clasificación de la Comisión Europea.

	Número	%
pequeñas	3,969	46.83
medianas	2,070	24.42
grandes	2,437	28.75
Total	8,476	100

La comparación de estas dos Tablas nos hace ver que, para nuestra muestra de empresas manufactureras españolas, el criterio de la Comisión Europea subestima el tamaño de las mismas. Por ejemplo, la proporción de Pymes según nuestra clasificación es algo superior al 61%, mientras que según la Comisión Europea es de más de un 71%, puesto que la proporción de empresas grandes es un 10% mayor en nuestra clasificación. Como podemos observar en la Tabla 13, algo más de un 35 % de las empresas clasificadas como pequeñas según la Comisión Europea pasan a ser medianas con la clasificación alternativa. Asimismo, poco más de un 40% de las empresas clasificadas como medianas según la Comisión Europea pasan a ser grandes con la clasificación alternativa.

Tabla 13: comparación entre la clasificación de la Comisión Europea y la clasificación alternativa.

		CLASIFICACIÓN ALTERNATIVA			Total
		pequeñas	medianas	grandes	
CLASIFICACIÓN CE	pequeñas	2,554 64.64%	1,395 35.31%	2 0.05%	3,951 100%
	medianas	0 0%	1,230 59.42%	840 40.58%	2,070 100%
	grandes	0 0%	0 0%	2,437 100%	2,437 100%

Sin embargo, si bien la proporción de empresas en cada categoría cambia ligeramente, las principales conclusiones que obtenemos con ambas clasificaciones son las mismas. Por ejemplo, realizando el análisis explicado en la sección 2.1 con la nueva clasificación produce resultados muy parecidos. Las mismas variables que resultan ser determinantes significativos de profesionalización e internalización de las empresas que se definen como Pymes con la clasificación del Comisión Europea, lo son también con esta clasificación, y el signo de su relación es también el mismo.

Referencias

- Batista Foguet, J.M. y M^a.R. Martínez Arias (1989). *Análisis multivariante: análisis en componentes principales*; Editorial Hispano Europea, Barcelona.
- Berger, A.N., N. Klapper, y G. F. Udell, (2001) "The ability of banks to lend to informationally opaque small businesses" *Journal of Banking and Finance*, 25, 2127–2167. Allen N.
- Berger, A. N., R. J. Rosen, y G. F. Udell (2007) "Does market size structure affect competition? The case of small business lending", *Journal of Banking & Finance*, 31, 11–33.
- Berger, A. N. y G. F. Udell (2006) "A more complete conceptual framework for SME finance", *Journal of Banking & Finance*, 2945–2966
- De Vicente y Oliva, M^a A., J. Bassa Manera y F.J. Blanco Jiménez (2000). *Análisis multivariante para las ciencias sociales*; Dykinson, Madrid.
- Peña, D. (1999). *Estadística: modelos y métodos*; Alianza Universidad Textos.
- Scott, J. A. y W. C. Dunkelberg (2003) "Bank Mergers and Small Firm Financing", *Journal of Money, Credit and Banking*, 35, 999-1017.