

Modelo discriminante para la industria de los fondos de cobertura (*HEDGE FUNDS*)

(Recibido: 11/febrero/2013 – Aceptado: 24/octubre/2014)

Elitania Leyva Rayón^{*}

Prosper Lamothe Fernández^{**}

Resumen

Los fondos de cobertura (*hedge funds*) desempeñan un papel fundamental en el sistema económico financiando proyectos de inversión con un nivel de riesgo que los bancos y otros intermediarios financieros no están dispuestos a asumir. No obstante, debido a este riesgo y a las características propias del sector, la quiebra es un factor latente. La temprana detección de una posible quiebra podría ayudar a tomar las medidas adecuadas por parte de inversores y autoridades reguladoras antes de que el evento se materialice. Sin embargo, predecir quiebras dentro de esta industria es aún una tarea difícil de realizar. Por ello, el objetivo de este trabajo es analizar factores internos para encontrar al de mayor influencia en la quiebra de los fondos de cobertura. El análisis empírico se realiza con un modelo discriminante aplicado a una base de datos construida manualmente con la información de la revista *MARHedgePerformance & Evaluation Directory* (1999-2006) en los horizontes de tiempo: año de quiebra, uno y dos años antes de la quiebra.

Clasificación JEL: G15, G23, G33.

Palabras Clave: fondos de cobertura, modelo discriminante, factor interno, quiebra.

^{*} Este trabajo forma parte de mi Tesis Doctoral con Premio Extraordinario realizada bajo la tutela del Dr. Prosper Lamothe Fernández en la Universidad Autónoma de Madrid. Profesora de Tiempo Completo, Departamento de Economía, Fundación Universidad de las Américas Puebla. Correo electrónico: <elitania.leyva@udlap.mx>.

^{**} Catedrático de Economía Financiera y Director del Centro Internacional Carlos V, Universidad Autónoma de Madrid. Correo electrónico: <prosper.lamothe@uam.es>.

Introducción

Los mercados financieros han evolucionado rápidamente en los últimos años, la innovación tecnológica y el proceso de globalización presentan grandes ventajas para la integración de la economía mundial. Sin embargo, esta integración también trae consigo problemas considerables debido a la estrecha vinculación financiera entre los diversos participantes de los mercados. La inestabilidad financiera puede provocar efectos negativos en la economía real, por lo que es fundamental comprender las causas que podrían generar una desestabilización económica. Una de las claves para entender los riesgos del comportamiento financiero actual se encuentra en la industria de los fondos de cobertura (*hedge funds*).

Los fondos de cobertura son instituciones financieras no bancarias con una operativa escasamente regulada que les permite endeudarse de forma considerable para aprovechar diversas oportunidades de inversión. Gracias a la liberalización de los flujos monetarios, pueden invertir en todo tipo de activos en los diversos mercados existentes. A pesar de que su existencia tiene más de 50 años, la aparición de este tipo de fondos se incrementó a mediados de los ochenta y durante la década de los noventa con la evolución de diversos instrumentos financieros como los derivados. Pero a pesar de su relevancia para el sistema financiero, el público en general continúa teniendo un limitado entendimiento de esta industria.¹

Caslin (2004) los describe como vehículos de inversión con una regulación más flexible que la de los fondos tradicionales, con una estructura de comisiones en dos niveles: una comisión fija y una comisión variable vinculada al éxito del gestor, quien persigue rentabilidades absolutas a través de estrategias que no están permitidas a los fondos tradicionales como las ventas en corte, el uso de derivados y fuerte apalancamiento. Los gestores también son generadores de alfas, es decir, persiguen rentabilidades absolutas en lugar de relativas basadas en un índice de referencia. Estos fondos cuentan con un periodo de permanencia mínima inicial en el que el inversor queda obligado a permanecer en el fondo debido a las posiciones largas de su inversión.²

¹ Liang (2001) señala que la principal razón es que la información sobre rentabilidades, riesgos y estructuras de comisiones de estos fondos no están disponibles para la consulta pública debido a su reducida regulación.

² Rajwade (2007) menciona que estos fondos cuentan con un Intermediario Principal a través del cual se realizan operaciones como préstamos de activos, financiamiento, etc. También cuentan con un: a) Custodio, que es el órgano no vinculado a los inversores que deberá velar por los activos gestionados; b) Gestor de Riesgos, que monitoriza y controla el riesgo de los activos, la cartera y las exposiciones; c) Asesor; y d) Analistas, entre las principales figuras.

Debido a las características propias de este sector, es innegable que nos encontramos frente a una industria con un fuerte potencial de riesgo sistémico. En este sentido, los reguladores financieros deben tratar de evitar las consecuencias sistémicas que tendrían para el sistema financiero internacional el colapso de un número relevante de fondos de cobertura o la quiebra de un fondo de cobertura de gran tamaño. No obstante, predecir la quiebra dentro de este sector es una tarea aun difícil de realizar. Por ello, estudiamos el fenómeno de quiebra dentro de esta industria a través del análisis de factores internos claves en la continuación o cese de operaciones de este tipo de fondos con ayuda de un modelo de “predicción de quiebra”.

Los modelos de “predicción de quiebra” son herramientas que nos permiten identificar a los factores con mayor relevancia en el proceso de liquidación de una entidad. A pesar de que estos modelos emplean diferentes metodologías (*logit*, *probit*, análisis discriminante, análisis de supervivencia, redes neuronales, entre otros), ninguno de ellos es capaz de hacer una “predicción” de quiebra *per se*, lo que todos ellos tienen en común es la clasificación de las entidades en activas o liquidadas de acuerdo a las características financieras propias de cada entidad, destacando al factor interno que posee la mayor contribución a dicha clasificación.

Por lo anterior, el objetivo de este trabajo es analizar la influencia de factores internos en la quiebra de fondos de cobertura mediante un modelo de “predicción de quiebra” para identificar al factor interno más importante en este evento. El resto del trabajo se desarrolla de la siguiente manera. La sección 1 se refiere a la revisión de literatura clásica sobre los modelos de “predicción de quiebra”. En la sección 2 se exponen la metodología, la base de datos, la selección de las variables y la construcción de las muestras. En la sección 3 se explican los resultados del análisis empírico. Finalmente, se presentan las conclusiones del trabajo.

1. Revisión de literatura

La importancia del estudio de la quiebra radica en que ésta no llega de un momento a otro, sino que se va gestando a lo largo de tiempo. Taffler (1982) señala que la finalidad de los modelos de “predicción de quiebra” es reconocer a tiempo las señales que emiten las entidades próximas a ser liquidadas para poder emprender las acciones necesarias antes de que sea demasiado tarde. A pesar de que ninguno de estos modelos predice una quiebra como tal, sí pueden indicarnos los factores internos más relevantes en el declive de una entidad.

Desde los trabajos pioneros de Beaver (1966) y Altman (1968) son varios los investigadores que continúan aplicando modelos de “predicción de quiebra” a

diferentes industrias y combinaciones de ellas, en varios países y en distintos periodos de tiempo. En esta revisión destacamos los trabajos clásicos más importantes de la literatura sobre este tipo de modelos de acuerdo a la gran cantidad de veces que han sido citados.

Comenzamos con Meyer y Pifer (1970), quienes presentan un modelo lineal de probabilidad para la “predicción de quiebra” bancaria. Utilizan una muestra de 39 bancos activos y 39 bancos liquidados entre 1948 y 1965, y calculan 10 variables financieras relativas a tendencia, variación, cambios inesperados y valores dos años antes de la quiebra. Encuentran que aproximadamente el 80% de los bancos son correctamente clasificados por su modelo en un horizonte temporal de uno o dos años antes de la quiebra. Sin embargo, en un horizonte de tiempo de 3 años la capacidad de discriminación de las variables financieras se reduce mucho.

Por su parte, Deakin (1972) compara los trabajos de Altman (1968) y Beaver (1968). Para ello, aplica el modelo de clasificación dicotómica de Beaver a una muestra de 64 entidades de diferentes sectores³ y el modelo discriminante de Altman a esa misma muestra, con este último obtiene buenos resultados en la “predicción de quiebra” con un horizonte de 3 años antes de la liquidación. Concluye que las dos técnicas, particularmente el análisis discriminante, pueden ser usadas para “predecir la quiebra” a partir de los reportes financieros de las entidades.

Otro que emplea el análisis discriminante es Blum (1974) para distinguir entre entidades sanas y entidades declaradas en quiebra bajo *The Federal Bankruptcy Act*. Utiliza 12 razones financieras referidas a liquidez, rentabilidad y variabilidad.⁴ Su modelo consigue una correcta distinción entre entidades del 94% un año antes de la quiebra, 80% dos años antes y 70% tres años antes de la quiebra. Al comparar sus resultados con los de Beaver (1966) y con los de Altman (1968) concluye que su modelo es más fiable que los de ellos aunque con sus limitaciones.

Por otro lado, Collins (1980) compara el modelo discriminante de Altman (1968) un año antes de la quiebra con el modelo de probabilidad lineal de Meyer y Pifer (1970), esta comparación es con la finalidad de comprobar si la sofisticación del modelo de Meyer y Pifer está justificada en un incremento de la capacidad de “predicción de quiebra”. Para ello, aplica los dos modelos a una muestra de 324 entidades de crédito⁵ y concluye que la capacidad de ambos modelos es buena, pero

³ Utiliza una muestra de 32 entidades liquidadas entre 1964 y 1970, emparejadas con 32 entidades sanas, el emparejamiento lo realiza por tamaño e industria a la que pertenecen.

⁴ Emplea una muestra de 115 entidades liquidadas en el periodo 1954-1968, emparejadas con 115 entidades sanas.

⁵ De las cuales 162 fueron liquidadas entre 1956 y 1976, y las 162 activas fueron seleccionadas al azar.

que el modelo de Meyer y Pifer no presenta resultados superiores al de Altman, por tanto, la sofisticación de su modelo no está justificada en un incremento de la capacidad de predicción.

Taffler (1982) también utiliza un modelo discriminante para la identificación de entidades británicas con riesgo de quiebra. Utiliza una muestra de 23 entidades liquidadas y dos grupos de entidades activas que cotizaban en el *London Stock Exchange* entre 1968 y 1973. Mediante un análisis factorial reduce la dimensionalidad de 50 razones financieras a 6 componentes referidos a rentabilidad, liquidez, apalancamiento y nivel de actividad. Encuentra que los componentes de rentabilidad y apalancamiento son los que tienen la mayor contribución a la función discriminante y, que los de liquidez de corto plazo son menos importantes en la determinación de la propensión de quiebra.

De igual manera, El Hennawy y Morris (1983) aplican dos modelos discriminantes a una muestra de 53 entidades británicas liquidadas emparejadas con 53 entidades sanas, durante el periodo 1960-1971. El primer modelo utiliza un horizonte temporal de 5 años antes de que ocurra la quiebra y el segundo tiene un horizonte de un año. Encuentran que las variables referidas a la rentabilidad son las más importantes en la “predicción de quiebra” en el corto plazo, mientras que para el largo plazo las variables referidas a la liquidez son primordiales.

Peel y Peel (1987) desarrollan un modelo de “predicción de quiebra” mediante un análisis *logit* aplicado a entidades británicas. Utilizan una muestra de 56 entidades liquidadas, 56 sanas y 34 con problemas financieros pero que aún están activas.⁶ Además, emplean variables con alta capacidad de predicción referidas a tamaño, edad e industria a la que pertenecen las entidades. Al comparar sus resultados con los obtenidos por el discriminante encuentran que la precisión en la clasificación del discriminante fue superior al del *logit* en la muestra original (83,1% y 78,9%, respectivamente) y en la muestra de validación (77,1% y 58,3%, respectivamente).

Por otro lado, Yam y Kiang (1992) comparan un modelo basado en el análisis discriminante con un modelo basado en redes neuronales, en lo referente a robustez, adaptabilidad y capacidad de predicción, utilizando diferentes probabilidades *a priori* y costos de clasificación errónea, con el objetivo de encontrar cuál de las dos metodologías tiene mayor capacidad en la “predicción de quiebra”. Encuentran que el modelo de redes neuronales tiene una ligera mayor precisión en la predicción pero muy cercana a los resultados del análisis discriminante.

⁶ También utilizan una muestra de validación formada por 12 entidades liquidadas, 24 sanas y 12 con problemas financieros pero que continúan activas durante el periodo 1982-1985.

Yim y Mitchell (2005) comparan la capacidad de “predicción de quiebra” de modelos basados en análisis discriminante, *logit*, *probit*, redes neuronales, y un modelo híbrido de redes neuronales. Estos modelos son aplicados a una muestra de 121 entidades brasileñas y utilizan 22 razones financieras. Con los modelos *logit* y *probit* obtienen resultados satisfactorios pero no mejores que con el modelo discriminante. De igual manera, encuentran que los resultados de las redes neuronales son inferiores a los del análisis discriminante.

Finalmente, Beaver, Correia y McNichols (2011) presentan una amplia investigación que expone los tres principales aspectos de la literatura académica sobre la predicción de quiebras financieras y su evolución: 1) las variables dependientes y explicativas que diferentes investigadores han utilizado; 2) las metodologías que se han empleado (regresión múltiple, análisis univariante, análisis discriminante, *logit* condicional, entre otros); y 3) la forma en la que se han modelado las quiebras financieras. Describen algunas limitaciones de la información contable (y de mercado) y mencionan que la forma en la que los modelos de predicción de quiebra deben estructurarse depende de lo que el investigador trate de predecir. Por último, hacen interesantes sugerencias para futuras investigaciones sobre predicción de fraude y estimación del riesgo de litigios.

Por lo anterior, encontramos que el análisis discriminante es una de las técnicas más populares en los modelos de “predicción de quiebra”. Aunque se han realizado comparaciones entre distintas metodologías (*logit*, *probit*, árbol de decisiones, discriminante bayesiano, análisis de supervivencia, redes neuronales, entre otros) con el objetivo de aumentar la capacidad de predicción, en la mayoría de los casos los resultados obtenidos con el análisis discriminante son mejores que los obtenidos por cualquier otra técnica (Altman y Narayanan, 1997).

2. Metodología

2.1 Análisis discriminante

El análisis discriminante es una de las modalidades más aplicadas en el ámbito de la economía de la empresa.⁷ Se trata de una técnica muy robusta y fiable que permite estudiar el grado en que diferentes grupos, establecidos *a priori*, difieren entre sí.⁸

⁷ Esta sección está basada en los libros de Análisis Multivariante de: Hair, Anderson, Tatham y Black (1999); Lévy-Mangin y Varela (2003); y De la Garza, Morales y González (2013).

⁸ Los grupos se establecen con anterioridad de acuerdo a la información, no se forman al azar. Por ejemplo: grupo A (Entidad Activa) y grupo B (Entidad Liquidada). Dado que se dispone de información sobre si la entidad está activa o si está liquidada, dicha entidad sólo puede pertenecer a uno de estos 2 grupos establecidos *a priori*.

Clasifica a las entidades en grupos a partir de los valores de un conjunto de variables, y cada entidad puede pertenecer a un sólo grupo. Tiene dos fines de aplicación: descriptivo y predictivo. El primero se emplea con la finalidad de entender como la combinación lineal de las variables independientes permite discriminar entre los dos grupos definidos en la variable dependiente y (que toma valor 0 si la entidad está activa y valor 1 si está liquidada). El segundo utiliza los valores de las variables para adscribir a cada entidad al grupo que le corresponda. A las variables que se utilizan para hacer la clasificación se les denomina variables de clasificación o discriminantes. La relación presenta la forma genérica:

$$y = f(x_1, x_2, \dots, x_m) \quad (1)$$

Donde: la variable dependiente y es categórica; las variables explicativas x_i son cuantitativas.

A diferencia de los modelos de regresión, el análisis discriminante no está basado en un modelo teórico previo, es decir, no parte de una hipótesis de tipo “el conjunto de variables x_i que explican el fenómeno y ”, sino que es un análisis exploratorio que busca aquella combinación de variables clasificatorias que optimice un resultado definido previamente. Su objetivo es obtener una puntuación en la variable dependiente y que se asocia con el conjunto de variables discriminantes. La información de las variables se sintetiza en unas funciones discriminantes que se utilizan en el proceso de clasificación. En el caso de utilizar k variables de clasificación, la función discriminante de Fisher D es la mejor opción, ya que se trata de una función lineal que representa la pertenencia a un grupo y se estructura como una combinación lineal de las k variables x en la que se plantea obtener los coeficientes de ponderación u_j . Si consideramos que existen n observaciones, podemos expresar la función discriminante para las n observaciones:

$$D_i = u_1 x_{1i} + u_2 x_{2i} + \dots + u_k x_{ki} \quad (2)$$

Donde: $i = 1, 2, \dots, n$

D_i es la puntuación discriminante correspondiente a la observación i -ésima.

Expresando las variables explicativas en desviaciones respecto a la media, D_i también lo estará. El criterio para obtener la función discriminante de Fisher es: “la maximización de la variabilidad entre-grupos sobre la variabilidad intra-grupos”. Con este criterio se trata de determinar un eje discriminante de forma que las distri-

buciones proyectadas sobre el mismo estén lo más separadas posible entre sí y, al mismo tiempo, que cada una de las distribuciones esté lo menos dispersa.

En el proceso de maximización se determinarán: a) los coeficientes normalizados (u_1, u_2, \dots, u_k); b) las puntuaciones discriminantes de cada fondo, que determinarán el grupo al que pertenecen; c) el autovalor asociado a la función discriminante, que nos indica la proporción de varianza total explicada por esta función; d) la contribución de cada variable a la función discriminante, que se calcula mediante las correlaciones entre las variables y las puntuaciones discriminantes dentro de cada grupo; e) las contribuciones de cada variable a la función discriminante, que se exponen en una matriz de estructura; y f) Los centroides (\bar{X}_I y \bar{X}_{II}), que se obtienen promediando las puntuaciones discriminantes para todos los fondos dentro de un grupo. Sustituyendo x_1, x_2, \dots, x_k en la función discriminante de Fisher D de cada grupo por los vectores de centroides tenemos:

$$\bar{D}_I = u_1 \bar{X}_{1,I} + u_2 \bar{X}_{2,I} + \dots + u_p \bar{X}_{k,I} \quad (3)$$

$$\bar{D}_{II} = u_1 \bar{X}_{1,II} + u_2 \bar{X}_{2,II} + \dots + u_k \bar{X}_{k,II} \quad (4)$$

El punto de corte discriminante C se calcula promediando \bar{D}_I y \bar{D}_{II} :

$$C = (\bar{D}_I + \bar{D}_{II}) / 2 \quad (5)$$

El criterio para clasificar a la entidad i es el siguiente:

Si $D_i < C$, se clasifica a la entidad i en el grupo de liquidados

Si $D_i > C$, se clasifica a la entidad i en el grupo de activos

Sin embargo, en este proceso de decisión se pueden cometer dos tipos de errores:

Error Tipo I: Se clasifica como liquidada a una entidad activa

Error Tipo II: Se clasifica como activa a una entidad liquidada

Para medir cuáles variables tienen mayor poder discriminante en la clasificación se calcula el coeficiente de correlación entre cada una de las variables y la función discriminante (en valor absoluto). Los métodos más utilizados para la jerarquización de variables son de carácter iterativo y requieren definir previamente una regla de decisión para medir la bondad de ajuste en cada paso o iteración. Una

regla de decisión suele ser la minimización del estadístico Lambda de Wilks (Λ). De manera que cuanto menor sea este estadístico mayor será el grado de ajuste, ya que este estadístico representa la proporción de varianza total de las puntuaciones discriminantes no explicadas por las diferencias entre grupos. Para determinar la capacidad discriminante del conjunto de variables, incluimos simultáneamente todas las variables en el modelo y utilizamos los coeficientes estandarizados para determinar la contribución de cada una a la discriminación.

Respecto a la bondad de ajuste, la matriz de clasificación determinará la capacidad de predicción de la función discriminante. En el análisis discriminante, la razón de aciertos es análogo al R^2 de la regresión, y el contraste F para la significación estadística de R^2 es, por tanto, análogo al contraste Chi -cuadrado de significación. No obstante, esta razón de aciertos podría estar sesgado a la alza si se utilizan las mismas observaciones. Para disminuir el sesgo utilizamos el método de validación cruzada (también conocido como *jackknife*), el cual consiste en estimar el modelo con la muestra pero dejando fuera una de las observaciones y a partir del nuevo modelo ($n-1$) se clasifica a la observación que se dejó fuera. El proceso se repite para las n observaciones dejando fuera una observación diferente cada vez.

Por otra parte, las probabilidades de clasificación en cada grupo se calculan utilizando el teorema de Bayes, que permite el cálculo de las probabilidades *a posteriori* a partir de las probabilidades *a priori* y de la información muestral contenida en las puntuaciones discriminantes. Considerando el caso de G grupos, este teorema establece que la probabilidad *a posteriori* de pertenencia a un grupo g con una puntuación discriminante D [$\text{Prob}(g/D)$] es:

$$\text{Prob}(g/D) = \frac{\pi_g + \text{Prob}(D/g)}{\sum_{i=1}^G \pi_i + \text{Prob}(D/i)} \quad (6)$$

En el segundo miembro aparecen las probabilidades *a priori* (π_g) y las probabilidades condicionadas [$\text{Prob}(D/g)$], estas últimas se obtienen calculando la probabilidad de la puntuación observada suponiendo la pertenencia a un grupo g . La clasificación de cada entidad se puede realizar mediante la comparación de las probabilidades *a posteriori*.

El objetivo del análisis discriminante es encontrar una combinación lineal de las variables independientes que permita diferenciar de la mejor manera a los

grupos o niveles que la variable dependiente tiene. Una vez que se encuentra la función, ésta puede utilizarse para clasificar nuevos elementos.⁹

2.2 base de datos

La mayor parte de los modelos de “predicción de quiebra” utilizan datos como los publicados en el balance general y en la cuenta de pérdidas y ganancias de las entidades, debido a que son datos que están disponibles para el público en general. También se utilizan datos remitidos a los reguladores, pero este tipo de datos no están a disposición de todos. Asimismo, se utilizan datos internos a los que sólo tiene acceso la propia entidad, y por ende, este tipo de trabajos tienen una frecuencia mucho menor en la literatura científica.

La información sobre los fondos de cobertura es uno de los mayores problemas a los que se enfrentan los investigadores, ya que son fondos de inversión que no están regulados por organismos oficiales y por ello no tienen la obligación de reportar sus datos a la opinión pública. La escasa transparencia de esta industria respecto a su información interna genera varias limitaciones para su estudio. Por fortuna, existen fondos de cobertura que envían su información de manera voluntaria a las principales bases de datos privadas.¹⁰ No obstante, la información enviada por los gestores es limitada, ya que envían sólo la información que consideran necesaria para atraer nuevos inversores.

La mayoría de las bases de datos comenzaron a recolectar información de esta industria a mediados de la década de los noventa, de ahí que la mayor parte de los trabajos empíricos sobre este sector sea relativamente nueva (Getmansky, Lo y Mei, 2004). Por su parte, Barry (2003) argumenta que analizar las bases de datos de fondos de cobertura es complicado por tres causas: 1) No todas las bases retienen los datos históricos acerca de los fondos que han sido liquidados o que han parado de enviar sus reportes por diversas razones; 2) Estos fondos tienen prohibido hacer promoción directa hacia cualquier inversor bajo la Ley de Sociedades de Inversión de 1940; 3) Los gestores de estos fondos son reservados en cuanto al reporte de sus rentabilidades.

En relación con la primera causa, Fung y Hsieh (2000, 2002) señalan que una vez que un fondo deja de enviar su información a la base de datos ésta lo

⁹ Para un análisis más detallado sobre cada una de las etapas del análisis discriminante puede consultar el capítulo 5 de De la Garza, Morales y González (2013).

¹⁰ Las principales bases comerciales que recolectan información sobre fondos de cobertura son: *Tremont Advisory Shareholders Services* (TASS), *Manager Account Reports (MARHedge)* y *Hedge Fund Research, Inc.* (HFR).

borra y es casi imposible obtener la información. Otra desventaja que encuentran es que únicamente los fondos que tienen o han tenido un buen desempeño son los que quieren ser incluidos en las bases, lo cual significa que las rentabilidades en estas bases podrían ser más altas que las rentabilidades de todo el universo de los fondos de cobertura, dando lugar al sesgo en la información llamado “tendencia de supervivencia”. Brealey y Kaplans (2001) argumentan que este sesgo se refiere al alza en la estimación de la rentabilidad media de los fondos de cobertura debido a que, a la hora de mirar las rentabilidades, los inversores no tienen en cuenta la información de los fondos liquidados.

Por otra parte, es común que la entrada de un fondo de cobertura a una base de datos traiga consigo rentabilidades históricas que se obtuvieron en un periodo de incubación antes de entrar a la base (Fung y Hsieh, 1997), lo que genera un sesgo en la información llamado “historias instantáneas de rentabilidad”. Este sesgo se refiere a que hay fondos que son incorporados a la base con casi dos años de historia. Debido a esto, muchas bases de datos completan los datos faltantes creando otro sesgo en la información llamado “complementado de datos” (Fung y Hsieh, 2000). Además, es probable que los fondos que comenzaron a operar el mismo año en el que fueron liquidados no sean incluidos en la base, causando otro sesgo en la información (Brown, Goetzmann y Park, 1997).

Aunado a lo anterior, los proveedores de estas bases ofrecen clasificaciones de fondos de cobertura que no tienen mucho en común, esto se debe a que los gestores emplean una amplia gama de estrategias de inversión, mientras que los proveedores deben proporcionar algún tipo de esquema de clasificación. El problema radica en que estas clasificaciones varían dependiendo de la base de datos.¹¹ La mayor parte de las estrategias que figuran en una base de datos se basan en la autoclasificación reportada por los gestores de acuerdo a las estrategias de cada proveedor.

Debido a que la inclusión de estos fondos en cualquiera de las bases es meramente voluntaria, el resultado es que no toda la información está disponible o completa, como los procesos de inversión, los activos en los que invierten, los niveles de apalancamiento, los términos específicos de los contratos, el periodo mínimo de permanencia, entre otros (Fung y Hsieh, 2002). Por ello, no es sorprendente que los estudios basados en diferentes bases de datos tengan conflictos en las conclusiones al encontrar distintos resultados. No obstante, es importante analizar el sector de los fondos de cobertura con diferentes bases de datos para corroborar o

¹¹ Por ejemplo, TASS es la única base de datos de fondos de cobertura que incluye la estrategia Mercados de Futuros (*managed futures*), los cuales tiene sus actividades limitadas al mercado de futuros.

hallar nuevos resultados con la misma metodología de otros autores pero con bases de datos y periodos de tiempo diferentes.

Para realizar nuestro análisis empírico hemos construido manualmente una base de datos¹² a partir de la información sobre fondos de cobertura publicada en la revista *MARHedge Performance & Evaluation Directory* perteneciente a la base de datos *MARHedge* (1999 – 2006). Esta base clasifica a los fondos en 10 estrategias de inversión generales:¹³ Hechos Relevantes, Mercados Emergentes, Mercados Internacionales, Mercados Establecidos, Mercados Globales, Posición Larga/Apalancamiento, Neutral al Mercado, Sectores, Ventas en Corto y Fondos de Fondos¹⁴ (*event driven, global emerging, global international, global established, global macro, long only/leverage, market neutral, sector, short-sellers y fund of funds*, respectivamente).

En 1994 fue lanzada la base de datos *MAR* durante la primera conferencia anual internacional sobre la inversión en fondos de cobertura en las Bermudas. En Marzo del 2002, la base de datos y la propiedad intelectual de *MAR* fue comprada por *Zurich Capital Markets Inc*, Finalmente, en el 2006 fue vendida a *Euromoney Institutional Investor*, cambiando su nombre a *MARHedge Institutional Investor*. Desgraciadamente, a partir del segundo semestre del 2006 la revista cambió su formato y dejó de publicar información cuantitativa interna de los fondos, en su lugar comenzó a publicar información relativa a relaciones públicas, congresos, eventos, cenas de gala, y diversas cuestiones sociales. Debido a esto, la información utilizada en nuestra investigación concierne únicamente al rango de años en los que se publicó información interna.¹⁵

Nuestra base de datos quedo conformada por 3,097 fondos de cobertura activos¹⁶ y 1,761 fondos de cobertura liquidados hasta junio del 2006. La información

¹² De la misma manera, Brown, Goetzmann e Ibbotson (1999) recolectan manualmente la información de los fondos de la publicación *The U.S. Offshore Directory*, y a partir de esta revista crean una base con información anual.

¹³ Rajwade (2007) explica algunas de las estrategias más importantes de la industria de los fondos de cobertura.

¹⁴ Los fondos de cobertura que siguieron la estrategia fondos de fondos no fueron contemplados en el análisis, ya que de acuerdo con Chan, Getmansky, Hass y Lo (2005) y Baquero, Horst y Verbeek(2005), la finalidad de no incluirlos es evitar la duplicación de fondos, ya que muchos de los fondos invierten en otros fondos de cobertura.

¹⁵ Ésta es una de nuestras principales limitaciones, ya que sería muy enriquecedor ampliar el periodo de análisis con el fin de contrastar los resultados del modelo discriminante para antes y después de la crisis financiera del 2007, analizar si los factores internos seleccionados cambian su influencia y significancia en la clasificación de fondos de cobertura activos y liquidados después de la crisis. No obstante, consideramos que el trabajo contribuye al estudio empírico de esta industria, ya que la mayoría de artículos en español sobre fondos de cobertura son de carácter divulgativo y/o descriptivo debido a la dificultad en la obtención de información interna.

¹⁶ De acuerdo con el sitio web de la empresa *Hedge Fund Reserch, Inc.*, en la actualidad existen alrededor de 10,000 fondos de cobertura. No obstante, esta cifra es sólo una estimación dado que los fondos de cobertura no requieren registrarse ante algún órgano regulador del sistema financiero internacional. Por ello, es difícil conocer con exactitud el número de fondos de cobertura existentes y el monto total del capital gestionado.

de los fondos de cobertura activos se obtuvo en la última publicación de la revista, mientras que la información de los fondos de cobertura liquidados se obtuvo mediante la construcción de un cementerio, debido a que la revista *MARHedge* no cuenta con una sección de fondos liquidados. El cementerio se construyó de la siguiente manera: comparamos la primera publicación de la revista (enero-junio/1999) con la segunda publicación (julio-diciembre/1999), los fondos de cobertura de la primera publicación que resultaron ausentes en la segunda publicación fueron catalogados como liquidados y agrupados por estrategia seguida; posteriormente, comparamos la segunda publicación (julio-diciembre/1999) con la tercera publicación (enero-junio/2000), los fondos de cobertura de la segunda publicación que resultaron ausentes en la tercera publicación fueron catalogados como liquidados y agrupados por estrategia seguida, y así sucesivamente. Este procedimiento se llevó a cabo con todas las publicaciones de la revista durante el periodo 1999-2006.

No obstante, la desaparición de un fondo de cobertura de una base de datos puede deberse a motivos distintos a la quiebra (fusiones, adquisiciones, cierre a nuevos inversores, etc.). Ésta es una limitación, ya que no podemos verificar cuales de los fondos de cobertura catalogados como liquidados realmente lo estaban. Sin embargo, no es un problema exclusivo nuestro, ya que Ackermann, McEnally y Ravenscraft (1999) y Gregoriou (2002) tampoco logran verificar cuales fondos de cobertura que pararon de enviar su información a las bases de datos MAR y HFR continuaban operando. Sin embargo, Liang (2001) encuentra que la razón principal por la que un fondo desaparece de una base es el pobre desempeño en la obtención de rentabilidad, el cual es uno de los principales motivos de quiebra dentro de esta industria.

2.3 selección de las variables discriminantes

Para llevar a cabo el análisis empírico construimos las variables propuestas y utilizadas por Chan, Getmansky, Hass y Lo (2005), quienes también analizan la quiebra de los fondos de cobertura.¹⁷ A partir de la información publicada en la revista *MARHedge* construimos las variables discriminantes (Tabla 1). La variable dependiente y_{it} se asocia con cada una de las variables de clasificación, indicándonos si el fondo de cobertura i está activo o liquidado en el año t . El fondo de cobertura i puede pertenecer sólo a uno de los dos grupos en el año t .

¹⁷ Sin embargo, nosotros empleamos otra metodología, una base de datos distinta y un objeto de estudio diferente.

Tabla 1
VARIABLES DE CLASIFICACIÓN PARA EL MODELO DISCRIMINANTE

<i>VARIABLES DE CLASIFICACIÓN</i>	<i>DEFINICIÓN</i>
<i>Edad</i>	Se refiere al número de años transcurridos desde la fundación del fondo hasta el momento del estudio.
<i>Rentabilidad</i>	Rentabilidad total del año actual en porcentaje. La rentabilidad total anual es la suma acumulada de las rentabilidades mensuales del año en porcentaje.
<i>Activos gestionados</i>	Logaritmo natural del total de los activos gestionados actuales en millones de dólares. ¹⁸
<i>Flujos de fondos</i>	Los flujos de fondos totales del año t divididos por los activos gestionados del año anterior, en el que $flujos\ de\ fondo_t = actgest_t - actgest_{t-1} (1 + rent_t)$, donde $actgest_t$ es el total de activos gestionados en el año t, $actgest_{t-1}$ es el total de activos gestionados del año anterior y $rent_t$ es la rentabilidad total del fondo (%) en el año t.

Fuente. Chan, Getmansky, Hass y Lo (2005).

El factor edad fue seleccionado debido a que regularmente los “viejos” fondos de cobertura con una trayectoria reconocida y una consolidación en el mercado financiero tienen mayor experiencia frente a las desavenencias de los mercados, en contraste con los fondos “jóvenes”. Por otro lado, un bajo nivel de flujos de fondos podría indicar problemas de competitividad en el sector, ya que los inversores regularmente se desplazan hacia los fondos que ofrecen mejores rentabilidades. Sin embargo, en un intento por parte del gestor de incrementar los flujos de fondos, éste podría asumir un mayor riesgo y terminar afectando la viabilidad del fondo.

El hecho de que un fondo de cobertura tenga problemas en la rentabilidad es muy importante debido a que –aunque las rentabilidades pasadas no aseguren rentabilidades futuras– este factor puede darnos una idea de la buena o mala gestión que está teniendo el fondo. Además, se trata de un factor relevante en la continuación o cese de actividades en esta industria. La importancia de los activos gestionados radica en que, aunque el tamaño del fondo no mida exactamente su solvencia, el fondo podría verse afectado si frente a una oportunidad de inversión no dispone de capital suficiente y debe recurrir al crédito, que podría estar restringido en un contexto de crisis.¹⁹

¹⁸ El propósito de utilizar el logaritmo natural es hacer comparable cada pareja de fondos y evitar problemas relacionados con el tamaño de los mismos. La transformación logarítmica disminuye los problemas de comparación relativa al tamaño entre las entidades debido a que estabiliza la varianza de la variable.

¹⁹ Cabe destacar que las variables relativas a la *rentabilidad* y a la *liquidez* son los factores más relevantes en los modelos de “predicción de quiebra” de la revisión de literatura. Por otro lado, la utilización de razones financieras sirve para hacer comparables entidades de diferentes sectores y nuestro análisis está enfocado exclusivamente al sector de los fondos de cobertura, por lo que utilizar razones financieras carece de sentido.

2.4 construcción de las muestras

A partir de la información de la base de datos construida desarrollamos un modelo discriminante para la industria de los fondos de cobertura con los siguientes propósitos:

- Determinar si existen diferencias estadísticamente significativas entre las medias de las variables de clasificación de los dos grupos (activos y liquidados).
- Determinar cuál de las variables de clasificación cuantifica mejor las diferencias entre ambos grupos.
- Establecer los procedimientos para clasificar a los fondos dentro de cada grupo en base a sus puntuaciones discriminantes sobre las variables de clasificación.
- Analizar la información anual disponible de cada uno de los fondos para determinar que factor interno tiene mayor relevancia en la liquidación de los fondos de cobertura.

De los 4,858 fondos de cobertura que componen nuestra base de datos, sólo empleamos los fondos que cuentan con información completa relativa las variables discriminantes seleccionadas y que estuvieran activos por lo menos un año. Los fondos seleccionados se ordenaron por estrategia seguida y número de años con información completa. En la Tabla 2 se observa que contamos con 744 fondos de cobertura que tienen información relativa a 3 años. En este grupo, para cada uno de los 269 fondos liquidados se tiene información referida a su año de liquidación, al año anterior a la liquidación y a dos años antes de la liquidación. De igual manera, para los 475 fondos activos de este grupo se cuenta con la información referida a 3 años de actividad.

Tabla 2
Fondos de cobertura por estrategia seguida y con información completa

<i>Estrategia</i>	<i>Fondos con información relativa a un año</i>		<i>Fondos con información relativa a dos años</i>		<i>Fondos con información relativa a tres años</i>	
	<i>Liquidados</i>	<i>Activos</i>	<i>Liquidados</i>	<i>Activos</i>	<i>Liquidado</i>	<i>Activos</i>
Hechos relevantes	14	24	32	32	24	35
Mercados emergentes	8	29	19	37	13	25
Mercados establecidos	44	71	50	110	56	117
Mercados internacionales	5	27	7	29	7	18
Mercados globales	17	14	9	17	20	18
Posición larga/apalancamiento	3	22	7	25	4	22
Neutral al mercado	58	163	64	195	105	189
Sectores	25	41	27	65	32	47
Ventas en corto	1	3	4	7	8	4
<i>Total</i>	<i>175</i>	<i>394</i>	<i>219</i>	<i>517</i>	<i>269</i>	<i>475</i>

Fuente: Elaboración propia a partir de la información de la revista *MARHedge Performance & Evaluation Directory* (1999-2006).

Sin embargo, en la selección de los fondos nos encontramos con dos problemas. En primer lugar, los grupos varían ampliamente en cuanto al número de fondos, por ejemplo, los fondos liquidados con información relativa a 2 años (219) representan el 42.3% de los fondos activos con información relativa al mismo número de años (517). Lo anterior afecta a la estimación de la función discriminante, ya que en la etapa de clasificación el grupo más grande (activos) tendría una posibilidad mayor de clasificación sobre el grupo de liquidados. En segundo lugar, no todas las observaciones corresponden al mismo periodo de tiempo dentro de cada subconjunto.

Para resolver estos dos problemas llevamos a cabo un proceso de emparejamiento de fondos por estrategia seguida. Los fondos activos son emparejados con los fondos liquidados que operaron durante el mismo periodo de tiempo y por estrategia seguida tomando como año base el año de liquidación del fondo liquidado. Con esto, se busca evitar que uno de los grupos (activos o liquidados) tenga más peso en un periodo de tiempo que en otro, ya que podría ser que la situación económica, financiera o legal haya cambiado en el tiempo y dichos cambios fuesen recogidos de forma diferente en uno y otro grupo. Enseguida, se busca aquella combinación de las variables discriminantes que optimicen la correcta clasificación de los fondos con la información referente al año de quiebra, al año anterior a la quiebra y a dos años antes de la quiebra para comprobar su poder clasificador en el tiempo.

No obstante, el emparejamiento de los fondos trae consigo algunos obstáculos, ya que un gran número de fondos no llega ni al tercer año de operación debido a las características propias de esta industria,²⁰ lo que limita el tamaño de las muestras. Por tanto, debido a que no se cuenta con información completa para el mismo periodo de tiempo para todos los fondos de la base de datos, hemos construido 3 muestras (de la M1 a la M3) para poder comparar el poder clasificador de las variables discriminantes en distintos horizontes de tiempo (Tabla 3).

²⁰ Brown, Goetzman e Ibbotson (1999), encuentran que la edad promedio de los fondos difícilmente llega a superar los 3 años de actividad debido principalmente a las pobres rentabilidades obtenidas.

Tabla 3
Número de parejas de fondos por estrategia y mismo periodo de actividad

Estrategias	Un año (año de quiebra)				2 años (año de quiebra y año anterior a la quiebra)				3 años (año de quiebra, año anterior a la quiebra y dos años antes de la quiebra)						
	2006	2005	2005	2002	M1 (Total de parejas)				M2 (Total de parejas)				M3 (Total de parejas)		
					2003, 2004	2004, 2005	2005, 2006		2002, 2003, 2004	2003, 2004, 2005	2004, 2005, 2006				
Hechos relevantes	2	2	1	39	44	3	9	2	14	5	14	2	21		
Mercados emergentes	0	0	0	20	20	0	1	0	1	0	0	0	0		
Mercados establecidos	4	6	1	76	87	1	8	11	20	9	10	9	28		
Mercados internacionales	0	0	0	21	21	0	0	0	0	0	0	0	0		
Mercados globales	0	1	0	17	18	0	0	1	1	7	7	0	14		
Posición larga/apalancamiento	0	0	0	8	8	0	0	1	1	0	0	1	1		
Neutral al mercado	3	17	2	93	115	3	26	7	36	12	55	6	73		
Sectores	2	1	0	30	33	0	1	0	1	5	9	2	16		
Ventas en corto	0	0	0	4	4	0	0	1	1	0	7	0	7		
Total de parejas	11	27	4	308	350	7	45	23	75	38	102	20	160		

Fuente: Elaboración propia a partir de la información de la revista *MARHedge Performance & Evaluation Directory* (1999-2006).

La muestra M1 está compuesta por 4 submuestras que corresponden a los fondos que cuentan con información referida sólo al año de liquidación (Año 0); la muestra M2 está compuesta por 3 submuestras de fondos que cuentan con información relativa a 2 años [año de quiebra (Año 0) y año anterior a la quiebra (Año 1)] y la muestra M3 está compuesta por 3 submuestras de fondos que cuentan con información de 3 años [año de quiebra (Año 0), año anterior a la quiebra (Año 1) y 2 años antes de la quiebra (Año 2)]. La elaboración de más de una muestra permite comparar la capacidad discriminante del modelo en diferentes horizontes de tiempo.

Las muestras están formadas por un total de 1,110 fondos de cobertura repartidos en 585 parejas, cada pareja está formada por un fondo activo y uno liquidado con información para los mismos años de actividad. El año base de cada pareja corresponde a la última publicación de los datos del fondo liquidado en la revista *MARHedge*. Finalmente, para detectar posibles observaciones atípicas que puedan causar errores en el análisis calculamos la distancia de Mahalanobis (dM_{ij}^2). Un

fondo es clasificado como atípico y eliminado de la muestra (junto con su respectiva pareja) cuando su dM_{ij}^2 al centroide del grupo es excesiva al 99,9% de confianza.²¹ Aunque la proporción de atípicos sea pequeña (3%) su presencia podría perjudicar los resultados.

3. Resultados

Para tener una idea general de la variable con mayor poder discriminante realizamos el cálculo de las medias y las desviaciones estándar de las variables discriminantes.²² En los resultados encontramos que en la muestra M1 (año de liquidación o año 0) las medias de todas las variables presentan valores más altos para el grupo de activos que para el de liquidados, las diferencias más significativas se encuentran en las variables edad y rentabilidad. En la muestra M2 las diferencias de medias son mayores en las variables edad y activos gestionados, en esta muestra las diferencias en las medias se hacen más pronunciadas en el año 0 y lo mismo ocurre en la muestra M3. De modo que a medida que nos vamos acercando al momento de la quiebra las diferencias en las medias de las 4 variables se van intensificando, lo que resulta normal debido a que nos acercamos al año en el que el fondo fue efectivamente liquidado.

El análisis descriptivo nos proporciona una idea intuitiva de las variables que tienen potencial de discriminación en cada muestra. Para confirmar o descartar esta primera impresión realizamos un análisis de la varianza (ANOVA) en el que contrastamos la significación estadística de cada variable a través de una prueba de igualdad de medias. Para considerar a las variables de modo individual utilizamos la Lambda de Wilks (Λ), si el valor de la Λ es cercano a cero la variable discrimina bien, mientras que valores próximos a la unidad sugieren la no existencia de diferencias entre las medias. La prueba de igualdad de medias nos revela que existen diferencias de medias entre los 2 grupos de la muestra M1 en todas las variables excepto flujos de fondos. En el Año 0 y el Año 1 de la muestra M2 existen diferencias de medias entre los 2 grupos para las variables edad y activos gestionados. En la muestra M3 las diferencias de las medias entre los grupos aumentan para las 4 variables conforme nos acercamos al año de liquidación (Tabla 4).

²¹ Otros que utilizan la distancia de Mahalanobis para eliminar los *outliers* de sus muestras son Laffarga, Martín, y Vázquez (1991) y López, Moreno y Rodríguez (1994) en sus respectivos modelos de "predicción de quiebra". Por excesiva nos referimos a una dM_{ij}^2 superior a tres veces la dM_{ij}^2 media de la muestra.

²² El análisis empírico se ha realizado con el paquete estadístico SPSS 15.0 (*Statistical Package of Social Sciences*).

Tabla 4
Contrastes de igualdad de medias (ANOVA)

	<i>Periodo</i>	<i>Edad</i>	<i>Rentabilidad</i>	<i>Activos gestionados</i>	<i>Flujos de fondos</i>
M1	Lambda de Wilks	0.89	0.98	0.98	1.00
	Estadístico <i>F</i>	84.67	11.56	16.95	2.51
	Niveles de significación	0.00	0.00	0.00	0.11
M2	Lambda de Wilks	0.62	1.00	0.94	1.00
	Estadístico <i>F</i>	88.49	0.35	8.39	0.06
	Niveles de significación	0.00	0.55	0.00	0.80
M2	Lambda de Wilks	0.63	0.99	0.93	1.00
	Estadístico <i>F</i>	85.22	1.64	10.21	0.52
	Niveles de significación	0.00	0.20	0.00	0.47
M3	Lambda de Wilks	0.98	0.98	1.00	1.00
	Estadístico <i>F</i>	6.69	7.40	0.16	1.28
	Niveles de significación	0.01	0.01	0.69	0.26
M3	Lambda de Wilks	0.98	0.99	1.00	1.00
	Estadístico <i>F</i>	6.96	3.33	1.19	0.20
	Niveles de significación	0.01	0.07	0.28	0.66
M3	Lambda de Wilks	0.96	0.94	0.98	0.99
	Estadístico <i>F</i>	11.42	20.73	5.40	3.03
	Niveles de significación	0.00	0.00	0.02	0.08

Fuente: Elaboración propia a partir de los resultados obtenidos en el análisis de la varianza (ANOVA) aplicado a las muestras de fondos de cobertura construidas con la información de la revista *MARHedge Performance & Evaluation Directory* (1999-2006).

3.1 Análisis discriminante

La información relativa al autovalor, al contraste de significación global de igualdad de medias y a la medida de bondad de ajuste aparece en la Tabla 5. El reducido nivel de significación de la *chi*-cuadrada para todos los años de las muestras M1, M2 y M3 implica rechazar la hipótesis de igualdad de medias de los grupos y aceptar que las funciones discriminantes son estadísticamente significativas y se preconiza la diferencia entre los dos grupos. Por tanto, para todas las muestras las variables de forma global ejercen un efecto significativo en la separación de grupos.

Por otra parte, la calidad de la función discriminante está vinculada a la presencia de valores relativamente altos en el autovalor. Los años de la muestra M2 presentan los autovalores más altos (0,65 en el Año 0 y 0,69 en el Año 1) asociados con una elevada correlación canónica (0,63 en el Año 0 y 0,64 en el Año 1), lo que indica que éstas son las mejores funciones discriminantes.

Tabla 5
Autovalor y Lambda de Wilks de cada muestra

<i>MI</i>			
<i>Función</i>	<i>Año 0</i>		
Lambda de Wilks	0.86		
Chi-cuadrada	102.34		
Nivel de significación	0.00		
Autovalor	0.16		
Correlación canónica	0.37		
<i>M2</i>			
<i>Función</i>	<i>Año 0</i>	<i>Año 1</i>	
Lambda de Wilks	0.61	0.59	
Chi-cuadrada	71.85	74.44	
Nivel de significación	0.00	0.00	
Autovalor	0.65	0.69	
Correlación canónica	0.63	0.64	
<i>M3</i>			
<i>Función</i>	<i>Año 0</i>	<i>Año 1</i>	<i>Año 2</i>
Lambda de Wilks	0.89	0.97	0.95
Chi-cuadrada	34.60	10.64	14.76
Nivel de significación	0.00	0.03	0.01
Autovalor	0.12	0.04	0.05
Correlación canónica	0.33	0.18	0.21

Fuente: Elaboración propia a partir de los resultados del análisis discriminante aplicado a las muestras de fondos de cobertura construidas a partir de la información publicada en la revista *MARHedge Performance & Evaluation Directory* (1999-2006).

Tabla 6
Funciones Discriminantes y Centroides de las muestras

<i>Periodo</i>		<i>Función Discriminante</i>					<i>Centroides</i>	
		<i>Constante</i>	<i>Edad</i>	<i>Rentabilidad</i>	<i>Activos gestionados</i>	<i>Flujos de fondos</i>	<i>Activos</i>	<i>Liquidados</i>
M1	Año 0	-3.09	0.32	0.02	0.09	0.00	0.40	-0.40
M2	Año 1	-3.80	0.37	0.01	0.12	0.00	0.82	-0.82
	Año 0	-4.05	0.33	0.01	0.14	0.00	0.80	-0.80
M3	Año 2	-0.88	0.19	0.04	-0.02	0.00	0.22	-0.22
	Año 1	-3.79	0.24	0.04	0.12	0.00	0.19	-0.19
	Año 0	-3.44	0.17	0.07	0.11	0.00	0.34	-0.34

Fuente: Elaboración propia a partir de los resultados del análisis discriminante aplicado a las muestras de fondos de cobertura construidas a partir de la información publicada en la revista *MARHedge Performance & Evaluation Directory* (1999-2006).

La Tabla 6 expone las funciones discriminantes obtenidas y los promedios de las puntuaciones discriminantes, es decir, los centroides. A partir de las puntuaciones discriminantes podemos clasificar a los fondos de cobertura en el grupo de activos o en el de liquidados. Los fondos que obtienen una puntuación discriminante negativa son asignados al grupo de liquidados, mientras que los que obtienen una puntuación positiva son asignados al grupo de activos.²³ Los componentes básicos de las funciones discriminantes nos permitirán el cálculo de las puntuaciones discriminantes relativas a cada función:

$$D_i = u_0 + u_1edad_i + u_2rent_i + u_3actgest_i + u_4flujos_i \quad (7)$$

Donde: $i = 1, 2, \dots, n$

u_0 se refiere al término constante de la función discriminante.

y u_1, u_2, u_3 y u_4 son los coeficientes de las funciones canónicas discriminantes.

Tabla 7
Coefficientes estandarizados de las funciones discriminantes canónicas

	<i>Periodo</i>	<i>Edad</i>	<i>Rentabilidad</i>	<i>Activos gestionados</i>	<i>Flujos de fondos</i>
M1	Año 0	0.92	0.28	0.15	0.28
M2	Año 1	0.98	0.10	0.23	0.21
	Año 0	0.95	0.12	0.27	0.01
M3	Año 2	0.60	0.74	-0.03	-0.30
	Año 1	0.77	0.54	0.20	-0.16
	Año 0	0.57	0.71	0.19	0.26

Fuente: Elaboración propia a partir de los resultados del análisis discriminante aplicado a las muestras de fondos de cobertura construidas a partir de la información publicada en la revista *MARHedge Performance & Evaluation Directory* (1999-2006).

Los coeficientes estandarizados de las funciones discriminantes canónicas tienen la misma interpretación que sus homónimos de la regresión lineal (los betas) y al ser estandarizados evitan el efecto escala (Tabla 7). La magnitud del coeficiente estandarizado –en *valor absoluto*– nos indica la importancia relativa de cada variable en el cálculo de la función discriminante y nos permite determinar que variables

²³ En este proceso de clasificación los centroides de los grupos determinan una “zona gris”, es decir, la zona intermedia entre el centroide del grupo de activos y el centroide del grupo de liquidados. Un fondo de cobertura es asignado al grupo para el que su puntuación discriminante esté más cercano al centroide de dicho grupo.

contribuyen más a la discriminación de la función, cuanto mayor sea el coeficiente mayor será la contribución de su variable asociada.

De modo que, en la muestra M1 después de la edad, la rentabilidad aparece como la variable más importante en la discriminación entre fondos activos y liquidados. Lo mismo ocurre en la muestra M2, en la que la edad tiene un peso relevante en la discriminación de fondos, seguida por los activos gestionados. Mientras que en la muestra M3 la rentabilidad aparece como la más importante en la capacidad discriminante de la función, seguida por la edad.

La Tabla 8 nos presenta la matriz de estructura, la cual expresa en qué medida cada variable contribuye a la función discriminante. Las variables de cada una de las matrices se presentan en orden de importancia dependiendo de su *valor absoluto*. De acuerdo con estas matrices, la *edad* del fondo es la variable que guarda una mayor correlación lineal con la función discriminante en las muestras M1 y M2, mientras que en la muestra M3 la *rentabilidad* –seguida por la *edad*– es la variable que tiene una mayor correlación con la función discriminante.

Tabla 8
Matriz de estructura

<i>M1</i>					
<i>Año 0</i>					
<i>Matriz estructura</i>	<i>Función</i>				
Edad	0.87				
Activos gestionados	0.39				
Rentabilidad	0.32				
Flujos de fondos	0.15				
<i>M2</i>					
<i>Año 0</i>			<i>Año 1</i>		
<i>Matriz estructura</i>	<i>Función</i>	<i>Matriz estructura</i>	<i>Función</i>		
Edad	0.95	Edad	0.94		
Activos gestionados	0.33	Activos gestionados	0.29		
Rentabilidad	0.13	Rentabilidad	0.06		
Flujos de fondos	-0.07	Flujos de fondos	0.03		
<i>M3</i>					
<i>Año 0</i>		<i>Año 1</i>		<i>Año 2</i>	
<i>Matriz estructura</i>	<i>Función</i>	<i>Matriz estructura</i>	<i>Función</i>	<i>Matriz estructura</i>	<i>Función</i>
Edad	0.75	Edad	0.80	Edad	0.70
Activos gestionados	0.56	Activos gestionados	0.55	Activos gestionados	0.66
Rentabilidad	0.38	Rentabilidad	0.33	Rentabilidad	-0.29
Flujos de fondos	0.29	Flujos de fondos	-0.13	Flujos de fondos	0.10

Fuente: Elaboración propia a partir de los resultados del análisis discriminante aplicado a las muestras de fondos de cobertura construidas a partir de la información publicada en la revista *MARHedge Performance & Evaluation Directory* (1999-2006).

Los coeficientes estandarizados (Tabla 7) y las correlaciones de estructura (Tabla 8) son indicadores de la capacidad discriminadora de la función obtenida para cada uno de los años que conforman las muestras. Los signos de ambos tipos de coeficientes no afectan a las clasificaciones, ya que solamente indican el sentido de la relación con la función. A partir de estos indicadores se confirma que la variable con un mayor poder discriminante en las muestras M1 y M2 es la edad, mientras que en la M3 es la rentabilidad.

Para la clasificación de los fondos utilizamos el análisis bayesiano. El teorema de Bayes permite calcular la probabilidad *a posteriori* o probabilidad de que un fondo pertenezca al grupo de activos o al de liquidados dada una puntuación discriminante determinada. La Tabla 9 presenta los resultados de la clasificación original y de la clasificación mediante el procedimiento de validación cruzada. Comparando el grupo de pertenencia real con el grupo de pertenencia asignado podemos comprobar el porcentaje de fondos correctamente clasificados en cada muestra, así como los Errores del Tipo I y del Tipo II.

Tabla 9
Resultados de la clasificación y validación cruzada

Periodo	Original					Validación cruzada					
	% Total (clasificación correcta)	% Liquidados (clasificación correcta)	% Activos (clasificación correcta)	Error tipo I	Error tipo II	% Total (clasificación correcta)	% Liquidados (clasificación correcta)	% Activos (clasificación correcta)	Error tipo I	Error tipo II	
M1	Año 0	63.6%	72.2%	55.1%	27.8%	44.9%	63.5%	72.2%	54.8%	27.8%	45.2%
	Año 1	82.9%	93.2%	72.6%	6.9%	27.4%	80.8%	91.8%	69.9%	8.2%	30.1%
M2	Año 0	81.1%	90.5%	71.6%	9.5%	28.4%	79.7%	90.5%	71.6%	9.5%	28.4%
	Año 2	59.7%	67.9%	51.6%	32.1%	48.4%	58.5%	67.3%	49.7%	32.7%	50.3%
M3	Año 1	56.1%	62.8%	49.4%	37.2%	50.7%	54.8%	62.2%	47.4%	37.8%	52.6%
	Año 0	63.1%	70.7%	55.4%	29.3%	44.6%	62.1%	69.4%	54.8%	30.6%	45.2%

Fuente: Elaboración propia a partir de los resultados del análisis discriminante aplicado a las muestras de fondos de cobertura construidas a partir de la información publicada en la revista *MARHedge Performance & Evaluation Directory* (1999-2006).

En los resultados de la validación cruzada, la muestra M1 presenta un elevado porcentaje de correcta clasificación de los fondos liquidados (72%). Por otra parte, la muestra M2 obtuvo los porcentajes de correcta clasificación más elevados (81% para el año anterior a la liquidación y 80% para el año de liquidación), lo que se corresponde con la mayor calidad de las funciones discriminantes obtenidas para esta muestra. En la muestra M3 -como era de esperarse- conforme nos acercamos al año de liquidación aumenta el porcentaje de correcta clasificación.

De acuerdo con los resultados obtenidos podemos concluir que los modelos discriminantes aplicados a las muestras funcionan razonablemente bien en la correcta clasificación de los fondos, dada su capacidad discriminadora. Por otro lado, hemos comprobado que la variable con mayor potencial discriminante cambia dependiendo de los años para los que se tiene información y del número de fondos dentro de una muestra. Así, el hecho de que un fondo de cobertura siga activo o sea liquidado en las muestras M1 y M2 depende de su edad, es decir, de su consolidación en el mercado. Mientras que en la muestra 3, la continuación o el cese de actividades del fondo depende en mayor medida de su rentabilidad. Por tanto, *la quiebra de los fondos de cobertura está muy influenciada por los factores rentabilidad (M3) y antigüedad del fondo (M1 y M2).*

Conclusiones

La importancia del estudio de la quiebra de los fondos de cobertura radica en las características propias del sector, como el elevado riesgo y apalancamiento que pueden llegar a asumir en la búsqueda de mayores rentabilidades y la opacidad en su información. En un contexto de crisis económica, si un importante número de fondos de cobertura fracasara y se viera insolvente para hacer frente a sus deudas, podría producirse un efecto contagio hacia el resto del sector y hacia otros participantes del sistema financiero, creándose así una fuente potencial de riesgo sistémico.

Por lo anterior, en este trabajo hemos aplicado una de las técnicas más utilizadas en la economía de la empresa para el estudio de la quiebra, nos referimos al análisis discriminante. Para llevar a cabo el análisis empírico hemos construido manualmente un base de datos con la información que fondos de cobertura envían de manera voluntaria a la revista *MARHedge*. A partir de esta información hemos construido 3 muestras de fondos de cobertura (de la M1 a la M3) y las variables propuestas por Chan, Getmansky, Hass y Lo (2005) (edad, rentabilidad, activos gestionados y flujos de fondos) con el objetivo de analizar dichos factores y encontrar al factor interno más influyente en la quiebra de los fondos de cobertura en distintos horizontes de tiempo.

La selección de las variables de clasificación es discutible, ya que otros factores internos como la edad del gestor, las comisiones, el apalancamiento, etc., podrían influir en la quiebra de los fondos de cobertura, desafortunadamente la limitada y escasa transparencia en la información de este sector es uno de los principales obstáculos en su estudio. Sin embargo, los factores seleccionados lograron ser de utilidad en la etapa de clasificación de los fondos de forma estadísticamente significativa.

En el análisis discriminante encontramos una combinación lineal de las variables clasificatorias que dieron origen a una función discriminante, con la que obtuvimos una serie de puntuaciones discriminantes para cada uno de los fondos de las muestras. A partir de estas puntuaciones calculamos la probabilidad de pertenencia de cada fondo al grupo de activos o al grupo de liquidados, y a partir de esa probabilidad se realizó su clasificación de manera satisfactoria.

Por otra parte, la obtención de los coeficientes estandarizados y de los coeficientes de estructura permitió seleccionar a la variable con la mayor capacidad discriminante en cada muestra. A partir de estos coeficientes, encontramos que la consolidación del fondo –medido por la edad– (muestras M1 y M2) y la rentabilidad (muestra M3) son los factores más relevantes en la continuación o cese de las operaciones de los fondos de cobertura. En los resultados de la muestra M3 observamos que los futuros fondos liquidados comienzan a manifestar problemas de rentabilidad desde 2 años antes de su liquidación.

Para complementar nuestros resultados, cabe destacar que Chan, Getmansky, Hass y Lo (2005) encuentran que el pobre desempeño en la obtención de rentabilidad es el indicador más importante en el evento de quiebra dentro de la industria de los fondos de cobertura. Además, proponen algunas medidas de riesgo de iliquidez para esta industria y hacen hincapié en los efectos perjudiciales del apalancamiento y del riesgo asumido por este tipo de fondos (dado el perfil de rentabilidad), los cuales difieren en modos muy importantes de la inversión tradicional.

Entre nuestras recomendaciones finales, al igual que en la mayoría de los trabajos de la revisión de literatura, consideramos que la metodología propuesta debe ser utilizada como una herramienta de ayuda para los reguladores y posibles inversores y nunca como un procedimiento exclusivo para “predecir” la liquidación de una entidad. Los reguladores internacionales necesitan crear sistemas de prevención de quiebras especialmente diseñados para instrumentos de inversión alternativa como los fondos de cobertura, con la finalidad de efectivamente anticipar y/o mitigar las consecuencias que el colapso de este tipo de fondos traería al sistema financiero debido a la creciente importancia que han experimentado en los últimos años.

Finalmente, conviene mencionar que el papel de los fondos de cobertura en México es mínimo debido a las características propias de esta industria y a la normativa vigente del mercado financiero local que restringe libertad a los gestores en el uso de derivados, ventas en corto y elevado apalancamiento, acciones que son clave en la obtención de altas rentabilidades. Además, los fondos de cobertura son instrumentos de inversión alternativa dirigidos a grandes patrimonios e inversores institucionales, quienes poseen portafolios diversificados y conocimientos especializados en la gestión de riesgos, por lo que no están disponibles para cualquier inversor.

Bibliografía

- Ackermann, C., McEnally, R., y Ravenscraft, D. (1999). "The performance of hedge funds: Risk, Return and Incentives", *The Journal of Finance*, 54, pp. 833-874.
- Altman, E. I. (1968). "Financial Ratios, Discriminant Analysis and Prediction of Corporate Bankruptcy", *The Journal of Finance*, Vol. XXIII, Núm. 4, pp. 589-609.
- Altman, E. I. y Narayanan, P. (1997). "An International Survey of Business Failure Classification Moldes", *Financial Markets, Institutions & Instruments*, Vol.6, Núm.2, pp.1-57.
- Baquero, G., Horst, J. y Verbeek, M. (2005). "Survival, Look-Ahead Bias and the Performance of Hedge Funds", *Journal of Financial and Quantitative Analysis*, 40, pp. 493-517.
- Barry, R. (2003). "Hedge Funds: A Walk through the Graveyard", *Working Paper No. 25*, Applied Finance Centre, Macquarie University, marzo.
- Beaver, W. H. (1966). "Financial Ratios as Predictor of Failure", *Journal of Accounting Research*, Vol. 4, Empirical Research in Accounting: Selected Studies, pp. 71-111.
- Beaver, W. H. (1968). "Alternative Accounting Measures As Predictors of Failure", *The Accounting Review*, January, pp. 113-122.
- Beaver, W. H., Correia, M. y McNichols, M. (2011). *Financial Statement Analysis and the Prediction of Financial Distress. Foundations and Trends in Accounting*, Ed. Now Publishers Inc.
- Blum, M. (1974). "Failing Company Discriminant Analysis", *Journal of Accounting Research*, 12 (spring) pp.1-25.
- Brealey, R. A. y Kaplains, E. (2001). "Hedge Funds and Financial Stability: An Analysis of their Factor Exposures", *International Finance*, Núm. 4, Vol. 2, pp. 161-187.
- Brown, S. J., Goetzmann, W. N. y Park, J. (1997). "Conditions for survival: changing risk and the performance of hedge funds manager and CTAs", Yale School of Management, *Working Paper No. F-59*.
- Brown, S., Goetzmann, W. y R. Ibbotson, (1999). "Offshore Hedge Funds: Survival and Performance 1989-1995", *Journal of Business*, Vol. 72, Núm 1, pp. 91-118.
- Caslin, J. J. (2004). *Hedge Funds*, Institute of Actuaries and Faculty of Actuaries, Ed. Bell & Bain Ltd., Escocia.
- Chan, N., Getmansky, M., Hass, S.M. y Lo, A.W., (2005). "Systemic risk and hedge funds", *Working Paper*, No.11200, National Bureau of Economic Research, March.

- Collins, R. A. (1980). "An Empirical Comparison of Bankruptcy Prediction Models", *Financial Management*, vol. 9 (2), summer, pp. 52-57.
- De la Garza, J., Morales, B.N. y González, B.A. (2013). *Análisis Estadístico Multivariante: Un enfoque teórico y práctico*, 1ra. ed., Ed. McGraw Hill, México.
- Deakin, E. B. (1972). "A Discriminant Analysis of Predictors of Business Failure", *Journal of Accounting Research*, Vol. 10, Núm. 1, pp. 167-179.
- Fung, W. y Hsieh, D. (1997). "Empirical Characteristics of Dynamic Trading Strategies: The Case of Hedge Funds", *The Review of Financial Studies*, Vol. 10, No. 2, pp. 275-302, summer.
- Fung, W. y Hsieh, D. (2000). "Performance Characteristics of Hedge Funds and Commodity Funds: Natural versus Spurious Biases", *Journal of Financial and Quantitative Analysis*, Vol. 35, No.3, September, pp. 291-307.
- Getmansky, M., Lo, A. y Mei, S. (2004). "Sifting Through the Wreckage: Lessons from Recent Hedge-Fund Liquidations", *Journal of Investment Management*, Vol. 2 (4), pp. 6-38.
- Gregoriou, G. (2002). "Hedge Fund Survival Lifetimes", *Journal of Asset Management*, Vol. 3 (3), pp. 237-252.
- Hair, J. F., Anderson, R. E., Tatham, R. L. y Black, W. C. (1999). *Análisis Multivariante*, 5ª ed., Ed. Prentice Hall Iberia, Madrid.
- El Hennawy, R. H. A. y Morris, R. C. (1983). "The Significance of Base Year in Developing Failure Prediction Models", *Journal of Business Finance and Accounting*, Vol. 10, No. 2, pp. 209-223.
- Laffarga, J., Martín, J. L. y Vázquez, M. J. (1991). "La predicción de la quiebra bancaria: el caso español", *Revista Española de Financiación y Contabilidad*, Vol. XXI, Núm.66, pp.151-166.
- Lévy-Mangin, J-P. y Varela, J. (2003). *Análisis Multivariable para las Ciencias Sociales*, Ed. Pearson Prentice Hall, Madrid.
- Liang, B. (2001). "Hedge Fund Performance: 1990-1999", *Financial Analysts Journal*, 57, pp. 11-18.
- López, D., Moreno, J., y Rodríguez, P. (1994). "Modelos de previsión del fracaso empresarial: Aplicación a entidades de seguros en España", *Esic-Market*, 84, pp. 83-125, abril-junio.
- MARHedge Institutional Investors, *MARHedge Performance & Evaluation Directory*, 1999-2006.
- Meyer, P. A. y Pifer, H. W. (1970). "Prediction of Bank Failures", *The Journal of Finance*, Vol. 25, No. 4, pp. 853-868.

- Peel, M. J. y Peel, D. A.(1987). “Some further empirical evidence on Predicting Private Company Failure”, *Accounting and Business Research*, Vol. 18, No. 69, pp. 57-66.
- Rasjwade, A. V. (2007). “Hedge Funds”, *Economic&Political Weekly*, Vol. 42, No. 13, pp.1147-1149.
- Taffler, R. J. (1982). “Forecasting Company Failure in the UK Using Discriminant Analysis and Financial Ratio Data”, *Journal of the Royal Statistical Society. Series A (General)*, Vol. 145, No. 3, pp. 342-358.
- Yam T. K. y Kiang M. Y. (1992). “Managerial applications of neural networks: the case of bank failure predictions”, *Management Science*, Vol. 38, Num. 7, pp. 926-947.
- Yim, J. y Mitchell, H. (2005). “A comparison of corporate distress prediction models in Brazil: hybrid neural networks, logit models and discriminant analysis”, *Nova Economia Belo Horizonte*, Vol.14, No. 1, pp. 73-93.