

Un procedimiento eficiente para la selección de variables en modelos factoriales de valoración de empresas

Francisco Guijarro Martínez • Ismael Moya Clemente¹
Universidad Politécnica de Valencia

RECIBIDO: 31 de enero de 2007

ACEPTADO: 30 de marzo de 2008

Resumen: En el presente trabajo se propone una metodología alternativa a la empleada en los modelos factoriales clásicos de valoración de empresas. Se muestra cómo la agrupación en factores de las diferentes variables explicativas del valor de la empresa, no debe hacerse en función de la correlación o variabilidad común entre ellas, sino teniendo en cuenta la correlación entre la variabilidad de la capitalización bursátil no explicada por cada una de ellas por separado. A través de la aplicación práctica de esta nueva metodología de selección de variables económico-financieras sobre datos de la Bolsa de Madrid (2000-2003) se constata la obtención de modelos más parsimoniosos, puesto que reducen el número de factores obtenidos respecto a los que aparecen en la literatura en este campo, y con menor multicolinealidad, cuantificada a través del índice de condición; en definitiva, modelos más adecuados para su empleo en valoración de empresas.

Palabras clave: Valoración de empresas / Valor bursátil / Modelos factoriales / Información financiera.

An Efficient Procedure for Variable Selection in Factor Firm Valuation Models

Abstract: This work is devoted to propose an alternative methodology to classic factor models in valuing firms. It is showed that grouping exogenous variables in factor models must be carried out not as a function of their common correlation or variance, but considering the variance of market capitalisation not explained by each exogenous variable. Through the application of this new methodology to financial information from the Bolsa de Madrid (2000-2003), we obtain more parsimonious models, provided that the number of factors is reduced compared to traditional factor models; and multicollineality is also reduced and quantified by means of the condition index. In summary, in our research we propose more appropriate models for valuing firms.

Key Words: Firms valuation / Market capitalisation / Factor models / Financial information.

INTRODUCCIÓN

De acuerdo con la teoría financiera contemporánea el objetivo de la empresa debe ser la maximización de su valor para los accionistas. El valor de mercado de una sociedad anónima puede ser aproximado por el precio de sus acciones —efectuando algunos ajustes como el valor de control—, el cual anticipa las principales decisiones empresariales, como son las de inversión, financiación y reparto de dividendos. Por lo tanto, y asumiendo la hipótesis de que el mercado de capitales funciona con perfección, el valor de las acciones es un parámetro que debe ser considerado por la dirección de la empresa en la asignación eficiente de los recursos. Este objetivo se concreta en la maximización de la diferencia entre el valor de la empresa y el capital invertido por los propietarios, o lo que es lo mismo, en la creación de valor para los accionistas.

La Bolsa es un mercado organizado que proporciona información sobre el valor que asignan los inversores a los resultados que se espera genere un título. Además de esta propiedad princi-

pal de utilidad para los analistas², la información bursátil es una referencia para los directivos de las empresas a la hora de tomar decisiones sobre los recursos que respaldan los valores.

Los modelos econométricos de valoración de empresas permiten determinar qué variables de los estados financieros, e incluso de naturaleza extracontable, influyen significativamente en la explicación de la variable representativa del valor. Para aquellas empresas cuyo capital social no cotiza en el mercado de valores (la gran mayoría) se ha propuesto la utilización de la metodología de valoración analógico-bursátil (Caballer y Moya, 1997) con la finalidad de estimar un valor similar al bursátil y proporcionar esta variable, junto con la información que le afecta significativamente³, que es clave en la dirección financiera actual, donde los directivos deben crear valor para los propietarios y conocer los parámetros que lo determinan y el efecto de sus decisiones sobre ellos. Por lo tanto, la información recogida en los modelos es la considerada por los partícipes de los mercados financieros a la hora de realizar inversiones y debe ser tenida

en cuenta por los directivos de las empresas que busquen la creación de valor para los accionistas.

Los métodos de regresión pretenden estimar el comportamiento de una variable, en este caso el valor bursátil referido a un determinado momento o período de tiempo, estimador del valor de mercado de la empresa, como una función matemática de otras variables consideradas explicativas. Implícitamente, se acepta que el valor bursátil está correlacionado con las características representadas por las variables explicativas. Una adecuada elección de dichas variables, partiendo de la teoría financiera, justifica el método desde la perspectiva teórica. Desde la perspectiva operativa, la precisión de estos métodos depende de la calidad de la estimación de la función que relaciona la variable dependiente con las explicativas, medida por parámetros estadísticos, tales como el coeficiente de determinación y otros tests de significación.

El presente trabajo pretende profundizar en la investigación de los denominados modelos econométricos de valoración de empresas, y más concretamente en aquellos que hacen uso de la técnica estadística de análisis factorial para la selección de variables explicativas del valor bursátil. A la hora de seleccionar modelos de valoración basados en regresión múltiple es habitual tener que confrontar dos criterios contrapuestos, como son el de obtener modelos con alta capacidad explicativa, y el de lograr que estos sean parsimoniosos, esto es, con el menor número posible de variables explicativas. Asimismo, se analizan desde una perspectiva crítica los modelos aportados por diferentes autores, y a la luz de los resultados se propone una metodología que supere los inconvenientes detectados.

Los modelos de valoración factoriales han sido ampliamente utilizados en la literatura financiera (Rushinek y Rushinek 1985, Steele 1995, Corielli y Marcellino, 2005), y más concretamente en el campo de la valoración de empresas, donde su aplicación se vincula en la práctica a los modelos de valoración econométricos. Morton (1977) emplea en su trabajo la técnica del análisis factorial en valoración para agrupar las variables explicativas del precio de la vivienda, según su grado de correlación. Su objetivo es obtener una ecuación de valoración donde las va-

riables explicativas estén poco correlacionadas entre sí, para evitar los problemas que frecuentemente suelen originarse en estos modelos originados por la multicolinealidad, como son la obtención de coeficientes carentes de significado económico en las variables explicativas, coeficientes de determinación que no se corresponden con la significación estadística de las variables consideradas, estimaciones mínimo cuadráticas de los coeficientes con varianzas muy grandes, entre otros. Con los modelos factoriales, además de superar este inconveniente sobre la redundancia en la información, también se persigue obtener modelos parsimoniosos (Caballer y Moya, 1997), en el sentido de que el objetivo de lograr modelos altamente explicativos coexista con un objetivo paralelo, y en la mayoría de las ocasiones en conflicto con el primero, de incluir el menor número de variables explicativas posible. Esta limitación en el número de variables es, en el contexto valorativo, de gran importancia práctica, pues permite al valorador poder realizar su trabajo teniendo que cuantificar un número lo más reducido posible de variables.

Con el análisis factorial, se persigue agrupar las variables según su nivel de correlación, de forma que las agrupadas dentro de un factor estén lo más correlacionadas posibles entre sí, y lo más débilmente correlacionadas con el resto de variables de otros factores. Debido a la diferente escala de medida de las variables explicativas, en el análisis factorial se emplea la matriz de correlaciones frente a la alternativa de utilizar la matriz de covarianzas. Puesto que el objetivo final es obtener una función de valoración, de cada factor se escoge como máximo una variable para dicha función, lo que permite disminuir y en algunos casos eliminar los mencionados problemas de multicolinealidad.

Esta propuesta metodológica fue aplicada en el ámbito español en los trabajos de Caballer y Moya (1997) y Moya (1995, 1996), bajo la denominación de valoración analógico-bursátil, con el objeto de establecer una función de valoración para empresas no cotizadas en bolsa a partir de las empresas sí cotizadas. A diferencia del trabajo de Morton, estos autores proponen considerar como variable representante de cada factor aquella que presente mayor correlación,

en valor absoluto, con la variable endógena. Posteriormente Miralles y Miralles (2002) emplean la misma metodología sobre las empresas cotizadas en el mercado de Lisboa. La principal diferencia respecto de los trabajos de Caballer, Morton y Moya, es que no restringen el número de variables seleccionadas por factor, que en los citados trabajos está limitado a una como máximo. La utilización de una única variable por factor queda justificada por el objetivo de obtener modelos con mínima multicolinealidad, si bien ninguno de estos autores proponen parámetros para su medición. El empleo de modelos de regresión *stepwise*, en sus dos versiones *forward* y *backward*, ya era recogida en el trabajo de Morton. En relación a los resultados, una diferencia entre los trabajos de Caballer y Moya y el de Miralles y Miralles está en el número de factores obtenidos: cinco en el caso español, frente a un máximo de nueve en el caso portugués.

En todos estos estudios se justifica la utilización del análisis factorial como una herramienta idónea para la reducción/eliminación de multicolinealidad en las funciones de valoración. Sin embargo, en ninguno de ellos se cuantifica el grado de multicolinealidad de las funciones resultantes. Tampoco se justifica si el procedimiento seguido para la obtención de las funciones de valoración puede considerarse óptimo, ya que la elección de la variable representativa de cada factor queda a elección del analista, sin que exista un procedimiento objetivo de selección⁴.

El objetivo de nuestro trabajo es doble: por una parte, presentamos un procedimiento que permite obtener de forma eficiente y automática las funciones de valoración; por otro lado, proponemos el índice de condición para validar el grado de multicolinealidad de las funciones obtenidas.

El resto del trabajo se estructura del siguiente modo: en el epígrafe 2 se presenta de forma somera la técnica estadística de análisis factorial, así como las principales debilidades de su aplicación clásica en los modelos factoriales de valoración de empresas. El epígrafe concluye con la propuesta de un procedimiento que supera estos inconvenientes. En el epígrafe 3 se ilustra la nueva propuesta sobre una base de datos real, y se comparan los resultados con los obtenidos a

través de otros procedimientos reflejados en la literatura. Finalmente, el trabajo termina con un epígrafe de conclusiones.

METODOLOGÍA

El modelo de análisis factorial supone una descomposición de las p variables originales observadas sobre una muestra de n individuos en la forma (1):

$$X = 1\mu' + F\Lambda' + \Psi \quad (1)$$

donde X es la matriz de datos original de dimensión $n \times p$ con distribución normal de media μ y matriz de covarianzas V , 1 es un vector de unos con dimensión $n \times 1$, μ es el vector de medias de las p variables, F es la matriz que contiene los m factores de los n individuos de la muestra, por tanto matriz de dimensión $n \times m$, Λ es la matriz de carga $m \times p$, y Ψ es la matriz $n \times p$ que recoge las perturbaciones, que se suponen normales con media 0 y matriz de covarianzas diagonal. A partir de (1) se comprueba que la matriz de carga Λ expresa la covarianza entre los factores latentes y las variables. Respecto de estas últimas, su matriz de covarianzas puede descomponerse siguiendo (2):

$$V = E[(x - \mu)(x - \mu)'] = \Lambda E[ff']\Lambda' + E[uu'] \quad (2)$$

y teniendo en cuenta que los factores están incorrelados, la esperanza de su covarianza ff' será la matriz identidad, con lo que la covarianza de las variables quedará descompuesta en dos partes independientes (3):

$$V = \Lambda\Lambda' + \Psi \quad (3)$$

La covarianza de la matriz de carga $\Lambda\Lambda'$ representa la variabilidad común entre las variables originales, en terminología factorial comunalidad, que depende de las covarianzas entre las variables y los factores. La segunda parte del sumando, Ψ , recoge la parte de variabilidad específica de cada variable, y por tanto independiente del resto (matriz diagonal). De esta forma se observa una clara analogía entre el análisis

factorial y la clásica descomposición de la varianza (ANOVA).

DEBILIDADES DE LOS MODELOS FACTORIALES ACTUALES DE VALORACIÓN DE EMPRESAS

Supongamos que se pretende explicar el valor bursátil y mediante un conjunto de p variables provenientes de la información económica-financiera publica generada por las empresas, de las que seleccionamos para su análisis en detalle las variables x_i y x_j como explicativas:

- y : vector que recoge el valor bursátil de las n empresas analizadas, con dimensión $n \times 1$.
- X : matriz que recoge las p variables económico-financieras de las n empresas, con dimensión $n \times p$. La variable x_i ocupará la columna i -ésima de dicha matriz.

La probabilidad de que, tras aplicar el análisis factorial en el formato clásico referenciado en la literatura sobre el conjunto de variables explicativas, las variables x_i y x_j se agrupen en un mismo factor es proporcional al coeficiente de correlación entre ambas. La probabilidad será tanto mayor cuanto más elevado sea este coeficiente, en valor absoluto. En cualquier caso, debe señalarse que en la agrupación no ha intervenido la variable endógena, ni su grado de relación con las variables explicativas. Podemos distinguir dos situaciones según el grado de relación entre las variables x_i y x_j : una correlación lo suficientemente significativa como para que se agrupen en un mismo factor; una correlación no significativa, con lo que se agruparán en factores diferentes. Analizando cada una de estas situaciones por separado:

♦ *Situación 1*: correlación altamente significativa entre las variables x_i y x_j .

Es la representada en las figuras 1.1 y 1.2, donde asumimos que ambas variables se agrupan en un mismo factor dado que la variabilidad común –intersección de la variabilidad de ambas– es muy elevada. La diferencia entre las figuras está en la capacidad explicativa conjunta de x_i y x_j sobre y .

En la figura 1.1 la variabilidad de la parte no explicada de la variable y (tono claro) es compa-

rativamente muy superior a la parte explicada simultáneamente por ambas variables (tono oscuro) o por cada una de ellas por separado (tono intermedio). Esto es, aún compartiendo factor, no es previsible que vayan a formar parte de la función de valoración. Obsérvese que una forma alternativa de analizar esta situación es a través de la parte no explicada por cada una de las variables x_i y x_j : puesto que ambas variables no están correlacionadas con y , la parte no explicada por x_i está muy correlacionada con la parte no explicada por x_j . En resumen, se puede encontrar un factor que aglutine a variables nula o escasamente correlacionadas con y , por lo que el factor, junto con sus variables, puede ser descartado del proceso de valoración. Tanto el análisis factorial clásico como el realizado sobre la parte no explicada llevaría a los mismos resultados.

En la figura 1.2 las variables x_i y x_j siguen estando igualmente correlacionadas, por lo que comparten factor, pero a diferencia de la figura 1.1, ambas explican una parte significativa de la variabilidad de y . Siguiendo el procedimiento de los modelos factoriales actuales, sólo una de las variables debería incluirse en el modelo de valoración (la más correlacionada con el valor bursátil y). Sin embargo, si analizamos la variabilidad explicada por cada variable, observamos que son considerablemente distintas. Esto es, que aún estando correlacionadas entre sí, cada una de ellas explica una parte sustancialmente diferente de la variabilidad de y : tienen *capacidad explicativa incremental*, por lo que presumiblemente ambas podrían formar parte de la función de valoración. Esta situación no es captada por los modelos factoriales en la literatura. Alternativamente, el análisis de la relación entre la parte no explicada por cada variable sí pone en evidencia esa capacidad explicativa incremental. El hecho de que la correlación entre la parte no explicada por x_i y la parte no explicada por x_j sea baja, indicaría que ambas variables deberían pertenecer a factores diferentes, y también ambas podrían formar parte de la función de valoración. En definitiva, el análisis factorial en su forma clásica –sobre las variables originales–, agruparía a ambas variables en un único factor, impidiendo que pudieran formar parte simultáneamente de la función de valoración. Un análisis factorial sobre la parte no explicada por cada variable, las

agruparía en distintos factores, permitiendo que aparecieran en una misma función de valoración.

♦ *Situación 2:* baja correlación significativa entre las variables x_i y x_j .

En el caso de las figura 1.3 y 1.4, la variabilidad común entre las variables x_i y x_j es escasa, con lo que presumiblemente se agruparían en distintos factores. Siguiendo el esquema de los modelos factoriales actuales, ambas variables podrían formar parte de la función de valoración. Analizando en detalle la figura 1.3, observamos que consiguen explicar casi perfectamente la variabilidad de y , y que además cada una de ellas por sí sola explica una parte significativamente distinta de la explicada por la otra. De nuevo, ambas tienen capacidad explicativa incremental, corroborando su posible inclusión en la función de valoración. A esta misma conclusión se llega si se analiza la variabilidad no explicada por cada una de ellas. La correlación entre ambas partes, que en la figura viene dada por la intersección de los tonos claros, es lo suficientemente pequeña como para permitir que ambas variables se incluyan en la función de valoración. En este caso, el análisis factorial clásico y el efectuado sobre la parte no explicada llegarían a la misma conclusión: ambas variables en distintos factores, pudiendo aparecer en una misma función de valoración.

En la figura 1.4 las variables x_i y x_j siguen estando débilmente correlacionadas, de igual forma en que lo estaban en la figura 1.3, por lo que volverían a formar parte de factores distintos. Sin embargo, es fácil constatar a través de la figura la escasa capacidad explicativa incremental de cada una de las variables, con lo que en realidad resulta plausible que sólo una de ellas formara parte de la función de valoración. Alternativamente, si analizamos la variabilidad de y no explicada por cada una de ellas, la correlación entre las mismas –de nuevo las zonas en tono claro– es mucho mayor que la detectada en la figura 1.3, con lo que, si hubiéramos realizado un análisis factorial tomando por cada variable la parte que no explica de y , en el resultado muy probablemente x_i y x_j hubieran compartido factor. En resumen, el análisis factorial sobre las variables originales las asigna-

ría a distintos factores, mientras que el aplicado sobre la parte no explicada las situaría en un único factor. Puesto que ninguna de ellas es relevante desde un punto de vista valorativo, el segundo análisis es más eficiente que el primero, pues las agrupa en un factor de “variables no relevantes”, en lugar de asignarlas a distintos factores que luego deben analizarse por separado.

En definitiva, y tras haber analizado diferentes posibilidades acerca de la correlación entre las variables exógenas, y su capacidad explicativa sobre y , podemos concluir que la aplicación de los modelos factoriales sobre la matriz de correlaciones de la variables explicativas no siempre ofrecerá un resultado satisfactorio. En unos casos porque agrupará en un mismo factor variables que podrían formar conjuntamente parte de la función de valoración; en otros porque agrupará en distintos factores variables que no podrán formar parte conjuntamente de la función de valoración.

Como criterio alternativo al clásico, podemos afirmar que la agrupación en factores de las diferentes variables explicativas no debe hacerse en función de la correlación o variabilidad común entre ellas, sino teniendo en cuenta la correlación entre la variabilidad de y no explicada por cada una de ellas por separado (tono claro en las figuras).

Figura 1.1.- Variables económico-financieras x_i y x_j , significativamente correlacionadas, con baja capacidad explicativa del valor bursátil y

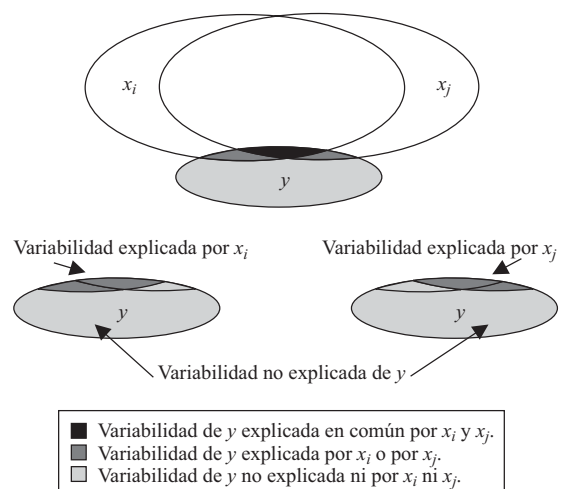


Figura 1.2.- Variables económico-financieras x_i y x_j , significativamente correlacionadas, con alta capacidad explicativa del valor bursátil y

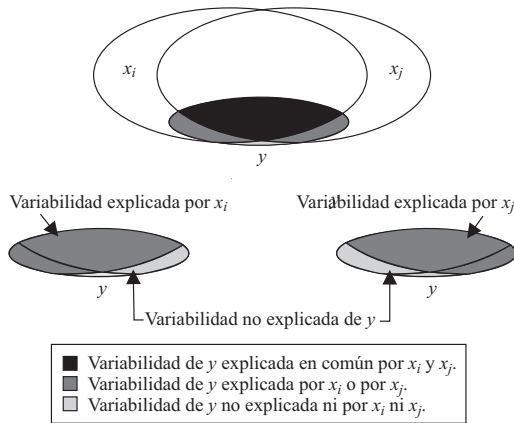


Figura 1.3.- Variables económico-financieras x_i y x_j , poco correlacionadas, con alta capacidad explicativa del valor bursátil y

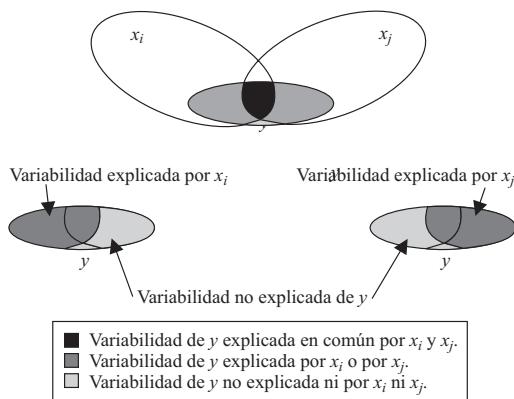
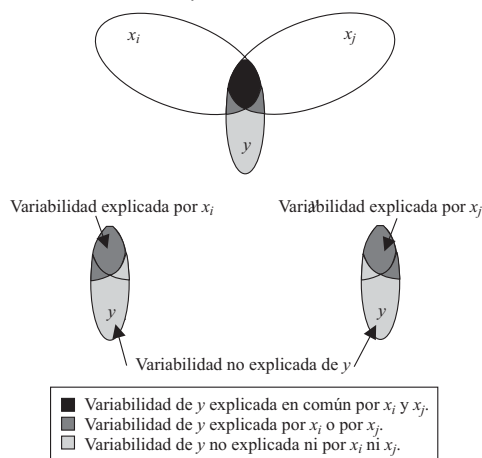


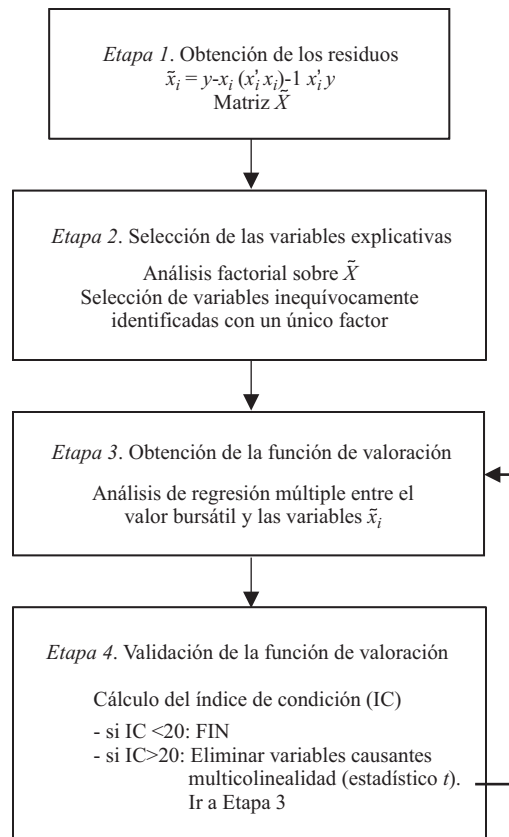
Figura 1.4.- Variables económico-financieras x_i y x_j , poco correlacionadas, con menor capacidad explicativa del valor bursátil y



UNA NUEVA PROPUESTA METODOLÓGICA

La propuesta metodológica para la aplicación de los modelos factoriales en valoración de empresas que se presenta en este trabajo ha sido estructurada de forma detallada en cuatro etapas consecutivas, lo que permite su aplicación de forma automática, además de reducir la discrecionalidad del analista a la hora de seleccionar las variables explicativas que han de incluirse en el modelo final de regresión. Con ello se superan los problemas puestos de manifiesto anteriormente en los modelos factoriales de valoración. Un esquema que resume lo realizado en cada etapa se muestra en la figura 2.

Figura 2.- Etapas para la obtención de la función de valoración



♦ *Etapa 1:* Obtención de los residuos por análisis de regresión simple. La propuesta metodológica que se realiza en este trabajo hace uso del análisis clásico de descomposición de la varianza. Dada la matriz de variables explicativas

X y el vector que recoge la variable endógena y , podemos realizar la siguiente descomposición (4):

$$y - \bar{y} = (y - \hat{y}_i) + (\hat{y}_i - \bar{y}) \quad (4)$$

donde el vector \bar{y} de dimensión $n \times 1$ contiene en todos sus elementos la media de la variable endógena y , mientras que el vector \hat{y}_i recoge la estimación obtenida por regresión lineal al emplear como única variable explicativa la i -ésima componente de la matriz X . Esto es, $\hat{y}_i = x_i(x_i' x_i)^{-1} x_i' y$.

Utilizando la terminología propia del análisis de la varianza, la componente $(\hat{y}_i - \bar{y})$ es conocida como la variabilidad de y explicada por la variable x_i , mientras que $(y - \hat{y}_i)$ recoge la parte no explicada o residual. De forma simétrica a como se construye la matriz de variables explicativas X , podemos obtener una matriz de idéntica dimensión \tilde{X} pero donde por columnas se recoja la componente no explicada por cada variable x_i (5). En definitiva, cada componente original x_i es sustituida por el residuo obtenido en la regresión entre la variable endógena y , y la exógena x_i .

$$\tilde{x}_i = y - x_i(x_i' x_i)^{-1} x_i' y \quad (5)$$

Así pues, tanto más próxima será la varianza de \tilde{x}_i a cero conforme mayor sea el nivel de correlación entre las variables x_i e y . Es previsible que cuando ambas variables estén poco correlacionadas, la estimación del vector $x_i(x_i' x_i)^{-1} x_i' y$ será muy próxima a \bar{y} , y por lo tanto la varianza de \tilde{x}_i tienda a la varianza de y .

En este trabajo, se propone que sea la matriz \tilde{X} la empleada en el análisis factorial, y no la original X habitualmente utilizada en la literatura, de forma que el análisis agrupará en distintos factores aquellas variables que expliquen partes significativamente diferentes de la variable endógena y . Esto supone una diferencia significativa respecto de los trabajos de investigación citados, donde el valor bursátil no era considerado

en la etapa del análisis factorial, sino únicamente en la última fase de determinación de la función de valoración. En esta propuesta, la variable y se emplea para la obtención de cada una de las columnas de la matriz \tilde{X} , por lo que forma parte del análisis desde el primer momento.

Así mismo, es recomendable emplear la matriz de covarianzas de \tilde{X} , en lugar de la habitual de correlaciones. Por la forma en que se ha construido la matriz \tilde{X} , todas sus columnas \tilde{x}_i vienen expresadas en la misma unidad económica⁵, la de la variable endógena y . La diferente varianza de cada \tilde{x}_i sí es relevante, puesto que es inversamente proporcional a la capacidad explicativa de la variable sobre la variable endógena.

Una vez obtenidos los factores a partir de la matriz \tilde{X} , y siguiendo la propuesta original de Caballer y Moya (1997), deben seleccionarse las variables que se incorporarán en el modelo de regresión múltiple para obtener la función de valoración. Estos autores recomiendan seleccionar una única variable por factor, para minimizar los posibles problemas de multicolinealidad. Aunque no es un procedimiento óptimo, sugieren escoger para cada factor la variable más correlacionada con el valor bursátil, sin que ello garantice obtener el modelo con mayor capacidad explicativa y libre de multicolinealidad. Nuestra propuesta es seleccionar las variables explicativas no sólo teniendo en cuenta su capacidad explicativa global, sino su capacidad explicativa incremental, de forma que se reduzca el solapamiento de información y, con ello, se minimice el grado de multicolinealidad de los modelos resultantes. Así pues, las funciones de valoración obtenidas tendrán que mostrar un equilibrio entre: 1) la capacidad explicativa, cuantificada a través del coeficiente de determinación corregido; 2) el grado de multicolinealidad, medido a través del índice de condición. Para evitar la subjetividad que supone que sea el analista quien seleccione las variables representativas de cada factor, proponemos obtener la función de valoración en tres pasos consecutivos.

♦ *Etapa 2:* Selección de variables explicativas. Una vez agrupadas las variables por factores, en la función de valoración sólo puede considerarse una variable por factor, para limitar el grado de multicolinealidad. Dicha variable re-

representativa del factor, debe estar altamente correlacionada con el mismo, y poco correlacionada con el resto de factores. Por lo tanto, se escogerá como variable representativa aquella con mayor coeficiente, en valor absoluto, dentro del factor⁶. De esta forma la elección de la variable representativa se lleva a cabo de forma totalmente automática y objetiva.

♦ *Etapa 3:* Obtención de la función de valoración. A partir de las variables económico-financieras seleccionadas en la etapa 2, se realizará un análisis de regresión múltiple, tomando como variable dependiente el valor bursátil. Empleando el método *stepwise regression* aseguraremos la obtención del modelo con mayor capacidad explicativa y en el que todas las variables explicativas sean estadísticamente significativas.

♦ *Etapa 4:* Validación de la función de valoración. El procedimiento *stepwise regression* no garantiza la obtención de funciones de valoración libres de multicolinealidad, si bien la acotación en el número de variables explicativas de la etapa 2 reduce considerablemente la posibilidad de encontrar este problema. La presencia de multicolinealidad está prácticamente asegurada en el caso de que en la función aparezcan como variables explicativas del valor bursátil variables identificadas con un mismo factor. Para evitarla, es necesario calcular el índice de condición (IC) a partir de la matriz simétrica $\tilde{X}'\tilde{X}$, siendo \tilde{X} la matriz de que recoge las columnas de las variables económico-financieras seleccionadas en la etapa 2, y no el conjunto completo \tilde{X} . El índice de condición se calcula con la expresión (6) a partir de los valores propios máximo (λ_{max}) y mínimo (λ_{min}) de matriz $\tilde{X}'\tilde{X}$.

$$IC(\tilde{X}'\tilde{X}) = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}} \quad (6)$$

Habitualmente se considera que valores del índice de condición superiores a 20 ponen de manifiesto problemas graves de multicolinealidad, mientras que valores superiores a 30 indicarían un nivel de multicolinealidad inaceptable.

Así, en el caso de encontrar valores superiores a 20, deberá eliminarse del análisis de regresión la variable económico-financiera con menor estadístico t , y repetir la etapa 4 hasta obtener

una función de valoración con un nivel de multicolinealidad aceptable⁷.

BASE DE DATOS Y RESULTADOS

Para contrastar empíricamente la metodología propuesta y compararla con los resultados obtenidos mediante otras aproximaciones, se ha recopilado información económico-financiera y bursátil de 104 empresas no financieras cotizadas en la Bolsa de Madrid, para cada uno de los años del periodo 2000-2003. El número de observaciones se mantiene constante para cada uno de los años considerados, habiendo eliminado de la base de datos aquellas empresas no disponibles para el periodo completo. Con ello se supera el efecto de supervivencia. Las variables analizadas junto con los principales descriptivos aparecen en la tabla 1, de las que se han excluido los ratios más habituales por presentar correlación próxima a cero y no significativa estadísticamente con la variable capitalización bursátil.

Tabla 1.- Información de empresas no financieras cotizadas en la Bolsa de Madrid (miles de euros)

	MEDIA	DESV. TÍPICA
Capitalización bursátil	2.225.510	8.081.536
Cuentas de Balance		
INMOVILIZADO	2.587.518	8.787.580
Inmovilizado inmaterial	264.968	1.535.959
Inmovilizado material	1.676.594	5.419.430
Activo circulante	989.039	2.242.528
Existencias	185.645	363.271
Deudores	518.138	1.174.556
Tesorería	230.022	714.085
Total activo	3.576.557	10.799.738
Fondos propios	1.191.686	3.816.659
Capital suscrito	198.472	612.565
Pasivo fijo	1.087.934	3.583.090
Acreedores largo plazo	1.077.711	2.812.808
Pasivo líquido	413.171	1.278.751
Deudas financieras	373.383	826.855
Acreedores comerciales	2.587.518	8.787.580
CUENTAS DE RESULTADO		
Resultado bruto	266.555	847.364
Resultado explotación	-69.380	272.269
Resultado financiero	168.631	591.805
Resultado actividades ordinarias	-85.527	1.124.281
Resultados activid. extraordinarias	83.334	834.905
Resultado del ejercicio	1.179.414	3.735.177
Gastos de mercaderías y mat. primas	236.393	596.395
Gastos de personal	181.199	786.051
Gastos por dotaciones para amortiz.	88.374	311.387
Gastos financieros y asimilados	263.064	1.144.157
Cash flow	2.106.081	5.683.408
Estadísticos obtenidos como promedio anual para el periodo analizado.		

En la aplicación de modelos econométricos en el ámbito de la valoración de empresas, es habitual tratar con observaciones muy heterogéneas en tamaño, por lo que para evitar problemas de heterocedasticidad⁸, sobre las variables originales se ha aplicado la transformación logarítmica (7):

$$x'_i = (x_i > 0) \times \ln(x_i) - (x_i < 0) \times \ln(-x_i) \quad (7)$$

donde $(x_i > 0)$ y $(x_i < 0)$ son pruebas lógicas que permiten calcular el logaritmo de la variable x_i en el caso de tomar valores negativos.

Con esta información, y después de haber aplicado la transformación (7), se han llevado a cabo tantas regresiones como variables explicativas de la capitalización bursátil, obteniéndose así las estimaciones mínimo-cuadráticas de la variable endógena, \hat{y}_i . Seguidamente se ha realizado el análisis factorial tomando los residuos de las anteriores regresiones, $(y - \hat{y}_i)$, empleando la matriz de covarianzas en lugar de la matriz de correlaciones, y practicando una rotación Varimax sobre los factores iniciales. Los resultados han sido agrupados por años en la tabla 2, donde también se recoge la varianza explicada por cada factor y el índice KMO⁹. A efectos de comparación, se han incluido en la tabla 3 los factores obtenidos para cada uno de los años analizados siguiendo la metodología planteada por Caballer y Moya (1997). En ambos casos, las variables han sido ordenadas respetando el orden obtenido para el año 2003.

Respecto del análisis practicado sobre la matriz \tilde{X} (tabla 2), podemos observar cómo, a excepción del año 2000 donde se obtienen 3 factores, en el resto de años las variables se agrupan en dos únicos factores. Con estos resultados queda garantizado el objetivo de conseguir modelos parsimoniosos, puesto que el número de factores es una cota superior del número de variables explicativas en la función de valoración (Caballer y Moya, 1997). Para el periodo 2001-2003, la composición de los factores se mantiene prácticamente invariable, lo que corrobora la robustez del análisis. Por ejemplo, en el caso del año 2003, el primer factor agrupa básicamente a variables de Balance, mientras que en el segundo

factor se encuentran la mayoría de cuentas relacionadas con el resultado en sus distintas dimensiones. Estos resultados contrastan en parte con los modelos de valoración propuestos por Ohlson (1995), donde el valor bursátil es función del valor en libros (cuenta de Balance) y los resultados no esperados (cuenta de Resultados), dejando la parte no explicada en función de otra información de carácter no financiera y con relevancia valorativa¹⁰.

Finalmente, debe destacarse la adecuación de los resultados obtenidos, tal y como confirma el hecho de que el índice KMO se sitúe por encima de 0,9 en todos los casos.

Al aplicar el análisis factorial sobre la matriz X (tabla 3), los valores obtenidos para el índice KMO se sitúan, a excepción del año 2003, por debajo de los registrados en la tabla 2. Además, el número de factores obtenidos para cada uno de los años es en todos los casos mayor que los alcanzados por la metodología propuesta.

Puesto que el objetivo final es obtener una función de valoración, en la tabla 4 se presentan los diferentes modelos obtenidos para cada uno de los años analizados siguiendo la metodología propuesta (análisis factorial sobre la matriz \tilde{X}) y la referida en la literatura (análisis factorial sobre la matriz X). Para poder comparar con metodologías totalmente automáticas, también se incluyen en la misma tabla los resultados obtenidos al aplicar la regresión *stepwise*. El principal inconveniente de este tipo de metodologías automáticas es que emplean un sólo criterio para la inclusión de variables en las funciones de valoración, el de la significación estadística. Como consecuencia, es muy probable que se consideren variables altamente correlacionadas entre sí, originando graves problemas de multicolinealidad, convirtiéndose en un inconveniente de cara a la aplicación de los modelos resultantes fuera de la muestra y a la aceptación de los mismos por parte del analista.

Con el objeto de poder cuantificar el grado de multicolinealidad de las diferentes funciones de valoración, en la última columna de la tabla 4 se recoge el índice de condición.

Al aplicar los modelos automáticos de regresión *stepwise*, se observa como, aún obteniendo

Tabla 2.- Resultados del análisis factorial sobre la matriz \tilde{X} . Coeficientes de los factores rotados

VARIABLES	AÑOS								
	2003		2002		2001		2000		
	Factor 1	Factor 2	Factor 1	Factor 2	Factor 1	Factor 2	Factor 1	Factor 2	Factor 3
Total activo	0,917	0,146	0,925	0,122	0,905	0,143	0,061	0,697	0,592
Acreedores L.P.	0,875	0,300	0,824	0,351	0,799	0,379	0,356	0,821	0,238
Gastos financieros y asim.	0,858	0,325	0,847	0,380	0,909	0,239	0,331	0,818	0,284
Inmovilizado	0,854	0,412	0,816	0,408	0,849	0,393	0,333	0,725	0,490
Pasivo fijo	0,841	0,429	0,795	0,486	0,819	0,357	0,373	0,752	0,419
Pasivo líquido	0,814	0,476	0,827	0,475	0,810	0,515	0,582	0,535	0,562
Activo circulante	0,806	0,419	0,825	0,417	0,820	0,411	0,397	0,517	0,632
Inmov. material	0,797	0,439	0,748	0,499	0,763	0,491	0,408	0,744	0,306
Capital suscrito	0,765	0,250	0,733	0,295	0,805	0,271	0,158	0,765	0,431
Gastos amort. inmov.	0,762	0,453	0,750	0,465	0,766	0,479	0,479	0,592	0,519
Deudas financieras	0,754	0,417	0,728	0,485	0,813	0,412	0,432	0,827	0,163
Deudores	0,737	0,556	0,768	0,550	0,756	0,580	0,536	0,492	0,632
Acreedores comerciales	0,726	0,536	0,762	0,525	0,764	0,485	0,591	0,465	0,550
Gastos mercaderías	0,687	0,489	0,751	0,542	0,638	0,509	0,444	0,536	0,519
Resultado bruto	0,673	0,622	0,639	0,735	0,644	0,736	0,62	0,422	0,620
Existencias	0,652	0,606	0,670	0,592	0,677	0,553	0,641	0,542	0,280
Inmovi. inmat.	0,614	0,518	0,629	0,518	0,673	0,457	0,339	0,309	0,698
Resultado actividades ord.	0,404	0,933	0,392	0,882	0,352	0,904	0,881	0,259	0,295
Resultado explotación	0,660	0,911	0,391	0,887	0,293	0,937	0,887	0,310	0,327
Resultado ejercicio	0,299	0,902	0,380	0,911	0,299	0,929	0,896	0,246	0,310
Cash flow	0,339	0,871	0,392	0,899	0,268	0,904	0,854	0,171	0,220
Resultado extraordinario	0,383	0,853	0,403	0,887	0,490	0,816	0,786	0,358	0,389
Fondos propios	0,406	0,748	0,274	0,902	0,630	0,270	0,031	0,515	0,517
Resultado financiero	0,539	0,739	0,514	0,756	0,607	0,628	0,693	0,627	0,117
Gastos personal	0,576	0,660	0,630	0,651	0,678	0,598	0,568	0,312	0,695
Número empleados	0,550	0,649	0,624	0,639	0,662	0,604	0,548	0,285	0,676
Tesorería	0,561	0,579	0,578	0,466	0,637	0,415	0,396	0,270	0,792
Varianza explicada por el factor	45,1%	36,1%	45,6%	38,2%	48,3%	33,1%	30,8%	30,5%	23,9%
KMO	0,916		0,925		0,913		0,915		

NOTA: Los coeficientes en negrita identifican las variables representantes de cada factor.

Tabla 3.- Resultados del análisis factorial sobre la matriz original X. Coeficientes de los factores rotados

VARIABLES	AÑOS													
	2003			2002			2001			2000				
	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	Factor 4
Acreedores LP	0,872	0,280	0,271	0,755	0,539	0,068	0,029	0,874	0,116	-0,132	0,569	0,724	0,179	0,037
Total activo	0,830	0,042	0,470	0,854	0,426	0,213	0,108	0,934	0,307	0,101	0,789	0,471	0,327	0,089
Inmovilizado	0,822	0,361	0,366	0,806	0,478	0,169	0,123	0,925	0,296	0,049	0,764	0,526	0,285	0,016
Pasivo fijo	0,819	0,394	0,327	0,787	0,530	0,060	0,072	0,902	0,188	-0,011	0,716	0,588	0,253	0,012
Gastos financieros y asim.	0,811	0,290	0,339	0,765	0,593	0,073	0,056	0,970	0,091	-0,001	0,661	0,659	0,294	-0,001
Capital suscrito	0,804	0,250	0,164	0,679	0,559	0,083	0,055	0,867	0,081	-0,034	0,709	0,573	0,075	-0,011
Inmov. material	0,802	0,408	0,289	0,724	0,529	0,240	0,146	0,887	0,342	-0,091	0,652	0,596	0,361	-0,026
Deudas financieras	0,726	0,417	0,258	0,596	0,695	0,124	0,124	0,916	-0,004	-0,100	0,439	0,800	0,192	0,165
Pasivo líquido	0,640	0,345	0,627	0,858	0,388	0,227	0,046	0,919	0,259	0,157	0,781	0,420	0,339	0,208
Gastos amort. inmov.	0,633	0,370	0,509	0,871	0,322	0,125	-0,037	0,888	0,246	0,182	0,782	0,462	0,321	-0,003
Gastos mercaderías	0,556	0,427	0,453	0,813	0,313	0,237	-0,104	0,716	0,346	0,058	0,705	0,435	0,308	0,312
Resultado actividades ord.	0,205	0,894	0,266	0,269	0,091	0,797	-0,407	0,215	0,863	0,116	0,242	0,136	0,815	0,169
Resultado explotación	0,294	0,855	0,356	0,218	0,188	0,826	-0,356	0,209	0,912	-0,007	0,285	0,176	0,929	-0,023
Resultado ejercicio	0,273	0,854	0,305	0,151	-0,145	0,846	0,281	0,128	0,939	-0,064	0,275	0,101	0,938	-0,082
Resultado extraordinario	0,315	0,818	0,337	-0,058	0,133	0,062	0,872	0,076	0,197	-0,769	-0,144	-0,017	0,080	-0,907
Cash flow	0,296	0,817	0,340	0,217	-0,061	0,780	0,306	0,247	0,865	-0,120	0,274	0,148	0,923	-0,088
Resultado financiero	0,467	0,713	0,346	-0,209	-0,799	0,050	-0,006	-0,640	0,208	0,240	-0,094	-0,890	-0,025	0,030
Fondos propios	0,358	0,711	0,310	0,133	0,170	0,791	0,034	0,819	0,364	0,080	0,768	0,415	0,298	0,043
Existencias	0,522	0,536	0,471	0,615	0,412	0,252	0,023	0,807	0,221	-0,096	0,467	0,558	0,389	0,344
Gastos personal	0,299	0,502	0,765	0,938	0,007	0,162	-0,138	0,778	0,277	0,465	0,892	0,165	0,247	0,067
Número empleados	0,268	0,495	0,758	0,909	-0,021	0,125	-0,200	0,746	0,248	0,489	0,858	0,155	0,201	0,097
Deudores	0,522	0,404	0,699	0,901	0,258	0,223	0,034	0,878	0,298	0,238	0,812	0,340	0,325	0,192
Resultado bruto	0,457	0,476	0,698	0,919	0,199	0,262	-0,114	0,857	0,355	0,295	0,848	0,319	0,351	0,073
Acreedores comerciales	0,506	0,394	0,681	0,864	0,287	0,224	-0,030	0,876	0,299	0,161	0,741	0,370	0,313	0,308
Inmov. inmaterial	0,380	0,397	0,648	0,834	0,167	0,151	-0,177	0,775	0,189	0,402	0,860	0,216	0,134	-0,006
Activo circulante	0,622	0,279	0,638	0,862	0,340	0,255	0,078	0,899	0,311	0,145	0,779	0,381	0,367	0,214
Tesorería	0,362	0,457	0,600	0,840	0,088	0,184	0,100	0,731	0,332	0,371	0,876	0,079	0,218	0,122
Var. explicada por el factor	33,2%	28,4%	23,8%	50,1%	14,9%	14,7%	5,3%	59,5%	17,6%	6,5%	44,6%	21,0%	18,5%	5,1%
KMO	0,916			0,901			0,902			0,895				

NOTA: Los coeficientes en negrita identifican las variables representantes de cada factor.

coeficientes de determinación cercanos al 90%, el grado de multicolinealidad es inaceptable. El valor mínimo del índice de condición se obtiene en el año 2003, con un valor de 107,6, muy por encima del 20 admitido como cota superior. Así mismo, las funciones de valoración incluyen un número excesivo de variables como para que puedan ser consideradas funciones parsimoniosas. Es el año 2001 donde este problema se observa con mayor claridad, al considerar un total de 9 variables en la explicación del comportamiento del valor bursátil. En el conjunto de dichas variables, la mayor correlación se da entre las variables Total activo y Gastos financieros y asimilados, que presentan un coeficiente de correlación de 0,923.

Los modelos factoriales, al considerar como máximo una variable por factor en la función de valoración, limitan la posibilidad de considerar variables tan altamente correlacionadas entre sí. En la tabla 4 aparecen las funciones de valoración al considerar como representante del factor a la variable con mayor coeficiente en el mismo. Con ello se acota el grado de multicolinealidad, que en el caso de aplicar los modelos de valoración clásicos basados en el análisis factorial sobre X (Caballer y Moya, 1997), obtienen índices de condición en todos los casos inferiores a 20. El promedio de los cuatro años considerados es de 11,9. Al considerar un menor número de variables explicativas, es lógico que la capacidad explicativa de las mismas se vea mermada respecto de los modelos automáticos. En promedio, el coeficiente de determinación corregido es de 69,5%, frente al 87,5% obtenido con la regresión *stepwise*. Sin embargo, tal y como se ha comentado anteriormente, la mayor capacidad explicativa de los últimos no implica un mejor comportamiento inferencial fuera de la muestra, como indican los inaceptables índices de condición que han obtenido. Es de destacar que, si en promedio para el periodo considerado se obtenían 3,5 factores, las funciones de valoración sólo han considerado un promedio de 2,25 variables. Esto es, no todos los factores han agrupado a variables relevantes desde un punto de vista valorativo, lo que corrobora una de las críticas realizadas a esta metodología. La agrupación de variables no sigue ningún criterio que contemple la relación entre el valor bursátil y las variables que

lo explican. Únicamente se consideran las correlaciones entre las variables explicativas, por lo que pueden aparecer diferentes factores que agrupen variables sin relevancia valorativa. Un caso extremo es el del año 2002, para el que se habían obtenido 4 factores, pero que únicamente considera una variable explicativa en la función de valoración. Esto implica que 3 de los 4 factores resultan ser irrelevantes. Con la nueva propuesta metodológica, las variables sin relevancia valorativa tienden a agruparse en un único factor, tal y como muestra el análisis realizado sobre la matriz \tilde{X} .

Si el análisis, en línea con lo propuesto en este trabajo, se lleva a cabo sobre la matriz de residuos \tilde{X} , el número de variables consideradas en las funciones de valoración es, en promedio, de 1,7. Esto es, se obtienen modelos más parsimoniosos que los obtenidos con las metodologías anteriores. El coeficiente de determinación corregido es mayor que el obtenido con los modelos que aplican el análisis factorial sobre X , con un promedio del 75%. Y si bien el índice de condición medio es ligeramente superior, 12,4 frente a 11,9, sigue estando por debajo del umbral máximo de 20. También se mejoran los errores absoluto y relativo promedios. En el caso de la metodología factorial clásica, el absoluto es, en promedio, de 2.176,5 miles de euros, mientras que para la metodología propuesta en este trabajo desciende hasta 1.720,3 miles de euros. Los valores relativos son de 8,0% y 6,6% para la metodología tradicional y la nueva propuesta, respectivamente.

Respecto de la relevancia de los factores, en 2 de los 4 años considerados en el periodo se obtienen el mismo número de variables que factores (años 2000 y 2003). Esto es, todos los factores tienen relevancia valorativa. En el caso de los años 2001 y 2002, de los dos factores obtenidos en cada caso, sólo uno obtiene representación en la función de valoración, lo que nos indica que la metodología ha agrupado a las variables no significativas en un solo factor. Es de destacar como con el análisis factorial sobre las variables originales, para el año 2002 se obtiene una agrupación de cuatro factores, si bien finalmente la función de valoración sólo considera una variable explicativa. Esta situación pone de manifies-

to la mayor eficiencia de la propuesta metodológica, que agrupa las variables no relevantes en un único factor, mientras que la metodología tradicional agrupa estas variables en tres factores distintos. En la metodología propuesta, de las 12 variables que agrupa el segundo factor, las 7 de mayor coeficiente, y por tanto las más representativas del mismo, coinciden con el conjunto de variables que el análisis factorial clásico agrupa en los factores 2, 3 y 4, a excepción de las deudas financieras.

De forma resumida, la metodología propuesta consigue obtener en promedio modelos más parsimoniosos que los obtenidos con la metodología clásica de los modelos de valoración basados en

el análisis factorial. Pese a considerar un menor número de variables explicativas, se obtienen mayores coeficientes de determinación corregidos, y ello sin sobrepasar el umbral de multicolinealidad marcado cuantificado por el índice de condición.

CONCLUSIONES

Los modelos teóricos que explican el valor bursátil en función de un restringido número de variables explicativas adolecen en la mayoría de los casos de suficiente capacidad explicativa o, cuando este problema se supera considerando un mayor número de variables, significativa multi-

Tabla 4.- Funciones de valoración

METODOLOGÍA	AÑO	VARIABLES INCLUIDAS EN LA FUNCIÓN DE VALORACIÓN	R ² CORREGIDO	ÍNDICE DE CONDICIÓN	ERROR ABSOLUTO PROMEDIO (miles euros)	ERROR RELATIVO PROMEDIO
Regresión <i>stepwise</i>	2003	Resultado bruto Inmovilizado Inmovilizado material Gastos de personal Pasivo fijo	88,0%	107,6	1.313,1	4,5%
	2002	Total activo Activo circulante Inmovilizado material Deudores Resultado bruto	86,4%	141,8	1.390,6	5,2%
	2001	Total activo Resultado explotación Inmovilizado material Deudores Gastos amort. inmov. Resultado act. ord. Resultado financiero Gastos financieros y asim. Gastos mercaderías	89,7%	200,6	1.387,4	4,3%
	2000	Total activo Resultado bruto Deudores Inmovilizado	86,1%	186,4	1.309,9	4,9%
Análisis factorial sobre \tilde{X}	2003	Acreedores LP Gastos de personal Resultado act. ord.	74,6%	16,7	2.078,7	6,7%
	2002	Gastos de personal	53,6%	9,2	2.775,9	10,1%
	2001	Gastos financieros y asim. Resultado ejercicio	70,4%	8,8	2.018,8	7,3%
	2000	Gastos de personal Resultado ejercicio Resultado financiero	66,1%	13,0	1.832,6	7,7%
Análisis factorial sobre \tilde{X}	2003	Total activo Resultado act. ord.	83,6%	15,4	1.507,4	5,5%
	2002	Total activo	81,5%	13,2	1.521,5	5,7%
	2001	Gastos financieros y asim.	65,7%	7,4	2.176,4	8,3%
	2000	Resultado ejercicio Deudas financieras Tesorería	69,2%	13,7	1.675,9	6,9%

NOTA: la columna "Error absoluto promedio" se ha calculado como el promedio del antilogaritmo de los residuos obtenidos con la regresión. La columna "Error relativo promedio" se ha calculado como el promedio del valor absoluto del ratio entre los residuos y el logaritmo de la capitalización bursátil.

colinealidad, lo que supone aceptar modelos donde los coeficientes de algunas variables carecen de sentido económico. Los modelos factoriales planteados en la literatura obtienen por lo general mayores coeficientes de determinación, al considerar más variables explicativas, además de suavizar significativamente los problemas de la multicolinealidad por el método seguido para la selección de las variables. En el presente trabajo se realiza un análisis crítico de estos últimos modelos de valoración bursátil de empresas, poniendo de manifiesto algunas debilidades, como la discrecionalidad en la selección de variables y la no medición de la multicolinealidad en los modelos de valoración finales.

Tras analizar diferentes posibilidades acerca de la correlación entre las variables económico-financieras, y su capacidad explicativa sobre la capitalización bursátil de las empresas cotizadas, puede concluirse que la aplicación de los modelos factoriales sobre la matriz de correlaciones de las variables explicativas no siempre ofrecerá un resultado satisfactorio. En unos casos porque agrupará en un mismo factor variables que podrían simultáneamente formar parte de la función de valoración; en otros porque agrupará en distintos factores variables que no deberían considerarse conjuntamente de la función de valoración. Como conclusión de lo anterior se propone que la agrupación en factores de las diferentes variables explicativas no debe hacerse en función de la correlación o variabilidad común entre ellas, sino teniendo en cuenta la correlación entre la variabilidad de la capitalización bursátil no explicada por cada una de ellas por separado (parte residual). De esta forma, el proceso de obtención de una función de valoración será más eficiente que el planteado por los modelos factoriales clásicos.

La propuesta metodológica se articula en 4 etapas consecutivas: obtención de los residuos, selección de variables explicativas, estimación de la función de valoración, y validación de la misma. Para la selección de las variables explicativas del valor bursátil, el analista considerará a la variable con mayor coeficiente en el factor (carga factorial), lo que permite identificar de forma inequívoca y automática qué variables se corresponden con cada factor obtenido. Las variables identificadas con un único factor repre-

sentan una parte de la variabilidad del valor bursátil significativamente distinta de la explicada por el resto de variables identificadas con otros factores. Otra diferencia respecto de los modelos factoriales clásicos es la etapa de validación del modelo. En la nueva propuesta la multicolinealidad es cuantificada a través del índice de condición. De esta manera, la función de valoración se obtiene de forma iterativa, y permite eliminar, en el caso en que se detectara multicolinealidad, aquellas variables que la originan.

A través de la aplicación práctica de esta nueva metodología sobre datos de la Bolsa de Madrid (2000-2003) se constata una mayor eficiencia respecto a los que aparecen en la literatura en este campo, al obtener modelos más parsimoniosos, puesto que reducen el número de factores obtenidos respecto al enfoque clásico de valoración con modelos factoriales, y con menor multicolinealidad, cuantificada a través del índice de condición. Como futura línea de investigación, se deja abierta la posibilidad de considerar información no estrictamente contable, y examinar a la luz de los resultados la relevancia valorativa de la misma.

NOTAS

1. Los autores desean agradecer las sugerencias y comentarios realizados por dos evaluadores anónimos, que han contribuido a mejorar el presente trabajo.
2. Arqués *et al.* (2003) analizan la relación existente entre la rentabilidad de los títulos y la información contable de la empresa para el caso español.
3. En esta metodología, se trata de formular modelos econométricos con una capacidad explicativa suficiente del valor bursátil de una empresa y su aplicación, mediante los adecuados tests estadísticos de homogeneidad, a empresas que por motivos jurídicos, económicos o técnicos no negocian sus acciones en los mercados de valores y en las que resulta necesario conocer este valor.
4. En el trabajo de Caballer y Moya (1997) se propone seleccionar para cada factor la variable más correlacionada con el valor bursátil, si bien se advierte que las funciones de valoración así obtenidas no tienen por qué tener el mayor coeficiente de determinación de los posibles.
5. La razón de trabajar con la matriz de covarianzas en lugar de con la matriz de correlaciones está en

que se utilizan los residuos en lugar de las variables originales, y todos ellos vienen expresados en la misma unidad que la variable endógena y , estos, en unidades monetarias, con independencia de que la variable exógena sea una partida o ratio contable/financiero.

6. Los coeficientes de los factores coinciden con el coeficiente de correlación entre dicho factor y la variable.
7. La multicolinealidad sólo podrá ser eliminada en su totalidad en los modelos de regresión simple (una única variable explicativa) o cuando los regresores sean ortogonales entre sí, algo poco probable teniendo en cuenta que se manejan variables económico-financieras.
8. De no corregir por la transformación logarítmica, los modelos de valoración se verían muy afectados por los valores extremos de las empresas más importantes en cuanto a tamaño. En cualquier caso, para poder hacer estimaciones a partir de la función de valoración obtenida deberá deshacerse dicha transformación. Otra posibilidad habría sido realizar un análisis cluster por tamaño para segmentar la muestra inicial de empresas en varias submuestras, pero esto habría limitado el número de grados de libertad para los análisis posteriores, y consecuentemente la capacidad inferencial de las funciones de valoración se habría visto reducida.
9. El índice KMO se obtiene como el cociente

$$\frac{\sum_{i=1}^n \sum_{j=1}^{i-1} r_{ij}^2}{\left(\sum_{i=1}^n \sum_{j=1}^{i-1} r_{ij}^2 + \sum_{i=1}^n \sum_{j=1}^{i-1} a_{ij}^2 \right)}, \text{ donde } r_{ij} \text{ es el}$$

coeficiente de correlación bivariada entre las variables i y j ; a_{ij} es el coeficiente de correlación parcial entre las variables i y j . El índice fluctúa dentro del rango $[0, 1]$, de forma que valores en general próximos a la unidad, y más concretamente por encima de 0,7, aconsejan la utilización del análisis factorial para la extracción de factores representativos de las variables.

10. Entre los trabajos de carácter empírico basados en el modelo de Ohlson (1995) podemos citar a Aboody *et al.* (2006), Amir y Lev (1996), Barth *et al.* (1998), Collins *et al.* (1997), Ely y Waymire (1999), Francis y Schipper (1999), Harris *et al.* (1994), Lev y Zarowin (1999), Nwaeze (1998), y Ryan y Zarowin (2003). En todos ellos se contrasta la relevancia valorativa de las variables analizadas por Ohlson, utilizando para ello el análisis de regresión. Los modelos obtenidos con nuestra propuesta coinciden en identificar dos grandes grupos de variables (periodo 2001-2003), que se pueden relacionar con Balance y Resultado res-

pectivamente. En este sentido, hubiera sido interesante incorporar información no estrictamente contable siguiendo la propuesta de Ohlson, y constatar si dicha información se agrupa en alguno de los factores anteriores o si, por el contrario, se crea un tercer factor que podríamos identificar con la variable "otra información" del citado modelo de Ohlson. Con todo, debemos señalar la diferencia entre este modelo, fundamentado desde un punto de vista teórico, y el que proponemos en este trabajo, que trata de inferir una función de valoración a partir de una evidencia empírica, por lo que la comparación debe ser realizada con prudencia.

BIBLIOGRAFÍA

- ABOODY, D.; HUGHES, J.; LIU, J. (2002): "Measuring Value Relevance in a (Possibly) Inefficient Market", *Journal of Accounting Research*, vol. 40, pp. 965-986.
- AMIR, E.; LEV, B. (1996): "Value Relevance of Non-financial Information: The Wireless Communications Industry", *Journal of Accounting and Economics*, vol. 22, pp. 3-30.
- ARQUÉS, A.; CALVO-FLORES, A.; GARCÍA, D.; MARTÍNEZ, P. (2003): "Relación entre la información contable y la valoración de acciones en el mercado", *Revista Europea de Dirección y Economía de la Empresa*, vol. 12 (1), pp. 143-152.
- BARTH, M.; BEAVER, W.; LANDSMAN, W. (1998): "Relative Valuation Roles of Equity Book Value and Net Income as a Function of Financial Health", *Journal of Accounting and Economics*, vol. 25, pp. 1-34.
- CABALLER, V.; MOYA, I. (1997): "Companies Valuation: An Analogical Stock-Market Empirical Approach", en: *Contemporary Developments in Finance*. París: ESKA.
- COLLINS, E.; MAYDEW, E.; WEISS, I. (1997): "Changes in the Value Relevance of Earnings and Book Value Over the Past Forty Years", *Journal of Accounting and Economics*, vol. 24, pp. 39-67.
- CORELLI, F.; MARCELLINO, M. (2005): "Factor Based Index Tracking", *Journal of Banking & Finance*, vol. 30 (8), pp. 2215-2233.
- ELY, K.; WAYMIRE, G. (1999): "Accounting Standard-Setting Organizations and Earnings Relevance: Longitudinal Evidence from NYSE Common Stocks, 1927-93", *Journal of Accounting Research*, vol. 37, pp. 293-317.
- FRANCIS, J.; SCHIPPER, K. (1999): "Have Financial Statements Lost Their Relevance?", *Journal of Accounting Research*, vol. 37, pp. 319-352.

- HARRIS, T.; LANG, M.; MÖLLER, H. (1994): "The Value Relevance of German Accounting Measures: an Empirical Analysis", *Journal of Accounting Research*, vol. 32, pp. 187-209.
- LEV, B.; ZAROWIN, P. (1999): "The Boundaries of Financial Reporting and How to Extend Them", *Journal of Accounting Research*, vol. 37, pp. 353-385.
- MIRALLES MARCELO, J.L.; MIRALLES QUIRÓS, J.L. (2002): "Factores determinantes del valor bursátil de las empresas portuguesas (1991-1999). Nuevas propuestas metodológicas", *Revista Española de Financiación y Contabilidad*, vol. 112, pp. 495-528.
- MORTON, T.G. (1977): "Factor Analysis, Multicollinearity and Regression Appraisal Models", *The Appraisal Journal*, (October), pp. 578-588.
- MOYA, I. (1995): "Valoración bursátil de empresas. Propuesta de una metodología", *Análisis Financiero*, vol. 66, pp. 92-105.
- MOYA, I. (1996): "Valoración analógico-bursátil de empresas. Aplicación a las cajas de ahorro", *Revista Española de Financiación y Contabilidad*, vol. 86, pp. 199-234.
- NWAEZE, E. (1998): "Regulation and the Valuation Relevance of Book Value and Earnings: Evidence from the United States", *Contemporary Accounting Research*, vol. 15, pp. 547-573.
- OHLSON, J. (1995): "Earnings, Book Values and Dividends in Equity Valuation", *Contemporary Accounting Research*, vol. 11, pp. 661-687.
- RUSHINEK, A.; RUSHINEK, S. (1985): "A Note to Barlev-Levy Theory of the Information Content of Accounting Data and the Management of Security Portfolios Which Include the Least Correlated Stocks: An Empirical Analysis", *Journal of Business, Finance & Accounting*, vol. 12, pp. 117-131.
- RYAN, S.G.; ZAROWIN, P.A. (2003): "Why has the Contemporaneous Linear Returns-Earnings Relation Declined?", *Accounting Review*, vol. 78, pp. 523-553.
- STEELE, A. (1995): "On the Eigen Structure of the Mean Variance Efficient Set", *Journal of Business, Finance & Accounting*, vol. 22, pp. 245-255.