

**UN ANÁLISIS EMPÍRICO SOBRE LA INCIDENCIA DEL
TAMAÑO EMPRESARIAL DE LAS PYMES CATALANAS
MEDIANTE RATIOS FINANCIEROS**

**Área de investigación: Administración de la micro, pequeña
y mediana empresa.**

Gerardo Reyes Ruiz

Universidad Autónoma del Estado de México
greyesruiz@hotmail.com

Alejandro Barragán Ocaña

Universidad Autónoma del Estado de México

Samuel Olmos Peña

Universidad Autónoma del Estado de México

XIX
CONGRESO
INTERNACIONAL
DE
CONTADURÍA
ADMINISTRACIÓN
E
INFORMÁTICA



Octubre 8, 9 y 10 de 2014 ♦ Ciudad Universitaria ♦ México, D.F.



ANFECA
Asociación Nacional de Facultades y
Escuelas de Contaduría y Administración



UN ANÁLISIS EMPÍRICO SOBRE LA INCIDENCIA DEL TAMAÑO EMPRESARIAL DE LAS PYMES CATALANAS MEDIANTE RATIOS FINANCIEROS

Resumen

El presente trabajo pretende explicar el tamaño empresarial de una muestra de PyMES españolas a través de un Análisis Discriminante, el cual considera como variables de insumo las dimensiones sintéticas generadas, a su vez, por un Análisis de Componentes Principales con base en una batería de determinados *ratios* financieros. La metodología propuesta en este artículo utiliza información de primera mano para explicar la variable de clasificación denominada tamaño para un determinado conglomerado de empresas españolas a través de su desempeño financiero. La valoración de la clasificación obtenida fue más que aceptable, sin embargo, se recomienda el uso de otras técnicas, como las redes neuronales, para reafirmar los resultados obtenidos.



UN ANÁLISIS EMPÍRICO SOBRE LA INCIDENCIA DEL TAMAÑO EMPRESARIAL DE LAS PYMES CATALANAS MEDIANTE RATIOS FINANCIEROS

1. Introducción

El Análisis de Estados Financieros¹, según Urías (1992), trata de investigar y enjuiciar, a través de la información contable, cuáles han sido las causas y los efectos de la gestión de la empresa para llegar a su actual situación y, así, predecir, dentro de ciertos límites, cuál será su desarrollo en el futuro, para tomar decisiones consecuentes. Desde la década de los 90 se propuso que el análisis de estados financieros se puede dividir en tres etapas: (1) examen, (2) análisis e interpretación, y (3) prescripción (Rivero 1991; Urías, 1992). Así, en la primera etapa, se enfatiza la verificación o revisión de los documentos contables, como son el balance, cuenta de pérdidas y ganancias, estado de origen, entre otros, utilizando para este fin una serie de procedimientos específicos, tales como: comparaciones de masas patrimoniales, porcentajes, números índices, *ratios*, tendencias, etc. Posteriormente, se analiza la información obtenida en la primera etapa para con ello emitir conclusiones de la situación actual de la empresa y, por supuesto, de la gestión realizada. Por último, en la tercera etapa, se realizan predicciones y se proponen soluciones con el fin de mejorar la situación de una determinada empresa en el corto y mediano plazo.

1.1 Objetivo del uso de *ratios* financieros

El uso de *ratios* financieros tiene por objetivo, grosso modo, facilitar la interpretación de ciertos estados financieros y, con ello, transcribir la información proporcionada en índices y/o porcentajes, que permitan, a su vez, desarrollar la mejor aproximación respecto a la situación financiera de la empresa, para con ello cimentar el denominado diagnóstico financiero.

Los *ratios* son expresiones aritméticas, establecidas a partir de cifras económicas o financieras extraídas, en este contexto, de la información contable. Estas expresiones pueden ser: a) Simples (entre cifras de un mismo tipo de balance, entre agregados de un cierto estado de resultados, etc.) y, b) Compuestas (entre cifras o agregados de diversas fuentes).

1.2 Las precauciones del uso de *ratios*

Un *ratio*, en términos muy generales, expresa una razón. Es decir, la relación entre dos números. En el ámbito contable, los *ratios* son un conjunto de índices, los cuales son el resultado de relacionar, al menos, dos cuentas del Balance o del Estado de Pérdidas y Ganancias. Los *ratios* suministran información de suma importancia, la cual tiene la finalidad de coadyuvar en la toma de las decisiones, generalmente acertadas, a quienes estén interesados en una determinada empresa. Como ya se ha señalado con anterioridad, los *ratios* corresponden a una expresión aritmética elemental, por consiguiente, su cálculo es relativamente sencillo. No obstante, la relativa simplicidad de su determinación puede transformar su uso en una simple aplicación técnica, lo que puede derivar en la indiferencia

¹ Se agradece al Servicio de Estudios de Pimec (Micro, petita i mitjana empresa de Catalunya) la cesión de la base de datos para realizar el presente trabajo. www.pimec.es



de lo que realmente es su aporte, es decir, su interpretación (Aching, 2005). Además, un error recurrente que se comete al tenor de facilitar el trabajo de los analistas, es buscar normas previamente fijadas o estándares para una interpretación más rápida y más cómoda pero que, en general, no logra el objetivo deseado al hacer uso de *ratios* financieros. Ello debido a que las realidades, y necesidades, del entorno de cada empresa son distintas.

Otro tema que por su trascendencia e impacto merece ser resaltado es que, la facilidad de construir o de crear *ratios* ha conllevado a que éstos sean cada vez más numerosos, lo que sin duda, genera un nuevo problema, y el cual se resume en la siguiente pregunta: ¿Cómo seleccionar los *ratios* a utilizar?² Este cuestionamiento es muy difícil de responder, ya que es una tarea complicada y que dependerá de distintos factores, entre ellos: el objetivo del análisis, la perspectiva del analista, la realidad económica preponderante, en la cual será fundamental la experiencia del analista para poder definir una tipología o batería de *ratios* que sea relevante para la realidad que se desea analizar.

1.3 Relaciones lógicas en la formación de *ratios*

Ya se ha dicho que los *ratios* son expresiones aritméticas, pero cuando se habla de la construcción de *ratios*, la literatura especializada define la existencia de algunas relaciones lógicas para los elementos a considerar entre ellos. Por ello, se han mantenido hasta el día de hoy tres tipos de relaciones lógicas a considerar cuando se desean construir *ratios* (Lev, 1974), y las cuales son:

- a. Es primordial que exista una relación económica entre los componentes del *ratio*.
- b. Sus componentes deben estar expresados en valores comunes.
- c. Se requiere una relación funcional entre sus componentes.

1.4 La definición de PyME

Uno de los principales problemas para estudiar la existencia de diferencias en la estructura, tanto económica como financiera, entre los distintos tamaños empresariales es el criterio para clasificar las empresas en medianas, pequeñas o micro. La problemática no sólo gira en torno de la variable que se debe utilizar, sino también en los límites o puntos de corte de cada una de las variables consideradas³.

Así, la revisión de los criterios utilizados para los efectos de esta investigación empírica permite realizar las siguientes observaciones: I) Los criterios utilizados son, en todos los casos, unidimensionales y los dos enfoques clásicos utilizados son: el número de empleados o la cifra de ventas. No obstante, el primer enfoque clásico presenta el inconveniente de que al utilizar tan solo un factor productivo (el trabajo), sin considerar la tecnología de cada sector, se infravalora el tamaño de las empresas basadas en el capital y se sobrevaloran aquellas cuya base es, precisamente, la mano de obra. Además, no se considera el trabajo temporal ni el

² En el trabajo de Sosa (2011) se propone una serie de *ratios* para ser utilizados en un análisis financiero empresarial y que pretenden dar respuesta a dicho cuestionamiento.

³ Sin duda existen múltiples factores que bien pudiesen influir para clasificar a una empresa, como: otros *ratios* sintéticos, el efecto territorio y el efecto entorno, particularmente el económico y el político, por citar algunos.



trabajo no remunerado, como por ejemplo el trabajo de los socios corporativos. El segundo enfoque clásico presenta la disyuntiva de que, además de captar el tamaño, incluye el distinto grado de eficiencia de las empresas, es decir, dos empresas con el mismo número de trabajadores y el mismo volumen de activo pudiesen pertenecer, como consecuencia del distinto grado de eficiencia en la organización del proceso productivo, a dos grupos distintos si su volumen de ventas fuese diferente, II) No existe un acuerdo general sobre los límites que se deben establecer dentro de cada criterio, por ello, en algunos trabajos se consideran a las PyMES como aquellas empresas con menos de 100 trabajadores, mientras que en otros se consideran a las que tienen hasta 500 trabajadores y, III) El uso de distintos criterios y puntos de corte para un mismo criterio dificulta la comparación de los resultados obtenidos en los múltiples trabajos realizados hasta el día de hoy. Por supuesto, esto impide la implementación, con cierto grado de rigor, de las regularidades existentes en los criterios de financiamiento para las distintas clases de empresas según su tamaño.

Esta falta de acuerdos también existe en las distintas definiciones de las PyMES, establecidas tanto por instituciones diversas así como por parte de las administraciones públicas, para delimitar a las empresas que pueden tener acceso a determinados tipos de subvenciones. Por ello, y con la finalidad de unificar criterios, la Unión Europea (UE) hace la siguiente recomendación: a) Que empleen a menos de 250 personas, b) Que el volumen de facturación no supere los 40 millones de euros o cuyo balance anual no exceda de los 27 millones de euros y, c) Que cumpla el criterio de independencia, es decir, que el 25% o más de su capital o de sus derechos de voto no pertenezcan a otra empresa o, conjuntamente, a varias empresas que no respondan a la definición de PyME.

Cuando se requiere distinguir entre empresas⁴ pequeñas y medianas, se define como empresa pequeña aquella que: a) No tenga más de 50 trabajadores, b) Tenga un volumen de negocio no superior a los 7 millones de euros o un balance general no superior a los 5 millones de euros y, c) Cumpla con el criterio de independencia (Calvo y Boedo, 2001).

2. Objetivo

El presente trabajo empírico pretende mostrar una clasificación de empresas catalanas respecto a su tamaño mediante el análisis de *ratios* financieros. El objetivo⁵ es exponer un criterio de medida del tamaño empresarial. Para ello se omitirá el enfoque clásico, el cual hace referencia principalmente al número de empleados (Galindo, 2005), y en su lugar se utilizará el enfoque de las denominadas técnicas multivariantes, las cuales servirán primero, para generar variables sintéticas que posteriormente serán incorporadas a un segundo análisis y, de esta manera, construir una medida del tamaño empresarial. Ello debido principalmente a que el establecimiento de límites cuantitativos, en términos de variables no responde, en general, al empleo de ningún método que discrimine por qué una empresa pertenece al conjunto de micro, pequeña o mediana.

⁴ En este trabajo de investigación se define lo siguiente: 1= micro empresa; 2= pequeña empresa y 3= mediana empresa.

⁵ La finalidad de esta investigación no es contable, más bien este trabajo muestra que diferentes técnicas de análisis de datos sirven para clasificar a un conglomerado de empresas en micro, pequeñas y medianas.



3. Datos

La base de datos⁶ está integrada por 12,658 PyMES de Cataluña, y para cada una de ellas se definieron 22 *ratios* financieros, los cuales se muestran en la tabla 1. Estos *ratios* financieros corresponden únicamente para el año 2002. Sin embargo, consideramos que la definición del año no limita los resultados obtenidos en este trabajo de investigación⁷. Posteriormente, y en la segunda etapa del presente análisis, se utilizaron las variables denominadas tamaño y número de empleados.

Tabla 1
Ratios financieros

No.	Concepto	Etiqueta	Ratio
1	Activo circulante	act_circ	Act circ explot / Actiu net
2	Activo neto	act_net	Act net explot / Actiu net
3	Rentabilidad de las ventas	altres_d	Altres desp explot / Ing explot
4	Liquidez	cash_flo	Cash Flow / creditors a c/t
5	Operación	consums_	Consums explot / Ing explot
6	Créditos	creditor	Creditors c/t / Passiu remunerat
7	Gastos de personal por empleado	desp_per	Desp personal / n treballadors
8	Fondos propios	fons_pro	Fons propis / passiu remunerat
9	Financiación ajena del inmovilizado	immob_ne	Immob net explot / Actiu net
10	Margen	marge	Marge
11	Endeudamiento	rec_ali	Rec aliens ll't / passiu remunerat
12	Recursos generales	recursos	Recursos Generats
13	Rentabilidad económica	rendibl	Rendibilitat econòmica
14	Rentabilidad financiera	rendibil	Rendibilitat financera
15	Cobertura de Gastod financieros	rene_d	Rene / Desp finances
16	Coberetura de efectivo	rene_i	Rene / Ing explot
17	Rotación	rotació	Rotació
18	Solvencia	tresorer	Tresoreria/Ingressos explot
19	Productividad del activo	vab_ac	VAB / Act net explot
20	Productividad del inmovilizado	vab_im	VAB / Immob net explot
21	Productividad del ingreso	vab_in	VAB / Ing explot
22	VAB (valor añadido bruto) por ocupante	vab_nom	VAB/ Nombre treb

Fuente: Elaboración propia con datos de Pimec.

4. Metodología

Para la primera fase del presente trabajo se aplicó el Análisis de Componentes Principales (ACP) a la batería de *ratios* financieros que resultaron elegidos de la tabla 1. Este método multivariante tiene como principal objetivo convertir un problema de información estadística muy compleja (muchas variables de tipo cuantitativo calibradas en diferentes unidades de medición) en otro casi equivalente pero más manejable (pocas nuevas variables) sin pérdida significativa de información (E. Pla, 1986).

La utilidad de un ACP es doble, porque por una parte permite representar óptimamente en un espacio de dimensión pequeña observaciones de un espacio general de mayor

⁶ El total de empresas hace referencia a diferentes sectores, y como la finalidad de esta investigación es pronosticar el tamaño empresarial entonces no se considera la diferencia por sectores o actividades productivas, ni por categorías de profesionales independientes, de pequeños negocios minoristas, bares, restaurantes, etc. Ello, sin lugar a dudas, es tema para trabajos futuros.

⁷ Las variables externas, como por ejemplo el entorno económico o el tiempo, que bien pudiesen contribuir para definir el tamaño empresarial son poco influyentes en este análisis, debido a que éste tan solo analiza información directa de cada empresa.



dimensionalidad. En este sentido, el ACP es el primer paso para identificar las posibles variables latentes, o no observadas, que generan los datos. Y por otra parte, permite transformar las variables originales, en general correlacionadas, en nuevas variables incorrelacionadas, facilitando así la interpretación de los datos (Fisher, 1936).

Con este tipo de análisis multivariante se obtienen combinaciones lineales de todas las variables originales que sean ortogonales entre sí. De esta forma se obtiene un subconjunto de nuevas variables que son independientes entre sí, pero que están altamente correlacionadas con las variables originales, denominadas Componentes Principales, las que ordenadas en forma decreciente según su varianza, permiten explicar el porcentaje de variabilidad de los datos, por cada componente (Montanero, 2008). Consecuentemente, cada observación de la muestra obtiene una puntuación en cada una de las componentes principales seleccionadas, lo que permite ordenar las observaciones en base a una información multivariante.

En este trabajo se obtuvieron las componentes principales diagonalizando la matriz de correlaciones, ello debido a que los *ratios* tienen distintas unidades de medida y cada variable participa con la misma importancia. Una vez generadas las variables latentes, éstas fueron el insumo en un Análisis Discriminante (AD), para con ello explicar con una medida sintética el tamaño empresarial de las PyMES catalanas. Es decir, con este análisis se pretende encontrar un patrón para separar a un grupo de empresas con relación a su tamaño. Se utilizó el AD debido a que es una técnica estadística que clasifica objetos en diferentes grupos basándose en las observaciones de algunas características de dichos objetos (Shih-Wei y Shih-Chieh, 2009). Básicamente consiste en obtener una combinación lineal de algunas variables independientes (esto lo garantiza el ACP previamente realizado), que maximicen las diferencias entre los grupos de clasificación y minimicen las diferencias entre cada uno de estos grupos (máximo valor del cociente definido por variabilidad entre grupos/variabilidad intra grupos). Esta función discriminante es una combinación lineal ponderada de los valores que toman las características observadas del objeto. Las ponderaciones representan, en esencia, la importancia relativa y el impacto de las características que forman parte de la función discriminante. Por lo tanto, el objeto se clasifica en un solo grupo o clase basándose en la puntuación obtenida con la función discriminante.

No obstante, el AD está sujeto a un gran número de hipótesis tales como: cada grupo debe de seguir una distribución normal multivariante, las matrices de covarianzas de cada grupo son idénticas y, las matrices de covarianzas, las probabilidades a priori y los errores de clasificación han de ser conocidos. Si no se cumplen estas hipótesis los resultados obtenidos pueden ser erróneos. Así, cabe mencionar que se utilizó el AD debido a que con este método se seleccionan las variables que posean mayor poder discriminante, las cuales previamente fueron generadas por un ACP.

5. Análisis de los *ratios* financieros

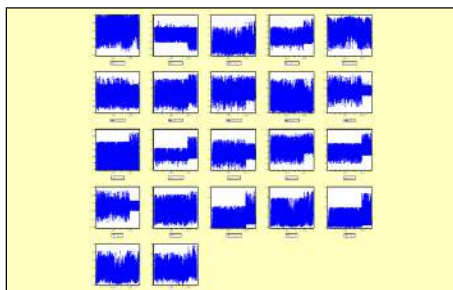
Trabajar con *ratios* financieros siempre será complicado, debido principalmente a la construcción de los mismos, ya que es muy común que existan problemas de dependencia lineal entre varios de ellos. Para los *ratios* que se propusieron en la tabla 1 el análisis fue el siguiente:



Por su construcción se apreciaron cuatro grupos. Sin embargo, el primer problema que mostraron estos *ratios* es que no siguen un comportamiento de una distribución normal univariante, por lo que de momento queda descartada toda inferencia estadística paramétrica -como el Análisis de Varianza (ANOVA) e incluso el coeficiente de correlación de Pearson, el cual supone normalidad para su cálculo- que se quisiera llevar a cabo, puesto que todo resultado llevaría a conclusiones erróneas. Si se supone normalidad en los *ratios* se obtendría un ANOVA que nos implicaría primero, la inexistencia de igualdad de varianzas entre los grupos del *ratio* tamany y segundo, que los *ratios* susceptibles a omitir en determinado momento serían: rotació, rec_alie, immob_ne y act_circ. Ello debido a que son los *ratios* que presentan la existencia de diferencias significativas entre los diferentes subgrupos en los que se han separado al factor tamany.

El siguiente problema fue el alto grado de multicolinealidad que presentan dichos *ratios*, lo cual se verificó al realizar un ACP con todos ellos, ya que se obtuvo una matriz sintética de correlaciones la cual no está definida positivamente. Ello implica la discriminación de algunos *ratios* para la obtención de resultados más robustos. Gráficamente resulta complicado hacer algún juicio de valor, ya que como puede apreciarse en la gráfica 1, no se perciben relaciones directas *a priori*.

Gráfica 1
Representación múltiple de los 22 *ratios* financieros



Fuente: Elaboración propia con base en Eviews.

El problema de la dependencia lineal se puede comprobar al hacer una regresión lineal múltiple, sin el término constante, utilizando cualquiera de los 22 *ratios* como variable endógena y el resto de ellos como variables exógenas. El resultado será siempre el mismo, es decir, no se puede llevar a cabo dicha regresión porque se obtendrá el mensaje de “Near Singular Matrix”. Es decir, algunos de los *ratios* son combinación lineal exacta de otras razones financieras, lo cual resulta lógico. Esta última relación lineal pudo apreciarse omitiendo una variable, por ejemplo el *ratio* denominado act_circ y, entonces, al explicar el *ratio* altres_d utilizando el resto de los *ratios* se obtuvo el resultado de la tabla 2.

Tabla 2
Dependencia lineal entre los *ratios*

Variable	Coefficient	Std. Error	t-Statistic	Prob.
CASH_FLO	-9.63E-14	1.10E-14	-8.744889	0.0000
CONSUMS_	-1.000000	1.14E-14	-8.75E+13	0.0000
CREDITOR	1.000000	3.26E-14	3.07E+13	0.0000
DESP_PER	-5.32E-15	3.21E-16	-16.58892	0.0000
FONS_PRO	1.000000	3.29E-14	3.04E+13	0.0000
IMMOB_NE	4.27E-15	7.61E-15	0.561940	0.5742
MARGE	-1.52E-11	2.35E-13	-64.64623	0.0000
REC_ALIE	1.000000	3.28E-14	3.05E+13	0.0000
RECURSOS	-2.97E-16	2.77E-17	-10.69459	0.0000
RENDIB1	1.61E-12	6.23E-14	25.88477	0.0000
RENDIBIL	-8.71E-14	8.24E-15	-10.57439	0.0000
RENE_D	6.05E-15	4.43E-16	13.64236	0.0000
RENE_J	1.21E-11	2.17E-13	55.69137	0.0000
ROTACI	4.21E-14	2.53E-15	16.61078	0.0000
TRESORER	4.56E-14	1.72E-14	2.650537	0.0080
VAB_AC	-2.53E-13	7.97E-15	-31.74186	0.0000
VAB_IM	-9.69E-15	5.08E-16	-1.904872	0.0568
VAB_IN	-1.000000	1.98E-14	-5.06E+13	0.0000
VAB_NOM	3.46E-15	2.43E-16	14.25310	0.0000
ACT_NET_	-2.79E-12	3.21E-14	-86.60744	0.0000

R-squared	1.000000	Mean dependent var	0.143915
Adjusted R-squared	1.000000	S.D. dependent var	0.080202
S.E. of regression	9.40E-14	Akaike info criterion	-57.15134
Sum squared resid	1.12E-22	Schwarz criterion	-57.13958
Log likelihood	381730.9	F-statistic	4.85E+26
Durbin-Watson stat	1.963626	Prob(F-statistic)	0.000000

Fuente: Elaboración propia con base en Eviews.

Con base en este análisis se observó una relación lineal muy alta entre casi todos los *ratios* involucrados, debido a que se obtuvo un coeficiente de determinación igual a uno y casi todos los estadísticos *t*'s fueron mayores que dos, es decir, resultaron ser significativos. Además, debido a que la prueba *d* de Durbin–Watson no es válida en el modelo de regresión sin el término constante, entonces no se pudieron hacer juicios de valor sobre la correlación de primer orden.

Dado que no se pudo asumir el supuesto de normalidad para un ANOVA, se procedió a utilizar pruebas no paramétricas, las cuales, aunque son menos potentes, permitieron llevar una discriminación entre los *ratios* y lo más importante, no hacen ningún supuesto sobre normalidad entre los datos. No obstante, al verificar la independencia de dichos *ratios* mediante el test de rachas se verificó que en ninguno de ellos se sostiene la hipótesis de aleatoriedad y, por lo tanto, de independencia. Este último resultado implicó que el muestreo de las empresas no fue aleatorio y que las pruebas no paramétricas pueden llevar a conclusiones erróneas. Sin embargo, y a manera de ejemplo, si se asumiera la independencia de estos *ratios*, al realizar un test de Kruskal-Wallis para *k* muestras independientes entonces se obtendría que los *ratios* denominados rotació, rec_alie, immob_ne, act_circ y tesorero muestran diferencias significativas entre las medias aritméticas para los grupos del *ratio* tamany, lo que podría sugerir la no inclusión de dichos *ratios* en el estudio. Por su parte, al realizar un test de la mediana se obtendrían los siguientes *ratios* con diferencias significativas entre los grupos del *ratio* tamany: rotació, rec_alie, immob_ne, act_circ, altres_d, tesorero y rene_d. Este último análisis bien podría llevar a una mala selección de *ratios*.

Así, y para la discriminación de todos los *ratios*, se optó por un análisis de distancias entre los *ratios* de cada grupo, con respecto al *ratio* tamany. Para ello se utilizó la distancia de Chebychev, estandarizada por su desviación típica para cada *ratio* y en valores absolutos, y de esta manera obtener una mejor apreciación de dicha distancia. Los resultados, para los *ratios* de la tabla 1 fueron los siguientes:



1. Para el bloque A, el *ratio* que mostró una mayor distancia con respecto al *ratio* tamany fue el denominado rec_alie, por lo que este *ratio* puede ser omitido para el correspondiente ACP.
2. En el bloque B, se obtuvo que el *ratio* denominado act_net_ presentó la mayor distancia con respecto al *ratio* tamany, por lo que también puede ser omitido para el ACP.
3. Para el bloque C, pasó algo poco esperado, se obtuvo que el *ratio* denominado rene_i mostró la mayor distancia con respecto al *ratio* tamany. Sin embargo, dicho *ratio* no agregaba información al ACP, ya que se obtenía una matriz de correlaciones que no estaba definida positivamente. Por su parte, el *ratio* denominado altres_d que mostraba la menor distancia con respecto al *ratio* tamany, al ser omitido se obtuvieron resultados más robustos en el ACP. Por lo que considerando estos resultados se optó por omitir al *ratio* denominado altres_d.

Con este análisis la batería de *ratios* financieros seleccionados, y que posteriormente serán el insumo para el ACP, se presenta en la tabla 3.

Tabla 3
Agrupación de los *ratios* financieros seleccionados

No.	Concepto	Etiqueta	Ratio
A	1 Fondos propios	fons_pro	Fons propis / passiu remunerat
	2 Créditos	creditor	Creditors c/t / Passiu remunerat
B	3 Financiación ajena del inmovilizado	immob_ne	Immob net explot / Actiu net
	4 act_circ	act_circ	Act circ explot / Actiu net
C	5 Operación	consums	Consums explot / Ing explot
	6 Productividad del ingreso	vab_in	VAB / Ing explot
	7 Coberetura de efectivo	rene_i	Rene / Ing explot
	8 Solvencia	tresorer	Tresoreria/Ingressos explot
D	9 Rentabilidad financiera	rendibil	Rendibilitat financera
	10 Rentabilidad económica	rendib1	Rendibilitat econòmica
	11 Margen	marge	Marge
	12 Rotación	rotació	Rotació
	13 Productividad del inmovilizado	vab_im	VAB / Immob net explot
	14 Recursos generales	recursos	Recursos Generats
	15 Liquidez	cash_flo	Cash Flow / creditors a c/t
	16 Cobertura de Gastod financieros	rene_d	Rene / Desp finances
	17 Productividad del activo	vab_ac	VAB / Act net explot
	18 VAB (valor añadido bruto) por ocupante	vab_nom	VAB/ Nombre treb
	19 Gastos de personal por empleado	desp_per	Desp personal / n treballadors

Fuente: Elaboración propia.

6. Resultados

En este apartado se presentan los principales resultados obtenidos a través de un ACP y posteriormente del AD. El ACP es una técnica utilizada para la extracción de factores y formar combinaciones lineales independientes de las variables observadas. La primera componente tiene la varianza máxima. Las componentes sucesivas explican progresivamente proporciones menores de la varianza y no están correlacionadas las unas con las otras. El ACP se utiliza para obtener la solución factorial inicial. Puede utilizarse también cuando una matriz de correlaciones es singular. En el ACP, llevado a cabo para los *ratios* de la tabla 3, se obtuvo la siguiente matriz de comunalidades con la extracción de 6 componentes (véase Gráfica 2) o variables latentes:



Tabla 4
Matriz de comunalidades para los *ratios* financieros seleccionados

	Inicial	Extracción
Rendibilitat financera	1,000	,722
Rendibilitat econòmica	1,000	,878
Marge	1,000	,919
Rotació	1,000	,934
Fons propis / passiu remunerat	1,000	,844
Creditors c/t / Passiu remunerat	1,000	,840
Immob net explot / Actiu net	1,000	,934
Act circ explot / Actiu net	1,000	,933
Consums explot / Ing explot	1,000	,922
VAB / Ing explot	1,000	,956
VAB / Immob net explot	1,000	,799
VAB / Act net explot	1,000	,943
Desp personal / n treballadors	1,000	,923
VAB/ Nombre treb	1,000	,937
Tresoreria/Ingressos explot	1,000	,269
Recursos Generats	1,000	,382
Cash Flow / creditors a c/	1,000	,735
Rene / Ing explot	1,000	,911
Rene / Desp financeres	1,000	,523

Método de extracción: Análisis de Componentes principales

Fuente: Elaboración propia con base en SPSS.

Es conveniente mencionar que el ACP fue más enriquecedor que un Análisis Factorial (AF), desde la perspectiva de un mayor número de variables latentes, debido a que el AF generaba únicamente cuatro factores latentes, explicando con ellos un 64.3% de la varianza total de los *ratios* seleccionados. Más aún, si el objetivo principal hubiese sido disminuir la dimensión lo más posible, entonces el AF habría dado mejores resultados, ya que con cuatro factores latentes se hubiesen obtenido los mismos estadísticos de prueba que en un ACP (véase Tabla 5).

Tabla 5
KMO y prueba de Bartlett para un AF

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,698
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	260998,4
	gl	171
	Sig.	,000

Fuente: Elaboración propia con base en SPSS.

Es de suma importancia resaltar que la factorización de ejes principales (Análisis Factorial) es un método utilizado para la extracción de factores que parte de la matriz de correlaciones original con los cuadrados de los coeficientes de correlación múltiple insertados en la diagonal principal como estimaciones iniciales de las comunalidades. Las saturaciones factoriales resultantes se utilizan para estimar de nuevo las comunalidades y reemplazan a las estimaciones previas en la diagonal de la matriz. Las iteraciones continúan hasta que el



cambio en las comunalidades, de una iteración a la siguiente, satisfaga el criterio de convergencia para la extracción. Sin embargo, lo que se llevó a cabo en este trabajo fue lo siguiente; a través de un mayor número de variables sintéticas posibles se planteó el objetivo de encontrar el mayor nivel de asociación de dichas variables sintéticas con respecto al tamaño de las empresas consideradas. Por lo tanto, se optó por adquirir los resultados del ACP. Así, se tomaron las puntuaciones del ACP, las cuales tienen una media 0 y una varianza igual al cuadrado de la correlación múltiple entre las puntuaciones estimadas y los valores verdaderos. Estas puntuaciones pueden estar correlacionadas incluso cuando los factores obtenidos sean ortogonales.

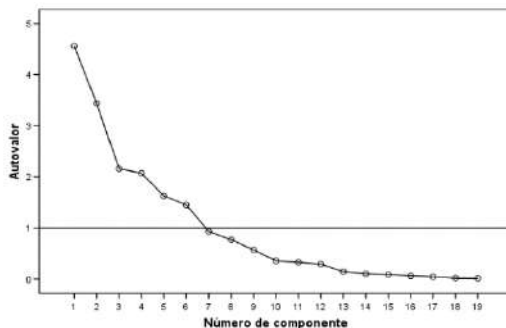
Aquí es importante comentar lo siguiente: como la matriz de correlaciones es calculada utilizando el coeficiente de correlación de Pearson, el cual es una medida de asociación lineal, entonces puede ocurrir que dos variables estén perfectamente relacionadas, pero si dicha relación no es lineal, el coeficiente de correlación de Pearson no será un estadístico adecuado para medir su asociación. Además, para el cálculo de dicho coeficiente se debe de asumir normalidad en los *ratios*. Por ello, y con la finalidad de apreciar el nivel de concordancia con el coeficiente de correlación de Spearman, el cual para su cálculo no presupone normalidad en los *ratios*, se llevó a cabo para la matriz de correlaciones de Pearson y la matriz de correlaciones de Spearman la diferencia que existe entre dichas matrices y posteriormente se calculó el determinante de dicha matriz de diferencias para conocer el nivel de “error” cometido en las mediciones.

No obstante, es de suma importancia mencionar que si se utiliza el coeficiente de correlación de Spearman no podrían hacerse juicios de valor sobre indicios de causa-efecto entre los *ratios*, debido principalmente a que este estadístico es no paramétrico. Así, el determinante calculado para la matriz de diferencias fue de 1.02763×10^{-30} , es decir, la distancia entre dichas matrices es muy cercana a cero, por lo que no afecta demasiado utilizar el coeficiente de correlación de Pearson en el ACP. De esta manera, para los *ratios* de la tabla 3 se llevó a cabo un ACP a través del cual se generaron unas variables sintéticas. Posteriormente, las puntuaciones de las componentes fueron utilizadas como insumo en un AD y obtener de esta manera una medida sintética para el tamaño de las empresas. Es decir, con el ACP se pretende disminuir la dimensión del problema mediante la creación de variables altamente correlacionadas y que, a su vez, sean el factor para diferenciar a una empresa de otra, con respecto a su tamaño.

Así, y con base en los resultados de la gráfica 2, se seleccionaron 6 componentes o variables sintéticas. La selección del número de componentes fue debido a que se usó la matriz de correlaciones, y también a que este número de componentes presentaron un valor superior a 1.



Gráfica 2
Línea de sedimentación



Fuente: Elaboración propia con resultados de SPSS.

También en el ACP se obtuvieron los estadísticos de la tabla 6 para con ello contrastar su representatividad.

Tabla 6
KMO y prueba de Bartlett

Medida de adecuación muestral de Kaiser-Meyer-Olkin.		,698
Prueba de esfericidad de Bartlett	Chi-cuadrado aproximado	260998,4
	gl	171
	Sig.	,000

Fuente: Elaboración propia con resultados de SPSS.

La medida de la adecuación muestral de Kaiser-Meyer-Olkin (KMO) contrasta si las correlaciones parciales entre las variables son pequeñas, por lo que para este estudio se obtuvo un coeficiente muy cercano al 70%, lo que implica un buen nivel de asociación entre los *ratios* seleccionados. La prueba de esfericidad de Bartlett contrasta si la matriz de correlaciones es una matriz identidad, lo cual indicaría que el modelo es inadecuado. Sin embargo, para este estudio se rechaza la hipótesis de que el determinante de la matriz de correlaciones sea igual a uno, debido a que se obtuvo una significancia menor al 5%. Estos últimos resultados, de manera conjunta, implican que el modelo es adecuado.

Con estas seis componentes se explica el 80.5% del total de la varianza para estos *ratios* (véase Tabla 7) y aunque no es relevante etiquetar estas variables sintéticas se realizó un ACP mediante una rotación de normalización de Varimax con Kaiser (véase Tabla 8).



Tabla 7
Matriz de varianza total explicada

Componente	Autovalores iniciales			Sumas de las saturaciones al cuadrado de la extracción			Suma de las saturaciones al cuadrado de la rotación		
	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado	Total	% de la varianza	% acumulado
1	4,559	23,996	23,996	4,559	23,996	23,996	3,695	19,449	19,449
2	3,439	18,098	42,094	3,439	18,098	42,094	2,837	14,933	34,382
3	2,159	11,362	53,456	2,159	11,362	53,456	2,684	14,125	48,507
4	2,070	10,894	64,350	2,070	10,894	64,350	2,445	12,867	61,374
5	1,626	8,556	72,906	1,626	8,556	72,906	2,048	10,781	72,155
6	1,450	7,630	80,536	1,450	7,630	80,536	1,592	8,381	80,536
7	,929	4,891	85,427						
8	,769	4,047	89,474						
9	,569	2,995	92,469						
10	,354	1,862	94,330						
11	,327	1,721	96,052						
12	,290	1,529	97,580						
13	,145	,766	98,346						
14	,101	,530	98,876						
15	,085	,448	99,324						
16	,059	,311	99,635						
17	,044	,234	99,869						
18	,015	,080	99,949						
19	,010	,051	100,000						

Método de extracción: Análisis de Componentes principales.

Fuente: Elaboración propia.

Los resultados para la matriz de componentes vs. la matriz de componentes rotados fueron los siguientes:

Tabla 8
Matriz de componentes vs. matriz de componentes rotados

	Matriz de componentes ^a						Matriz de componentes rotados ^a					
	Componente						Componente					
	1	2	3	4	5	6	1	2	3	4	5	6
Rendibilitat financera							,779					
Rendibilitat econòmica	,660						,883					
Marge	,739						,909					
Rotació						,789						,949
Fons propis / passiu remunerat										,912		
Creditors c/t / Passiu remunerat											,869	
Immob net explot / Actiu net		,679							,918			
Act circ explot / Actiu net		,675							,915			
Consums explot / Ing explot								,916				
VAB / Ing explot								,945				
VAB / Immob net explot									,824			
VAB / Act net explot								,823				
Desp personal / n treballadors				,692							,934	
VAB/ Nombre treb											,949	
Tresoreria/Ingressos explot			,615		,600							
Recursos Generats										,622		
Cash Flow / creditors a c/	,754						,908					
Rene / Ing explot	,738						,635					
Rene / Desp financeres												

Método de extracción: Análisis de componentes principales.
Método de rotación: Normalización Varimax con Kaiser.

a. 6 componentes extraídos
a. La rotación ha convergido en 6 iteraciones.

Nota: se han suprimido correlaciones menores a 0.6 en valor absoluto.

Fuente: Elaboración propia.



Con estos últimos resultados se ha logrado reducir la dimensión del problema y ahora se utilizarán 6 variables independientes en lugar de 19. Para las seis componentes o variables latentes se calcularon sus puntuaciones factoriales tipificadas, y con ello estandarizarlas y garantizar su homogeneidad en el Análisis Discriminante. Así, las variables que se utilizaron como insumo para el AD se definen en la tabla 9.

Tabla 9
Variables sintéticas para el AD

No.	Etiqueta	Variable
1	F01	Factor 01
2	F02	Factor 02
3	F03	Factor 03
4	F04	Factor 04
5	F05	Factor 05
6	F06	Factor 06
7	número e	Número empleados 2002
8	tamany	Tamaño empresarial

Fuente: Elaboración propia con resultados de SPSS.

Para las seis variables sintéticas, obtenidas con el ACP, y la variable que hace referencia al número de empleados en 2002, al realizar el test de Jarque-Bera se aprecian problemas de normalidad univariante (véase Tabla 10).

Tabla10
Resultados de la prueba Jarque-Bera

	FACTOR1	FACTOR02	FACTOR03	FACTOR04	FACTOR05	FACTOR06	EMP 2002
Mean	-2.31E-12	-1.81E-12	9.08E-13	1.03E-12	-4.74E-13	8.53E-13	10.70398
Median	-0.031284	-0.104791	0.056657	-0.120248	-0.113794	-0.050867	8.000000
Maximum	4.022478	3.421737	3.510298	4.032051	3.996283	4.347965	245.0000
Minimum	-4.535290	-2.336568	-2.917688	-3.200732	-2.946192	-4.499070	1.000000
Std. Dev.	1.000000	1.000000	1.000000	1.000000	1.000000	1.000000	10.67817
Skewness	-0.142793	0.416481	-0.117654	0.484072	0.485847	0.213561	4.709507
Kurtosis	4.260840	2.562898	2.690508	3.084496	3.070916	2.981031	57.15522
Jarque-Bera	881.4587	466.7019	79.72179	498.1149	500.6331	96.40785	1593593.
Probability	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
Observations	12658	12658	12658	12658	12658	12658	12658

Fuente: Elaboración propia con base en Eviews.

Por su parte, la variable que presentó una mayor relación lineal con respecto a la variable tamany fue definitivamente la variable que hace referencia al número de empleados en 2002, ello debido a que su coeficiente de correlación ascendió a poco más del 70% (véase Tabla 11).



Tabla 11
Matriz de correlaciones

Correlation Matrix								
	FACTOR1	FACTOR02	FACTOR03	FACTOR04	FACTOR05	FACTOR06	EMP_2002	TAMANY
FACTOR1	1.000000	-6.27E-12	-2.14E-12	1.87E-12	-2.44E-12	2.19E-12	-0.038403	-0.024366
FACTOR02	-6.27E-12	1.000000	2.98E-12	-5.31E-12	1.88E-12	8.07E-13	0.146135	0.116763
FACTOR03	-2.14E-12	2.98E-12	1.000000	4.56E-13	3.87E-12	-3.52E-12	0.028377	0.019847
FACTOR04	1.87E-12	-5.31E-12	4.56E-13	1.000000	-3.22E-12	-2.28E-12	0.058828	0.070179
FACTOR05	-2.44E-12	1.88E-12	3.87E-12	-3.22E-12	1.000000	-6.08E-12	0.045620	0.085520
FACTOR06	2.19E-12	8.07E-13	-3.52E-12	-2.28E-12	-6.08E-12	1.000000	0.049876	0.049326
EMP_2002	-0.038403	0.146135	0.028377	0.058828	0.045620	0.049876	1.000000	0.713159
TAMANY	-0.024366	0.116763	0.019847	0.070179	0.085520	0.049326	0.713159	1.000000

Fuente: Elaboración propia con base en Eviews.

Este último resultado pudo corroborarse con la matriz de estructura (véase Tabla 12), en la cual puede apreciarse que la variable que presentó mayor información para la función discriminante 1 fue la variable que hace referencia al número de empleados en 2002, y la cual retuvo el 99.4% de la varianza total explicada.

Tabla 12
Matriz de estructura

	Función	
	1	2
Número empleados 2002	,998*	-,049
REGR factor score 5 for analysis 1	,051	,687*
REGR factor score 4 for analysis 1	,040	,626*
REGR factor score 6 for analysis 1	,034	,272*
REGR factor score 2 for analysis 1	,097	,154*
REGR factor score 1 for analysis 1	-,026	,136*
REGR factor score 3 for analysis 1	,018	-,021*

Correlaciones intra-grupo combinadas entre las variables discriminantes y las funciones discriminantes canónicas tipificadas
Variables ordenadas por el tamaño de la correlación con la función.

*. Mayor correlación absoluta entre cada variable y cualquier función discriminante.

Fuente: Elaboración propia con base en SPSS.

En el AD la variable de agrupación fue la denominada tamany, y para la cual se definieron tres subgrupos; 1 (micro), 2 (pequeña) y 3 (mediana). La tabla 13 muestra la integración de la variable de agrupación.



Tabla 13
Variable Tamany (frecuencias)

	Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos 1	7642	60,4	60,4	60,4
2	4901	38,7	38,7	99,1
3	115	,9	,9	100,0
Total	12658	100,0	100,0	

Fuente: Elaboración propia con resultados de SPSS.

Al llevar a cabo la técnica multivariante de AD a las variables de la tabla 9 se obtuvo que, para la prueba de igualdad de medias, todas las variables resultaron ser significativas. Sin embargo, con la prueba M de Box, que contrasta la igualdad de las matrices de covarianzas, se obtuvo que dichas variables mostraron una clara heterogeneidad entre sus varianzas. Este último resultado pudo verificarse mediante el contraste de la tabla 14.

Tabla 14
Resultados para la prueba de Box

M de Box		12025,123
F	Aprox.	212,305
	gl1	56
	gl2	259671,731
	Sig.	,000

Contrasta la hipótesis nula de que las matrices de covarianzas poblacionales son iguales.

Fuente: Elaboración propia con resultados de SPSS.

Este último resultado implica que existen diferencias sustantivas entre los subgrupos analizados, con respecto a su varianza para explicar el tamaño empresarial, lo cual resulta lógico, debido a que la dispersión de los datos del subgrupo que está integrado por las micro empresas no tiene por qué ser homogéneo con respecto al subgrupo que está integrado por las medianas empresas. Otro resultado importante obtenido con este análisis fue la Lambda de Wilks para las funciones canónicas discriminantes, ya que al obtenerse un nivel de significancia menor al 5% para cada una de ellas, significa que las dos funciones discriminantes (al definirse tres grupos de tamaño para las PyMES, a lo más pueden obtenerse dos funciones lineales que discriminen dichos grupos) son estadísticamente significativas. Ello implica que la discriminación es posible y tiene sentido. Así, y como puede apreciarse en la tabla 15, la función discriminante 1 absorbe el mayor porcentaje de variabilidad, mientras que la función discriminante 2 no presenta mucha asociación entre las puntuaciones discriminantes y los grupos involucrados. Es decir, la discriminación no es tan clara para distinguir una empresa del subgrupo 2 (pequeña) de una del subgrupo 3 (mediana).



Tabla 15
Resultados para los autovalores

Función	Autov valor	% de varianza	% acumulado	Correlación canónica
1	1,596 ^a	99,4	99,4	,784
2	,009 ^a	,6	100,0	,095

a. Se han empleado las 2 primeras funciones discriminantes canónicas en el análisis.

Fuente: Elaboración propia con resultados de SPSS.

Este último resultado puede verificarse con la tabla 16, la cual hace referencia a los puntajes discriminantes para los centroides de los subgrupos:

Tabla 16
Funciones para los centroides

Tamany	Función	
	1	2
1	-,839	-,044
2	1,103	,087
3	8,775	-,747

Funciones discriminantes canónicas no tipificada evaluadas en las medias de los grupos

Fuente: Elaboración propia con resultados de SPSS.

Es claro que para la función 1 se aprecia una mejor discriminación, mientras que para la función discriminante 2 no lo es del todo. No obstante, las dos funciones discriminantes resultaron ser significativas, lo cual puede verificarse con el test de la lambda de Wilks (véase Tabla 17), ya que al obtenerse una significación estadística (*p value*) menor al 5% para cada función discriminante, implica que cada función, de manera individual, es estadísticamente significativa.

Tabla 17
Resultados para la Lambda de Wilks

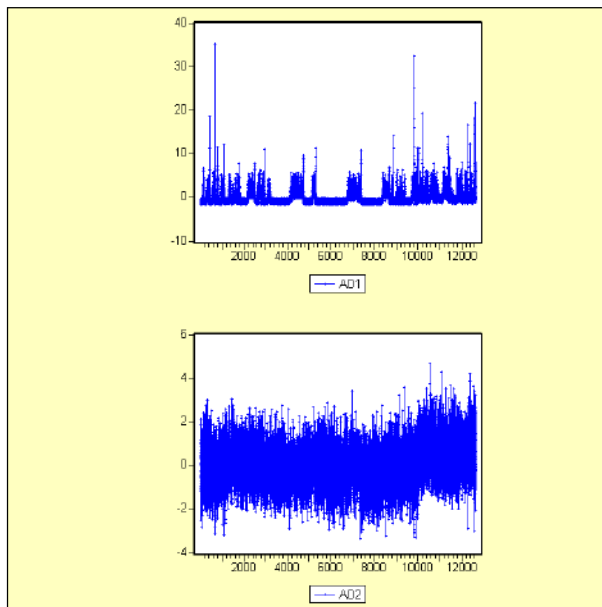
Contraste de las funciones	Lambda de Wilks	Chi-cuadrado	gl	Sig.
1 a la 2	,382	12183,980	14	,000
2	,991	115,488	6	,000

Fuente: Elaboración propia con resultados de SPSS.

En Tabla 18 se muestran los puntajes calculados mediante AD para las dos funciones discriminantes.



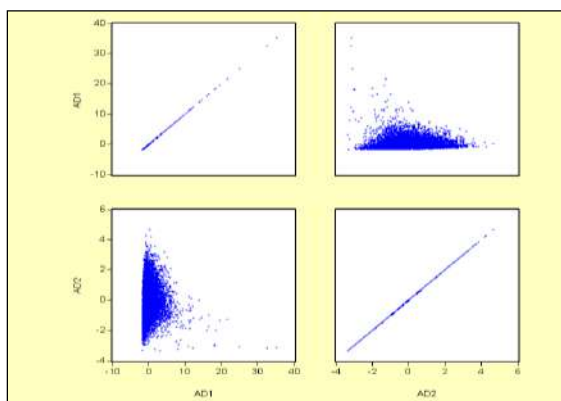
Tabla 18
Puntajes mediante un AD para las funciones discriminantes



Fuente: Elaboración propia con base en Eviews.

En la tabla 19 se muestra el *scatter* de los puntajes calculados mediante AD para las dos funciones discriminantes:

Tabla 19
Puntajes calculados para las funciones discriminantes



Fuente: Elaboración propia con base en Eviews.

Cabe mencionar que con el análisis multivariante AD mediante variación cruzada se obtuvo un 89.9% de clasificación correcta. De la misma manera, en la tabla 20 se muestran los estadísticos de frecuencias para la variable tamany y su correspondiente grupo pronosticado.

Tabla 20
Tamaño empresarial real y pronosticado

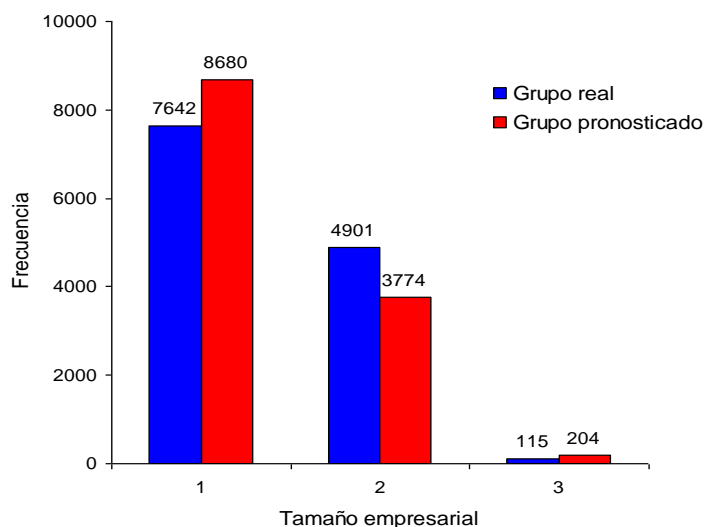
Tamany					
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	1	7642	60.4	60.4	60.4
	2	4901	38.7	38.7	99.1
	3	115	.9	.9	100.0
	Total	12658	100.0	100.0	

Grupo pronosticado para el análisis 1					
		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válidos	1	8680	68.6	68.6	68.6
	2	3774	29.8	29.8	98.4
	3	204	1.6	1.6	100.0
	Total	12658	100.0	100.0	

Fuente: Elaboración propia con resultados de SPSS.

Por último, en la gráfica 3 se presenta la valoración de la clasificación.

Gráfica 3
Clasificación a través de un AD



Fuente: Elaboración propia.

7. Conclusiones

Los resultados obtenidos en este trabajo empírico demuestran que la clasificación de una empresa, a través de su información contable-financiera, es posible. La implicación inmediata de estos resultados bien podría ser una mejor captación de impuestos, debido a que algunas empresas consideradas, mediante otros criterios clásicos, como micro bien pudiesen ser ubicadas en el conglomerado de pequeñas. A su vez, si con esta metodología se detecta que una empresa muestra tener un “mayor” tamaño bien podría considerarse como una



potencial compañía generadora de empleo. Por su parte, y aunque dista mucho de lo obtenido en esta investigación, la importancia de los resultados obtenidos en las últimas décadas para las PyMES han tenido como justificante el crecimiento económico y el potencial innovador de un determinado país o región.

Los ratios seleccionados para el ACP permiten hacer un AD considerable, aún cuando estos *ratios* presentan problemas estadísticos desde su construcción. Para verificar los resultados obtenidos con los *ratios* financieros seleccionados en ACP sería conveniente realizar muestreos independientes (como el muestreo aleatorio sistemático, debido a que ya se conoce el rango de las observaciones) para con ello comprobar la consistencia de dichos resultados.

Existen evidencias estadísticas que permiten la estimación de dos funciones lineales para discriminar a los tres grupos de empresas. Más aún, la discriminación del tamaño de las empresas para 2002 con base en las variables sintéticas generadas mediante un ACP y la variable que hace referencia al número de empleados en 2002, tiene un nivel de confiabilidad de un 89.9% mediante la técnica de validación cruzada. La discriminación realizada permitió identificar considerablemente a una micro empresa de una pequeña empresa. Sin embargo, entre una pequeña empresa y una mediana empresa la discriminación no fue muy clara.

Por su parte, la variable utilizada en el AD que hace referencia al número de empleados en 2002, definitivamente proporciona la mayor información para poder realizar la discriminación. En este sentido, se recomienda reforzar los resultados obtenidos con un análisis contable.

La medida sintética obtenida para explicar el tamaño empresarial mediante los *ratios* seleccionados podría ser sesgada, debido a la naturaleza de los datos principalmente. Lo cual sugiere una transformación de los mismos o, en su caso, la recolección de una base de datos más enriquecida. Por ello, una alternativa sería la selección conveniente de otros *ratios* financieros más significativos para un mismo sector y, con ello, validar los resultados obtenidos.

La búsqueda de un concepto o medida empresarial a partir de datos numéricos por procedimientos estadísticos tradicionales de ninguna manera limitó la finalidad del estudio. Sin embargo, y debido a la naturaleza de los datos, resultaría oportuno la aplicación de otras técnicas para su análisis, como por ejemplo un modelo *Probit (Logit)* con tres categorizaciones, o un análisis de redes neuronales. Este último serviría para fortalecer la relación de los *ratios* o para facilitar la construcción de otros *ratios* sintéticos que proporcionen mayor información respecto al tamaño empresarial.

Referencias

Amat S., O (1994), *Análisis de Estados Financieros. Fundamentos y Aplicaciones*, Gestión 2000, Barcelona.

Aching G., César (2005), *Ratios financieros y matemáticas de la mercadotecnia*, Prociencia y cultura S. A.



- Aluja, T. y Morineau, A. (1999), *Aprender de los datos: El análisis de componentes principales*, EUB, Barcelona.
- Bernstein, L. (2007), *Análisis de Estados Financieros, Teoría, Aplicación e interpretación*, Cuarta Edición, Editorial Colombia S.A.
- Calvo S., A.y Boedo V., L. (2001), *Incidencia do tamaño sobre o comportamento financeiro da empresa. Unha análise empírica con pemes galegas*, *Revista Galega de Economía*, vol. 10 (2): 1-23.
- E. Pla, L. (1986), *Análisis Multivariado: Método de Componentes Principales*, Departamento de Asuntos Científicos, Secretaría General de la Organización de los Estados Americanos, monografía 27, págs. 94.
- Fisher, R. A. (1936). *The use of multiple measurements in taxonomic problems*, *Annals of Eugenics*, 7: 179-188.
- Galindo L., A. (2005), *El tamaño empresarial como factor de diversidad*. Universidad de Málaga. Grupo SEJ-366.
- Lev, Baruch (1974), *Financial Statement Analysis: A New Approach*. Prentice-Hall Inc., EnglewoodCliffs, N. Y.
- Montanero F., J. (2008), *Análisis Multivariante*, Universidad de Extremadura Servicio de Publicaciones, 59.
- Rivero T, P. (1991), *Análisis de Balances y Estados Complementarios*. Edit. Pirámide, Madrid.
- Shih-Wei, L. y Shih-Chieh, C. (2009). *A particle swarm optimization approach for enhancing classification accuracy rate of linear discriminant analysis*, *Applied Soft Computing*, 9: 1008-1015.
- Sosa S., M.C. (2011), *Breve inventario de los principales ratios financieros utilizados en el análisis financiero empresarial*, *EconPapers*, 152: 1-23.
- Urías V., J. (1992), *Análisis de Estados Financieros*, McGraw-Hill, Madrid.
- Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H. et al. (2008). *Top 10 algorithms in data mining*, *KnowlInfSyst*, 14: 1-37.

